# 1AC

## 1AC---Technology

### 1AC---Innovation ADV

#### Contention 1 is Innovation.

#### The unmoderated, permissive nature of present merger policy sanctions Big Tech to engage in predatory behavior that disrupts Schumpeterian competition.

Michael L. Katz 21, Sarin Chair in Strategy and Leadership at the University of California, Berkeley's Haas School of Business. Professor, Economics, US Berkeley, "Big Tech Mergers: Innovation, Competition for the Market, and the Acquisition of Emerging Competitors," Information Economics and Policy, Vol. 54, March 2021, ScienceDirect. error edited.

1. Introduction

Commentators across the political spectrum have called for new and more forceful approaches to antitrust enforcement with respect to Big Tech—especially Amazon, Facebook, and Google. A wide range of proposals have been put forth, including breaking firms up, subjecting them to pervasive regulation, and being much more wary of mergers between incumbents and either recent or potential entrants.1 The last concern is driven, in part, by the combination of apparently dominant market positions coupled with large numbers of acquisitions. For example, Amazon is estimated to have accounted for almost 40 percent of all 2019 ecommerce sales in the U.S., and Facebook and Google together account for over 60 percent of U.S. digital ad spending.2 These companies made hundreds of acquisitions in the previous decade.3

To determine if acquisitions by Big Tech are a big problem that justifies heightened or different scrutiny, it is useful to begin by asking: what is special about these firms, beyond their tremendous success? One answer is that all are in industries with very strong increasing returns and positive feedback loops. There are several sources of increasing returns—often present simultaneously. One source is network effects. Another is the collection and use of big data, which can give rise to economies of scale, scope, and experience. Lastly, the creation of software and intellectual property (including proprietary hardware designs) is typically characterized by large economies of scale, with high fixed costs and very low marginal costs.

The presence of such strong increasing returns can limit the number of viable competitors and even create a tendency to tip toward monopoly. When increasing returns are large relative to product differentiation, competition may be for the market rather than in the market. Competition for the market, whereby firms compete by innovating to attain temporary market dominance, is often referred to as Schumpeterian competition.

The need for scale can make entry difficult in the very markets in which entry is critical because competition takes place for the market. In a market in which a large base of users is essential to a firm’s ability to offer an attractive value proposition—say because network effects are strong and/or the value of user big data is high—the only economically viable means of entry may be to build up a base of users in an adjacent market and then provide the new service to that base of users—what is sometimes called a two-stage entry strategy.4 Some commentators believe that Instagram and WhatsApp would have used two-stage entry to become strong competitors to Facebook in social networking if that firm had not acquired them in 2012 and 2014, respectively.5

Another strategy, to use alone or in conjunction with two-stage entry, is to offer a product with higher quality than that of the incumbent firm. Most directly, an entrant’s quality advantage may outweigh its scale and installed-base disadvantages. There can also be an indirect effect. In markets subject to strong network effects, a firm may gain significant competitive advantage from favorable consumer expectations—when consumers expect that firm to have high sales, their expected value of patronizing that firm rises due to the anticipated network benefits associated with a larger user base. One might expect incumbents generally to have expectations advantages. However, a highly visible innovation might tip expectations in favor of an entrant.6 \*\*\*FOOTNOTE BEGINS\*\*\* Farrell and Katz (1998) provide an early analysis of the role innovation can play to facilitate entry by shifting consumer expectations in markets with network effects. \*\*\*FOOTNOTE ENDS\*\*\* If this pattern prevails, then leapfrog innovation might allow an entrant that is small today to generate a positive feedback cycle and overcome an incumbent’s various scale advantages.

Both two-stage and innovative entry strategies require the entrant to amass complementary resources (e.g., users and intellectual property) to have a chance of overcoming the advantages of incumbency. An entrant also needs a strong growth trajectory that allows it to achieve viable scale. An entrant’s need to acquire complementary assets and attain a promising growth trajectory may allow an incumbent to identify potential rivals before they become major actual competitors.

Below, I explore the role of merger policy in a model in which competition is for the market and an incumbent can identify—and merge with—an emerging or potential competitor before the entrant becomes the new dominant firm. After Section 2 briefly reviews related literature, Section 3 presents the overall analytical framework: a discrete-time, infinite-horizon game in which each period a new potential entry opportunity arises with exogenous probability. Entry requires making a sunk investment in product development. If entry occurs and the firms do not merge, then the incumbent and entrant compete for the market, and the market is ultimately monopolized as one of the firms is driven to exit.

Section 4 argues that incumbents’ acquisitions of emerging or potential competitors should be subject to heightened antitrust scrutiny when competition is for the market. Entry is a critically important means of promoting market performance under Schumpeterian competition but acquisition of a nascent competitor can be an especially effective way to avoid Schumpeterian competition, to the detriment of consumers. Marino and Zábojník (2006) have shown that, when competition is in the market (i.e., multiple incumbents can be profitable simultaneously), the threat of rapid entry can sometimes serve as a substitute for merger policy by making merger unprofitable absent efficiencies. To see why, suppose that there are two incumbents. In the absence of additional entry, merging to monopoly raises profits by eliminating product-market competition. However, there can be an offsetting share-dilution effect: when the two firms merge, they may create room for a subsequent firm profitably to enter the market because they weaken their bargaining position—instead of collectively receiving two-thirds of the continuation profits in symmetric bargaining over merger with the entrant, they receive only half. When subsequent entry is rapid, the share-dilution effect renders mergers unprofitable absent efficiencies. By contrast, there is no share-dilution effect when competition is for the market: even absent merger, the market is monopolized after a period of competing for dominance following entry. Because it eliminates a period of competition to be the dominant firm, merger is profitable even if ensuing entry is rapid.

Sections 5-8 examine the effects of merger policy on potential entrants’ pre-merger innovation incentives. When mergers are banned, entry occurs only when a potential entrant has a sufficiently valuable innovation that it can overcome any incumbency advantages and become the new dominant firm. As Rasmussen (1988) identified, when mergers are feasible, a firm may enter the market solely to induce the incumbent to purchase the entrant in order to avoid dissipating product-market profits through competition—so-called entry for buyout.7 \*\*\*FOOTNOTE BEGINS\*\*\* Entry for buyout can occur even if the entrant could otherwise become dominant—competition to attain that position could prevent the firm from recovering its entry costs. \*\*\*FOOTNOTE ENDS\*\*\* It is often argued that a benefit of allowing mergers is that, by facilitating entry for buyout, they can promote innovative entry.8 Below, I show that allowing mergers can also discourage entrant innovation.

As long as the incumbent’s actions conditional on not merging are independent of whether mergers are permitted, the option to merge can only increase an entrant’s profits. This finding, however, does not imply that this option everywhere increases an entrant’s incentives to marginally improve its product. In fact, the possibility of merger may diminish innovation by increasing the relative profitability of entering with a less ambitious product. Moreover, the option to merge might reduce an entrant’s profits by facilitating incumbency for buyout, whereby an incumbent invests in strengthening its competitive position solely to induce the entrant to merge on more favorable terms. Finally, a permissive policy that also applies to future mergers can reduce a current entrant’s profits by facilitating entry for buyout by subsequent entrants.

Section 5 analyzes the effects of merger policy on an entrant’s incentives to invest in marginal product improvements. When there is no incumbency for buyout and certain other conditions are satisfied, an entrant’s disagreement profits when bargaining over whether to merge equal its profits when mergers are banned. Hence, the difference between the entrant’s profits with and without the possibility of merger equals the entrant’s share of the gains from merger. The effect of permissive merger policy on innovation incentives thus depends on how a change in the entrant’s product quality affects those gains. In general, increasing the entrant’s product quality can raise or lower the gains from merger, which implies that option to merge can raise or lower an entrant’s incentives to invest in quality.

The analysis summarized above relies on reduced-form profit functions. Sections 6-8 present examples in which product-market competition is explicitly modeled. These examples illustrate how permissive merger policy can harm innovation incentives and that the competitive effects of mergers can be complex and highly fact specific. The example of Section 6 exhibits same-side network effects, such as arise with social networks. Among other things, this example demonstrates that, in situations where the merged firm retains the older technology (to avoid splitting users and losing network benefits), an increase in the entrant’s product quality lowers the gains to merger. Intuitively, the higher is the entrant’s product quality, the greater the opportunity cost of not using it. The example of Section 7 considers a market without network effects in which a firm must make a new investment each period in order to be an active seller. In this example, the entrant chooses a higher product quality when mergers are banned because the possibility of merger leads the entrant to put weight on the effects of its innovation on monopoly profits, which are less sensitive to the entrant’s quality than are the profits the entrant earns when competing with the incumbent. Section 8 shows by example that permissive merger policy can facilitate incumbency for buyout that discourages innovative entry.

Section 9 discusses some policy implications of this analysis. Although acquisitions of emerging competitors in Schumpeterian markets should be subject to heightened scrutiny generally, in some cases merger promote innovation and efficiency. It is thus important to examine the facts of each case at hand even if doing so is difficult.

Finally, a technical appendix examines a benchmark case of imitative entry to compete in the market, which serves as a benchmark for the text’s model of innovative entry to compete for the market.

2. Related literature

Before turning to the analysis, it is useful to put it in context and discuss related literature. The model presented below is not intended to apply to all Big Tech mergers.9 It is highly unlikely that the Big Tech incumbents considered each of the hundreds of firms they acquired to be a significant potential competitor, and any given incumbent may have a variety of motives underlying different acquisitions. For example, some commentators are concerned that a digital platform can expand into various industries that rely on it and then create artificial advantages for its subsidiaries while harming competition from rivals to those subsidiaries. Variants of this complaint have been made against Amazon Basics, Google Shopping, and Apple apps sold on its App Store.10 My focus is on mergers aimed at preventing successful entry into a firm’s core market(s), rather than mergers that allow a firm to expand into other markets. The latter type of merger is more appropriately analyzed using models developed in the literature on vertical mergers. Church (2008) surveys the main theories relevant to vertical mergers, and Slade (2020) surveys empirical studies of the effects of vertical mergers.

Big Tech firms often operate multi-sided platforms, which can have strong implications for the welfare effects of a merger. Recent analyses include Anderson and Peitz (2020) and Correia-da-Silva et al. (2019), and Foros et al. (2015) and Jullien and Sand-Zantman (2020) offer surveys of the two-sided merger literature. Below, I abstract away from multi-sidedness and focus on the implications of strong increasing returns.

There has been a long and inconclusive debate regarding the general relationship between horizontal mergers and innovation. See, for example, Baker (2007), Jullien and Lefouili (2018), Katz and Shelanski (2007), and Shapiro (2012), and the papers cited therein. This literature tends to be concerned with the effects of merger on the level of post-merger innovation. By contrast, I consider situations in which innovation is necessary to launch an entry attempt, and I focus on the effects that the prospect of merger has on premerger innovation incentives.

Several recent papers also consider premerger incentives.11 Cabral (2020), Letina et al. (2020), Motta and Peitz (2020), and Hollenbeck (2020) all find that prohibiting mergers can reduce innovation by reducing entry for buyout. The first three papers all examine models in which additional investment increases the probability that an innovation project is successful but does not affect the outcome conditional on success.12 Hence, the possibility that merger discourages entrant investment in marginal quality improvements does not arise. Hollenbeck (2020) considers a computational oligopoly model in which entrants’ investment decisions do affect their marginal product qualities, but for the functional forms that he utilizes, he finds that the prospect of merger never decreases entrants’ innovation levels.

Kamepalli et al. (2020) also examine the effects of merger policy on premerger innovation. They find that permissive merger policy can reduce entrant innovation but for a reason very different than the ones I identify below. Specifically, Kamepalli et al. examine a market subject to network effects in which prospective early adopters shy away from an entrant that they believe will be merged out of existence.

Finally, Bryan and Hovenkamp (2020) examine a model of startup acquisitions by duopoly incumbents in which a startup does not have the ability to enter the market. Although related, the analysis focuses on the effects of exclusive licensing rather than the elimination of a potential Schumpeterian competitor that I examine.

3. The model

I explore the effects of merger policy by examining Markov perfect equilibria (i.e., subgame perfect equilibria in which players’ strategies depend only on the current, payoff-relevant state variables) in the infinite-horizon game illustrated in Fig. 1 below. Firms and consumers have a common per-period discount factor, δ. The structure of the game and specific parameter values are common knowledge.

At the start of each period, any existing incumbent chooses whether to remain active in the market. Then a new potential entry opportunity arrives with exogenous probability, ρ¯. An opportunity can be thought of as arising due to the discovery of a new generation of product or process technology that is proprietary to the potential entrant. When a new generation arrives, the potential entrant associated with that generation chooses how much to invest in developing the new technology, where greater investment leads to higher quality.13

Let qk(I) denote the quality level obtained for the generation-k technology when the entrant invests I. qk(I) is assumed to be increasing in k as well as I. To simplify the notation, equate each technological generation with the corresponding period index. That is, generation k refers to the technology that allows entry commencing in period k. This notation implicitly assumes that generations advance over time independently of both whether an entry opportunity arises in a particular period and what amounts that entrants have invested in past generations. Once a given generation of technology has arrived, it does not improve (i.e., for a given k, qk( · ) does not change with the passage of time), and there is only a single entry opportunity for any given firm: a firm must either choose to enter when an opportunity arises or otherwise stay out forever.

Conditional on having an entry opportunity and choosing to enter, a generation-l entrant facing an incumbent with product quality qf chooses investment I to maximize

[Equation omitted]

where VE l (qf, ql) is the expected discounted productmarket profits that will be earned by a generation-l entrant with quality ql facing an incumbent with quality qf. 14 One can think of f as standing for “follower” and l standing for “leader” because the entrant relies on the leading technology. Assume that that there is a unique optimal investment level conditional on the incumbent’s quality, I E l (qf), with associated quality level qE l (qf) ≡ ql(I E l (qf)). The potential entrant chooses to enter the market if

[Equation omitted]

and will stay out of the market otherwise (with I = 0). Let ρE l (qf) denote ρ¯ times the conditional probability that entrant with generation-l technology will enter if the opportunity arises and there is a single incumbent having quality qf.

After the entry decision, any incumbents and the entrant (if there is one) choose whether to merge, which generates a new state of the market. A merger is assumed to take place whenever doing so maximizes the net present value of the merging parties’ joint profits. The exit, entry, and merger decisions occur sequentially but instantaneously, followed by a production period. The next period repeats this structure.

I consider settings in which, if entry occurs and the incumbent and entrant do not merge, then they compete for the market for one period, after which one of them exits the market. The losing firm puts up a fight for one period before exiting because it has a depreciating asset (e.g., consumer brand, installed user base, or plant and equipment) in which the firm stops investing after entry but that has sufficient residual value for the firm to constrain the winner’s pricing for one period.15 Once the loser has exited the market, it exerts no competitive pressure.16

[Figure omitted]

Let πI t(qf, ql) and πE t (qf, ql) denote the per-period profits of the incumbent and entrant, respectively, when they compete for the market in period t with products having qualities qf and ql, respectively.17 These profit functions may also reflect horizontal product differentiation as long as that differentiation is insufficient to allow the firms to coexist profitably.18 After one firm has exited the market, the remaining firm earns π M t (qk ) per period until the next entry event occurs, where k ∈ {f, l} denotes the winner’s technology

Assume that maxI πE t (qf, ql(I)) − I < 0, so that it is unprofitable to enter the market, lose the competition for the market, and exit. Given this assumption, and because the model is deterministic with respect to product-market competition, entry occurs only if the entrant anticipates either winning the competition for the market or merging.19

If entry occurs and the incumbent and entrant merge, then the combined firm earns πT t (qf, ql) in the entry— or “transition”—period given the incumbent’s and entrant’s product qualities. The transition period reflects lags in redeploying competitive assets (e.g., brands, installed bases, or physical plant). The merged firm earns π M t (qk ) per period in subsequent periods until the next entry event occurs, where k ∈ {f, l} is the technology that the merged firm chooses to adopt going forward. The assumption that the merged firm has to choose one or the other technology is consistent with my focus on effects that arise when mergers do not generate productive efficiencies.

Even absent merger efficiencies, there is private incentive to merge to avoid dissipating profits through competition: both πT t (qf, ql) and max k∈{f,l} π M t (qk ) are assumed to be larger than πI t(qf , ql) + πE t (qf , ql).

4. Acquisitions of emerging competitors: a case for heightened scrutiny

As described in the Introduction, to be successful, an entrant needs to amass complementary resources (e.g., users and intellectual property) to overcome any incumbency advantages, and it needs a strong growth trajectory to achieve viable scale. The entrant’s need to acquire complementary assets and attain a strong growth path may allow an incumbent to identify and acquire a potential rival before it has entered into direct competition with the incumbent, or while the entrant still has a very small share of the market in which the incumbent competes.20 This possibility challenges the traditional antitrust approach to assessing mergers in the U.S., under which it is difficult to prevail in court when trying to block a merger based solely on the loss of potential future competition.21 \*\*\*FOOTNOTE BEGINS\*\*\* See, e.g., Bush and Massa (2004) and Werden and Limarzi (2010) for discussions of relevant cases. In addition, the lack of assets or sales may result in a transaction’s falling below the thresholds that trigger mandatory notification of antitrust authorities. (See Wollman (2019).) \*\*\*FOOTNOTE ENDS\*\*\* Moreover, the earlier the dominant firm can identify such a rival, the more problematical the outcome is for competition policy—both because the earlier a merger involving a nascent or potential competitor occurs the harder it is to challenge and because such a merger allows the firms more fully to avoid competing.

Although the conditions just described make it particularly difficult for antitrust enforcers to challenge acquisitions, these conditions also make it particularly important that enforcers do so: absent entry and dynamic competition, there will be little competition at all. Moreover, when competition is in the market, the share-dilution effect described in the Introduction and illustrated by Proposition A.2 in the Appendix can serve as a substitute for antitrust enforcement in limiting inefficient mergers.22 By contrast, when competition is for the market, the share-dilution effect does not arise.

The intuition for why the share-dilution effect doesn’t limit mergers when competition is for the market is as follows. Suppose that, in the event of merger, the merged entity adopts the entrant’s technology. If entry occurs, then after one period, the continuation game looks the same whether or not the entrant and incumbent have merged— either way there will be a single firm, which relies on the entrant’s technology. From the perspective of the current incumbent and most recent entrant, the only effect of their merger is to avoid a period of competition to be the dominant firm. If the merged firm would adopt the incumbent’s technology, then merger must be even more profitable.

To see this point formally, suppose that mergers are allowed and that the current entrant comes into the market in period l with a quality advantage sufficient for it to win the market if the firms do not merge. Absent merger, the incumbent earns

πI l (qf, ql) (1)

for one period and then exits the market, while (gross of the entry cost) the entrant earns

[Equation omitted]

where ρE l (·) ≡ 0 and VI l+j (ql, ql+j) denotes the continuation value that a firm with quality ql earns as the new incumbent when the next entrant comes into the market in period l + j with quality ql+j. πE l (qf, ql) in expression (2) equals the profits the entrant earns in the entry period. The terms involving π M l+j (ql) are the discounted profits the current entrant earns in later periods before subsequent entry occurs.

If the current incumbent and entrant merge and adopt the entrant’s technology, then they collectively earn

[Equation omitted]

Denote the gains from merger by Gl(qf, ql). Comparing the sum of expressions (1) and (2) with expression (3), the only difference is in the first period following entry, when merger allows the firms to avoid competing. Thus, if the merged firm adopts the entrant’s technology

[Equation omitted]

If the merged firm would find it more profitable to retain the incumbent’s technology, then Gl(qf, ql) would be larger than the right-hand side of eq. (4). A similar argument applies to cases in which the incumbent would prevail absent merger. Hence,

Proposition 1. When competition is for the market, it is profitable to merge regardless of the rate at which subsequent entrants arrive or firms discount future profits.

The finding that merger is profitable when competition is for the market is somewhat more general than Proposition 1. The key factors are that: (a) until entry next occurs, the incumbent and current entrant earn greater profits if they are merged than if they are independent competitors, and (b) their joint continuation value when entry next occurs is not lowered by their merger.23 The example presented in Section 7 below demonstrates that, as long as the incumbent would place no competitive pressure on the subsequent—as opposed to current—entrant, condition (b) is satisfied even if a non-merging incumbent would remain active in the market until the next entry event occurs.

5. Merger policy and product development

Now, consider the effects of merger policy on a potential entrant’s investment incentives. When entry is profitable if and only if it leads to merger, the ability to merge must increase the entrant’s investment incentives.24 Similarly, when entry is profitable if and only if mergers are prohibited, permissive merger policy must decrease the entrant’s investment incentives.

To further examine the effects of merger policy on marginal investment incentives, consider a firm that finds entry profitable regardless of whether or not it subsequently merges. Recall that entry is assumed to be profitable absent merger only if the entrant anticipates winning the market. Hence, when mergers are prohibited and entry occurs

[Equation omitted]

Using Eq. (5), any firm other than the first one to enter the industry earns

[Equation omitted]

Next, suppose mergers are allowed and that the incumbent always gets share σ of the gains from merging with the entrant. Then

[Equation omitted]

In general, the functions ρE t (·) and qE t (·) can vary across the two merger-policy regimes, which makes comparison of the entrant’s profits in the two regimes difficult. Assume that, for any given value of ql, terms of the form ρE l+j (ql) and πI l+j (ql, qE l+j (ql)) are identical under the two regimes.25 Under this assumption, the difference in the entrant’s objective function due to a merger—Eq. (8) minus Eq. (6)— is:

[Equation omitted]

l(ql) equals the expected net present value of the sum of the entrant’s share of the gains from merger with the incumbent, (1 − σ )Gl(qf, ql), and its share of what would have been the gains from merger with a subsequent entrant had it not merged with the original incumbent, σ Gl+j(ql, qE l+j (ql)). The latter term arises because it is a component of the entrant’s disagreement payoff with respect to the initial merger with the incumbent.

The effect of merger on the entrant’s investment incentives depends on the derivative of l(ql), which in turn depends in part on how Gl(qf, ql) and Gl+j(ql, ql+j) vary with ql. In general, the gains from merger are non-monotonic in the quality levels. For example, there are no gains from merger if: (a) the entrant’s product quality is so low that the entrant places no competitive pressure on the incumbent, or (b) the entrant’s product quality is so high that the incumbent places no competitive pressure on the entrant and the merged firm would not choose to adopt the incumbent’s technology.26 But there are gains from merger when the two firms have quality levels sufficiently close to one another that each firm would put competitive pressure on the other. This fact strongly suggests that the effects of merger on marginal innovation incentives are ambiguous. The derivative of l(ql) also depends on how the entrant’s choice of quality affects decisions by later potential entrants with respect to entry probabilities and investments in quality (i.e., terms of the form ρE l+j (ql) and qE l+j (ql)). Models that consider a single merger in isolation and only an all-or-nothing innovation project miss these effects.

Given these complexities, it is useful to look at explicit examples of product-market competition in order to identify possibilities and say more about welfare effects.

6. An example with network effects

The first example is of a market subject to network effects in which competing networks are incompatible. In keeping with the assumption that there are no production efficiencies from merger, assume that, even if an incumbent and an entrant merge, the incumbent’s existing users cannot be migrated to the entrant’s network and the incumbent’s existing network cannot be upgraded to the entrant’s quality. A merged firm must choose between adopting the entrant’s technology and stranding existing users, or retaining the incumbent’s network and not using the entrant’s technology.

Each consumer lives for two periods, and any given consumer purchases at most one subscription to a network. Each period, a unit mass of consumers enters the market, and each consumer chooses whether to subscribe to a network for two periods. I assume that a consumer cannot choose to wait one period before making a purchase. Hence, at any given time, half of the consumers in the market are choosing a network.

A consumer enjoys per-period gross consumption benefits of qβ(m) − p, where q is the chosen network’s quality, m is the network’s size or “member” base, and β(m) is the network benefit function, which is increasing with β(0) = 0.

When choosing a network, a consumer must forecast the sales of any currently available network in each period of his or her life. Consumers are forward looking and account for the possibility of entry in the second period of their lives. I assume that, if there are two firms actively selling output, consumer expectations “track quality” in that consumers expect the higher-quality network to win all of the sales in any given period (if qualities are equal, consumers expect the incumbent to win). This expectations process is very favorable to innovative entry because any degree of product superiority neutralizes the incumbent’s installed base advantage.27

6.1. Product-market equilibrium

Consider the product-market equilibrium when an entrant with quality ql competes against an incumbent with quality qf, where ql > qf. 28 In the entry period, the incumbent has an installed base equal to 1—the previous cohort of consumers who are in the second year of their lives. However, because consumer expectations track quality, consumers currently making purchases do not expect the incumbent to make any future sales.29 The incumbent thus offers a new consumer expected gross consumption benefits equal to qfβ(1).

The entrant has no installed base but has a projected base equal to 1 in the entry period due to its superior quality and the consumer expectations process. Consumers also form projections of the entrant’s user base in the next period. To simplify the analysis, assume that each subsequent generation of technological opportunity is a sufficient improvement over earlier ones that entry always occurs when an opportunity arises: ρE j (ql) = ρ¯ for all j ≥ l. 30

Suppose mergers are prohibited. If there is no entry next period, then the current entrant will have a user base equal to 2. If there is entry next period, then the current entrant is expected to have a future user base equal to 1 because the next entrant will have a superior product. Hence, the current entrant offers a new consumer expected gross consumption benefits equal to ql{β(1) + δβˆ}, where βˆ ≡ ρβ¯ (1) + (1 − ρ¯)β(2).

Consumers are assumed to recognize that, when mergers are allowed, users will not be stranded if entry occurs next period but the entrant merges with the incumbent and the merged firm uses the incumbent’s technology. Let γ t(q) denote the probability that, conditional on entry in period t and the incumbent’s having quality q, the resulting merged firm will use the incumbent’s technology. When mergers are permitted, a firm with quality q offers a new consumer expected gross consumption benefits equal to

[Equation omitted]

Note that βˆ t(q) ≥ βˆ, with strict inequality whenever γ t(q) > 0. In equilibrium, consumer expectations are fulfilled.

Now consider consumers’ purchase decisions. If mergers are allowed but the current entrant and incumbent do not merge, then consumers in the entry-period cohort choose the entrant’s product if and only if

[Equation omitted]

where pj is the price charged by a network that offers the generation-j product.31 \*\*\*FOOTNOTE BEGINS\*\*\* Firms either merge immediately or never merge; there would be no gains from delay. Whenever entry next occurred, at least one of the current entrant and incumbent would have an installed base of 0 and thus be competitively irrelevant to the next entrant. Hence, the firm without a base would have no effect on subsequent merger bargaining. \*\*\*FOOTNOTE ENDS\*\*\* The incumbent would be willing to charge a price below 0 only if it could later charge a positive price. However, if a negative price is needed to win in this period (when the incumbent has an installed base of users), then a negative price would also be needed to win in every subsequent period, when the incumbent would face either the current entrant or a subsequent entrant having an even-higher-quality product. Hence, the incumbent will price as low as 0, but no lower, to win sales. By inequality (10), the entrant can profitably win sales even if pf = 0. Therefore, in equilibrium, the entrant wins all of the sales in the period in which it enters, with

[Equation omitted]

Because the two cohorts of users are split across the two networks, the resulting gross consumption benefits are equal to {ql + qf}β(1).

Next consider the periods that follow period l but are prior to the period in which entry next occurs. The previous incumbent (generation f) has an installed base of 0 and consumers do not expect the firm to make future sales. Given β(0) = 0 and its unwillingness to price below zero, the previous incumbent offers consumers no surplus and, thus, places no constraint on the entrant’s pricing beyond that already imposed by consumers’ option to purchase nothing. When selling in period t, the most recent entrant has expected network benefits in the next period equal to βˆ t+1(ql). Consumer cohorts making purchase decisions in these periods thus choose the recent entrant’s product if and only if ql{β(2) + δβˆ t+1(ql)} − pl ≥ 0, and the entrant makes sales at pl = ql{β(2) + δβˆ t+1(ql)}. Because all users are on the generation-l network, gross consumption benefits are equal to 2qlβ(2) per period.

If mergers are banned, a similar analysis applies but with terms of the form βˆ t(ql) replaced by βˆ. In summary,

Proposition 2.

(i) Suppose that mergers are permitted, but the current entrant and incumbent do not merge. Then the continuation equilibrium has the following form:

a) in the entry period, the entrant wins all of the sales at a price of (ql − qf)β(1) + δqlβˆ l+1(ql) and gross consumption benefits are equal to {ql + qf}β(1); and

b) in subsequent periods prior to the next entry event, the generation-l entrant wins all of the sales at a price equal to ql{β(2) + δβˆ t+1(ql)} and gross consumption benefits are equal to 2qlβ(2) per period.

(ii) If mergers are prohibited, then the same results hold with βˆ replacing βˆ l+1(ql).

Now suppose that mergers are permitted and the current entrant and incumbent merge. Recall that the incumbent’s existing users cannot enjoy the benefits of the entrant’s higher-quality technology. The merged firm must choose whether to: (1) maintain two separate networks for one period and shut down the network based on the older technology at the end of the transition period,32 or (2) have new cohorts of consumers join the incumbent’s network even though the entrant’s network could offer higher quality.

If the entrant and incumbent merge, then there is no competition until the next entry event occurs. Thus, until then, the merged firm sets its prices to fully appropriate the expected consumption benefits. Straightforward calculations establish:

Proposition 3. Suppose mergers are permitted and the current entrant and incumbent merge.

(i) If the merged firm operates both networks during the entry (or transition) period and shuts down the network utilizing the older technology at the end of that period, then in the continuation equilibrium:

a) in the entry period, the merged firm makes sales on the new network at a price of ql{β(1) + δβˆ l+1(ql)} and gross consumption benefits are {ql + qf}β(1); and

b) in the subsequent periods prior to the next entry event, the merged firm makes sales on the new network at a price of ql{β(2) + δβˆ t+1(ql)} and gross consumption benefits are 2qlβ(2) per period.

(ii) If the merged firm operates a single network using the older technology, then in the continuation equilibrium the merged firm makes sales at a price of qf{β(2) + δβˆ t+1(qf)} and gross consumption benefits are 2qfβ(2) in the entry period and all subsequent periods prior to the next entry event

Propositions 2 and 3 demonstrates that, in this example, for ql > qf,

[Equations omitted]

which satisfy the reduced-form profit assumptions in Section 3 above.

The conditions of this example are favorable to entry. When mergers are prohibited, every firm chooses to enter given the opportunity and is an active producer for two periods. Hence, much of the time the industry is subject to product-market competition if ρ¯ is near 1. Proposition 1 implies that, by contrast, if mergers are freely allowed, then entry is immediately followed by merger, and there are no periods of product-market competition—the favorable entry conditions alone cannot protect product-market competition.

6.2. Inefficient entry and killer acquisitions

The favorable entry conditions can, however, promote the introduction of innovative new technologies. Given the assumed nature of consumer expectations, even a minor innovation can allow the entrant to appropriate a large fraction of the monopoly profits. As a result, there can be socially excessive entry.33

One situation in which this arises is when entry occurs and the merged firm uses the older technology: The only effect of entry on total surplus is to reduce it by the amount the entrant invests to develop its product. In some respects, the outcome in which the merged firm operates a single network using the older technology is similar to what is known in the pharmaceutical industry as a “killer acquisition,” where an incumbent buys a new drug but never introduces it to the market (Cunningham et al. (2019)). By contrast to pharmaceutical markets, however, there can be consumer and efficiency benefits of sticking with the older technology in order avoid the loss of network benefits that occurs when users are split across networks

In fact, a firm may fail to stick with the older technology when doing so would be efficient.34 Consider the technology-adoption decision of a firm newly created through merger. The private cost to the merged firm from splitting the current cohorts is equal to the loss in the consumption benefits it is able to appropriate: qf{β(2) − β(1)}. By contrast, the social cost from splitting the cohorts is 2qf{β(2) − β(1)}. The difference arises because the firm does not internalize the losses suffered by consumers who are in the second half of their lives during the transition period.

One also needs to account for the effects of the quality choice on profits in future periods. Adoption of the new technology generates gross consumption benefits prior to the next entry event, which the merged firm fully appropriates in expected value: the firm charges ql{β(2) + δβˆ t(ql)} instead of qf{β(2) + δβˆ t(qf)}. This effect induces no bias.

Adoption of the new technology also improves the merged firm’s bargaining position with respect to the next entrant: that entrant’s non-merger profits are decreasing in the merged firm’s quality. This bargaining effect is a purely private benefit.

The net effect of the first three components is to bias the firms toward adopting the new technology. However, there may be an effect running in the opposite direction: by adopting the new technology, the current firm may generate surplus for the next entrant because, when the current firm merges with that entrant, the new merged firm will have a more attractive option when deciding whether to use the current firm’s network. When δ is sufficiently small, failure to internalize the consumer losses due to stranding dominates, and the current incumbent and entrant are biased toward adopting the new technology. It is an open question whether, in other situations, merging firms can be biased toward the old technology.

In some situations, permissive merger policy can prevent inefficient stranding. As shown in Proposition 2(ii), when mergers are prohibited, the market will move to the new network and associated technology whenever the entrant’s product quality is at all higher than the incumbent’s. However, when mergers are allowed, it can be seen from Proposition 3 that the merged firm will stick with the old technology whenever the quality levels are sufficiently close together. Hence, there is a range of quality differentials such that total surplus is maximized when consumers remain on the old network and this happens if and only if the firms merge. In summary,

Corollary: By facilitating killer acquisitions, permissive merger policy can prevent socially inefficient stranding.

Although the corollary identifies a logical possibility, it is not clear how important it is in practice; the entrant’s innovation must be a large enough improvement that the entrant provides a credible competitive threat to the incumbent but not so large that the merged firm will adopt it.35 It should also be noted that even a merger that prevents inefficient stranding harms consumers through the loss of competition in the transition period.

6.3. Merger policy and entrant innovation

In addition to showing that killer acquisitions can sometimes be efficient, the network effects example illustrates that the option to merge can raise or lower the entrant’s marginal investment incentives. First, suppose that the merged firm adopts the entrant’s technology. Then, by Propositions 2 and 3,

[Equation omitted]

Similarly, if the next merger leads to adoption of the then-new technology,

[Equation omitted]

Recall that, in this example, ρE l+j (ql) = ρ¯ for all j > 0. Hence, by Eq. (9), when merged firms always adopt the newer technology,36

[Equations omitted]

Anticipation of the initial merger has no effect on the entrant’s marginal investment incentives; the gains from that merger arise because consumers do not get the surplus that the incumbent would have offered them as an independent competitor, and this amount does not depend on the entrant’s quality. However, for the same reason, an increase in the current entrant’s product quality does raise the gains from the merger with the next entrant if the current entrant does not merge with the incumbent. The current entrant’s anticipation of sharing the increased gains from a future merger raises its product-development incentives. The same argument applies to the first firm to enter the industry because it anticipates merging with a later entrant.

Now, suppose that the current merged firm retains the incumbent’s technology but again assume that, when the next entry event occurs, the resulting merged firm will adopt the technology of that entrant. Given the latter assumption, Gl+j(ql, ql+j) = β(1)ql and βˆ t(ql) = βˆ for all t from the entry period until the next entry event. As in the case just considered, an increase in the entrant’s product quality raises the gains from the merger that would take place between the current entrant and the next entrant if the former did not merge, which raises the entrant’s disagreement profits. Now, however, the gains to the initial merger, Gl(qf, ql), are decreasing in ql. Intuitively, the higher is the entrant’s product quality, the greater the opportunity cost of not using it. Formally, the effects can be broken into three components using Propositions 2 and 3. First, the gains from the initial merger realized during the transition period,

[Equation omitted]

are decreasing in the entrant’s quality. So too are the merger’s effects on the profits earned in each subsequent period prior to the next entry event,

[Equation omitted]

Third, because the merged firm adopts the incumbent’s technology, the gains to the current merger arising from effects on the gains to agreement associated with a subsequent merger, gS l+j (qf, ql) ≡ β(1)(qf − ql) are decreasing in ql.

Using the fact that ρE l+j (ql) = ρ¯ for all j > 0,

[Equation omitted]

Differentiating and collecting terms,

[Equation omitted]

The right-hand side of Eq. (11) is negative for any σ ≤ 0.5 or when σ < 1 and δ is sufficiently close to 0.37 Summarizing this discussion:

Proposition 4. When product-market competition exhibits network effects, permitting mergers: (a) can increase an entrant’s product-development incentives when the merged firm adopts the entrant’s technology; and (b) can reduce an entrant’s product-development incentives when the merged firm retains the incumbent’s technology.

7. An example with per-period fixed costs

Next, consider a market in which there are no network effects but, at the start of each period, a firm must incur a fixed cost of F in order to be an active producer in that period. An incumbent’s investment decision for each period is made before that period’s entry opportunity arises. A new entrant sinks its initial per-period investment at the same time that it makes its one-time entry investment, I. For simplicity, assume there are constant marginal costs of production, which are subsumed in the demand function.

Suppose there is no horizontal product differentiation and all consumers identically value vertical quality improvements, so that all consumers choose the same product in any period. Let X = qi − pi denote the per-period market demand function, where i is the index of the active producer offering the lowest quality-adjusted price, pi − qi. Any firm not charging the lowest quality-adjusted price makes no sales.

Consider a single incumbent with quality qf facing an entrant with quality ql. 38 If ql ≥ 2qf, the entrant’s innovation is drastic (i.e., because the monopoly price is ql/2, the current incumbent provides no competitive constraint even if it sets its nominal price at 0). In this case, the firms have is no incentive to merge, and the incumbent will make no further investments. Until the next entry event occurs, the current entrant earns profits of

[Equation omitted]

per period. If ql ∈ (qf, 2qf), the entrant’s innovation is non-drastic. If the firms do not merge and the next entry event has not yet occurred, then the current entrant earns

[Equation omitted]

in each period that the incumbent remains active and π M(ql) per period if the incumbent has dropped out.

Whether or not it remains active, the incumbent makes no sales in any period following entry. Assume that πC(qf, ql) > 0, so that the current entrant would find it profitable to pay the per-period fixed cost even if it anticipated the incumbent’s doing so as well.

If there is no possibility of merger, then the incumbent never invests F after entry has occurred and, thus, is a competitive constraint at most during the transition period. But suppose that merger is allowed. If the firms merge, they jointly earn product-market profits π M(ql) in each period until the next entry event. The gains from merger depend on what the incumbent would do absent merger. In the network-effects example above there is nothing that an incumbent can do to remain competitively relevant after the entry period because consumers’ perceived value of its product falls. By contrast, in the present example, an incumbent chooses whether to remain a competitive constraint (i.e., whether to invest F), which raises the possibility that an incumbent might sink F solely to induce the current or later entrants to pay it more to merge. That is, an incumbent might engage in incumbency for buyout.

To avoid some of the complications associated with such strategies, assume that the next entry event after the current one will entail a drastic innovation relative to the current incumbent (and possibly to the current entrant as well). This assumption implies that the presence of the incumbent as an independent firm would have no effect on the subsequent entrant’s incentives, and as an independent firm the current incumbent would have no ability to extract rents from any later entrant. Hence, merging does not give rise to a share-dilution effect. It follows that the incumbent and current entrant will merge to avoid dissipating rents in one or more periods until the next entry event occurs.

There remains the issue of the acquisition price. The cooperative-game-theoretic approach to bargaining with respect to the merger is problematical because the disagreement payoffs are not obvious when F is sufficiently small relative to π M(ql) − πC (qf, ql). For example, if subsequent entry occurs whenever the opportunity arises, then, depending on what the incumbent would do off the equilibrium path absent merger, the gains from merger could range from π M(ql) − πC (qf, ql) (when the incumbent drops out after the initial entry period)39 to πM (ql)−πC (q f ,ql)+δF 1−δ+δρ¯ (when both the incumbent and entrant invest in remaining in the market as long as they have not merged and there has been no subsequent entry in any previous period). But regardless of the acquisition price, the firms have incentives to merge, and a merger reduces consumer and total surplus by eliminating transition-period competition.

The value of the acquisition price can, however, affect an entrant’s product-development incentives. Suppose that a firm has an opportunity to enter using the generationl technology and face an incumbent with product quality qf. Moreover, suppose that technological opportunities are such that ql(I) ∈ (qf, 2qf) for all I, and with probability one an entry opportunity will arise in period l + 1 with ql+1(I) > 4qf for all I (so that ql+1(Il+1) will be drastic relative to ql(Il)). Under these assumptions, the gains from a merger between the incumbent and the generation-l entrant equal the increase in profits from the elimination of competition in the entry period. After that, a new firm will enter and displace the merged firm with a drastic innovation, which eliminates any incentive for a further merger. Hence, l(ql) = (1 − σ ){π M(ql) − πC (qf, ql)}. Differentiation yields l (ql) = 1 2 (1 − σ ){ql − 2qf}, which is negative given that ql(I) < 2qf for all I. Intuitively, the entrant chooses a lower product quality when mergers are allowed because the possibility of merger leads the entrant to put weight on the effects of its innovation on monopoly profits (which are a component of the gains from merger), which are less sensitive to the entrant’s quality than are the profits the entrant earns when competing with the incumbent.

In summary:

Proposition 5. When there is undifferentiated Bertrand product-market competition, permitting mergers can reduce an entrant’s product-development incentives even when its technology is adopted by the merged firm.

8. Incumbency for buyout

As long as an incumbent’s actions (e.g., investment levels or exit decisions) conditional on not merging are independent of whether mergers are permitted, the option to merge weakly increases an entrant’s profits. Hence, in this case, allowing mergers encourages innovative entry, and there can be situations in which entry is profitable only if entry for buyout is feasible. But the finding that the option to merge can only increase an entrant’s profits does not apply if there is a possibility that the incumbent can engage in conduct to appropriate a larger share of the post-entry profits through merger. As the following example shows, by facilitating incumbency for buyout, permissive merger policy can discourage innovative entry.

The structure of this example is the same as the one in the previous section, with three differences.40 \*\*\*FOOTNOTE BEGINS\*\*\* Permissive merger policy can discourage innovative entry in the example in Section 7 if acquisition prices are based on the belief that the incumbent will sink F in future periods if the firms have not yet merged. Intuitively, in the periods after the entry period but before the next entry event, the current entrant has to share π M(ql) with the incumbent rather earning it all itself. I consider the modified example in the text to avoid the issue of multiple equlibria based on beliefs regarding moves off the equilibrium path. \*\*\*FOOTNOTE ENDS\*\*\* First, F = 0, so there is no issue of the incumbent’s dropping out in future periods. Second, whenever an entry opportunity arises, it is drastic: ql ≥ 2qf. Third, if entry occurs, an incumbent can make a one-time, sunk investment of Z that allows it partially to imitate or catch up to the entrant, with resulting quality level qZ f , where ql ∈ (qZ f , 2qZ f ). The incumbent can imitate only the next entrant to arrive. The choice of investing Z is made after entry but before any merger discussions. There are gains from merger if and only if the incumbent invests Z.

The incumbent has no incentive to invest Z if mergers are prohibited—once entry has occurred, the incumbent makes no sales whether or not it invests Z. But when mergers are allowed, the incumbent will find it profitable to invest whenever Z < σGl(qZ f , ql) because doing so induces the entrant to merge and share some of the post-entry profits. Investing Z lowers the current entrant’s profitability, but the probability of later entry is unaffected in this example because the current incumbent is irrelevant as a standalone firm and neither the entrant’s nor merged firm’s costs are affected.41 Conditional on entry’s occurring, the incumbent’s investment has no effect on either gross consumption benefits or consumer welfare. For any given stream of entry events, the net present value of industry profits is less when mergers are allowed than when they are banned; the difference is the net present value of the rent-seeking investments in Z. It follows that, when Z is positive but sufficiently small that incumbents make the investment in response to entry when mergers are allowed, there must be at least some entry events that are less profitable when mergers are allowed than when they are banned. Thus, there exist parameter values for which merger discourages at least some entry.42

Summarizing the discussion of this section,

Proposition 6. Depending on the specifics of the market, banning mergers may: (a) reduce innovation by preventing entry for buyout, or (b) promote innovation by preventing incumbency for buyout that would otherwise deter entry.

This example only scratches the surface of actions an incumbent could take to improve its bargaining position. Here, the investment takes place only after entry has occurred. An incumbent might also make pre-entry investments designed to appropriate rents should entry occur. Such investments might also affect the probability that entry occurs at all. An interesting line of future research would examine how the possibility of merger affects an incumbent’s pre-entry investments, which can serve both to deter entry and to extract rents should entry occur.

Although the example presented above is very specific, it is sufficient to establish the general point that arguments that permissive merger policy supports innovation are too simple: banning acquisitions can, at least in theory, increase or decrease innovative entry. Moreover, as discussed above, entry incentives can be socially excessive, so that, promoting additional innovation does not necessarily promote greater total surplus.

9. Implications for merger policy

It is often argued that, under Schumpeterian competition, a firm that appears to be dominant today based on its market share may actually face strong competitive pressures from the threat of being displaced by an innovative entrant with little or no current share.43 This argument does not imply that dominance is not a concern. Instead, the importance of innovative entry as a driver of market performance provides a rationale for paying increased attention to harm to emerging or potential competition when assessing acquisitions by incumbents in such markets: acquisition of a firm that does not yet have a substantial—or even any—share of the incumbent’s immediate product market may nonetheless substantially harm future competition. Moreover, in markets subject to Schumpeterian competition, enforcement authorities should be wary of claims that the threat of future entry will discipline the market by discouraging inefficient mergers. As shown above, due to the absence of the share-dilution effect, acquisitions are especially likely to be a profitable strategy for avoiding Schumpeterian competition.

It is also often argued that mergers facilitate innovation by allowing entry for buyout. The analysis above, however, demonstrates that the possibility of merger can have positive or negative effects on entrants’ innovation incentives. Hence, in some instances, blocking mergers will promote both dynamic and static efficiency, rather than sacrificing the former for the latter.

Several recent reports address competition policy toward large digital platforms. They generally call for heightened scrutiny of acquisitions by dominant firms in markets with strongly increasing returns to scale, innovation, and competition for the market, especially with regard to acquisitions of potential or nascent competitors.44 The analysis of the present paper broadly supports these proposals.45 \*\*\*FOOTNOTE BEGINS\*\*\* It also supports the conclusion that merging parties should not be exempted from notifying antitrust authorities based on having low current sales and/or market share. Furman et al. (2019, pp. 12 and 94-95) and Stigler Committee on Digital Platforms (2019, p.16) express concern with underreporting. The U.S. Federal Trade Commission has recently requested information on past un-notified acquisitions by major tech firms (U.S. Federal Trade Commission, “FTC to Examine Past Acquisitions by Large Tech Firms: Agency Issues 6(b) Orders to Alphabet Inc., Amazon.com, Inc., Apple, Inc., Facebook, Inc. Google, Inc., and Microsoft Corp.,” press release, February 11, 2020), and the European Commissioner for Competition has stated that the European Commission is developing means for screening smaller acquisitions by Big Tech (Arjun Kharpal, “EU says it will look closer at smaller acquisitions made by big tech firms after ‘shopping spree,’” CNBC, February 18, 2020, available at cnbc.com, accessed March 13, 2020). \*\*\*FOOTNOTE ENDS\*\*\*

#### That entrenches slow growth---only a dynamic technology sector can rebuke the trend.

Boris Kheyfets & Veronika Chernova 21, Doctor of Economic Sciences, Professor, Chief Researcher Institute of Economics of the Russian Academy of Sciences Nakhimovskiy; PhD, Associate Professor Department of International Economic Relations Peoples' Friendship University of Russia Miklukho-Maklaya, "Comparative Assessment of the Influence of a Technological Factor on Economic Growth," Eastern-European Journal of Enterprise Technologies, Vol. 1, No. 109, February 2021, pg. 6-13.

A growing role of science and state-of-the-art technology in ensuring sustainable economic growth has become obvious lately [1, 2]. The innovation type of development has placed a special emphasis on the use of the leading-edge technologies, the production of high-tech products, the implementation of progressive organizational and management decisions [3]. Technology has fundamentally and quickly changed the structure of the world economy and has become one of the primary factors in economic progress. The shifts have outlined the radically new global space, novel conditions for competition in world markets, and modern principles of interaction between enterprises.

The role of technology in today’s economy has long been debated among researchers [4, 5]. However, there is still a lack of studies on the reasons behind technological inequality between countries. Currently, one can observe a new bipolar configuration of the global technological space forming, where the USA and China are taking the lead and all other countries are unable to close this gap in the short term [6, 7]. The spread of technological innovations is uneven, which causes technological inequality to emerge that represents a new challenge to sustainable economic development. The availability of technology and capital exacerbates the problem of economic differentiation. At that, the modern form of uneven development can no longer be represented using the common schemes, since it is widely manifested in various fields. Such indicators as labor productivity, living standards, GDP per capita, etc. characterize the overall state of national economies, but do not specify the factors which contributed to obtaining this position. Structural analysis highlights that the technological factor is among the most significant ones determining the objective pattern of uneven development [8]. However, the question remains about the constituent parts of the technological factor (its component base), methods and approaches to assessing the influence of this factor on economic growth. Researchers have different approaches to the selection of a set of technological factor indicators. This poses a problem of methodological consistency that precludes comparative research. For this reason, the topic of this study is becoming relevant, related to the study of the influence of the technological factor on differences in economic growth and inequality between countries.

Thus, the relevant studies point to a distinctive primacy of manufacturability as the main factor in sustainable economic development. Then, we aim to clarify the role of the technological factor. However, even now one can argue that the aggravated cross-country competition implies the need for tools to assess and determine the key determinants of technological economic growth. The results are expected to confirm the significance of the technological factor, allow identifying its parameters and setting their priorities for improving economic policy aimed at sustainable development. These circumstances understood will open up opportunities for countries to narrow the technology gap.

2. Literature review on the technological factor of economic development

Economic theory pays special attention to issues of development and sustainable growth, as well as the causes of differences and factor changes. The sources of economic growth through GDP were specified in [9–13]. These researchers agree that sustainable economic growth is driven by factors such as new technologies and globalization. However, with the availability and access to these factors, it becomes important to build optimal management. The dynamics of economic growth is believed to be based on the results of structural transformations, mastering new technological principles, the introduction of innovations and an increase in labor productivity. At that, the seemingly insignificant differences in the economic growth rates bring about the substantial divergence in countries’ economic potential. Determining these discrepancies becomes a relevant scientific task.

It is becoming increasingly obvious that if the economy is not focused on technological innovation, it has no prospects for long-term development [14–17]. Some researchers, such as [14], focus on fundamentally new solutions (patents) that have commercial implementation potential. We can agree with this opinion, because it is innovation that should ensure accelerated economic growth at the expense of competitive advantages. A similar opinion is expressed by [15]. The publication [16] proves that renewed industrialization becomes an important condition for the development of technology. According to [17], entrepreneurial skills are needed to support industrialization.

Numerous studies [18–20] demonstrate that there is a direct correlation between the technological preparedness of a country and its ranking in the global economy. Research results on this issue are coordinated. These trends, if underestimated, lead to the fact that some countries can find themselves lagging behind. Here, it is important to realize the essence and the role of the technological factor, as well as the opportunities for managing the level of technological effectiveness of the economy. However, in [18, 19] there are no clear indications of quantitative measures of the technological factor.

We agree with [21], who claims that the technological factor is new technologies or their clusters that underlie the changes in the relative cost of production factors, stimulate the development of new industries and enhance the efficiency of traditional ones. Historical regularities in the emergence of fundamental technological innovations give impetus to structural changes in the economy [22]. Therefore, it is important to identify the determinants of economic growth that occurs against the background of technological structural changes. As practice shows, national economies, which for one reason or another were unable to independently create high-tech products, first applied imitation strategies within the country, and then entered foreign markets by occupying particular niches [23–25]. These researchers note the role of R&D spending and high-tech exports in economic growth. However, factor quantitative estimates are not given. The development of the USA and China are interesting cases here. For example, from a country that had mainly copied innovations, China turned into one of the leading innovation-generating nations leaving behind most other countries in terms of the level of technological development. In this context, the patterns of production, distribution, exchange and consumption of goods are largely predetermined by the peculiar nature of the technological processes [8]. At the same time, the observed temporal reduction of cycles is formed precisely due to the technical progress and the use of innovations [26].

The study of the reasons behind technological inequality is believed to lend some insight into the mechanisms that underlie economic changes. According to [27, 28], the choice of a model of economic growth should focus on mobilizing the potential to follow the technological path of evolution. Since the modern development of the theory of evolutionary economics is based, first of all, on the neo-Schumpeterian theory, which determines the need for structural technological changes in ensuring sustainable economic development, such changes provide for the formation of new industries with a high degree of processing of primary raw materials and an increase in the efficiency of traditional ones. Therefore, the issue of developing an integral strategic management system aimed at ensuring innovative structural changes becomes relevant. As we see it, these changes are of a technological nature.

Thus, the literature review demonstrates that economic growth is significantly affected by the flows of developed and exported technologies [29], as well as R&D costs [30–32]. The presence of stable patterns for these factors allows us to use them in the assessment model. The indicators proposed by the researchers (the share of ideas with the potential for commercialization [33], the share of R&D funding in GDP [34], indicators of science, technology and innovation development [35], the number of patents [36]) often reflect the multidirectional dynamics of the technological factor’ financial aspects and its qualitative components.

The review confirmed the significance of the technological factor for economic growth. At the same time, there is a clash of researchers’ opinions on key determinants. In the context of the literature review, the indicators of the technological factor need to be revised. The question about the approaches to assessing the impact of the technological factor on economic growth is left unanswered, which proves the relevance of the present research.

3. The aim and objectives of the study

The aim of this study is to develop an integrated approach to assessing the impact of a technological factor on economic growth. This will provide an opportunity for a comparative analysis on the countries for technology gaps.

To achieve the stated goal, we aim to fulfill the following objectives:

– to determine the leading countries and outsiders in terms of digitalization of the economy;

– to assess the dependence of economic growth on the technological factor.

4. Materials and methods

In the present study, technological effectiveness refers to the ability of a country to implement structural reorganization in accordance with the model of innovation development and realize its scientific and technological potential. We evaluate the level of technological effectiveness of the economy using the relevant index that serves as the basis for ranking countries. The set of technological factor indicators that will be used in our approach will be adjusted taking into account the literature review.

To calculate the Index (Ii), we use the indicators characterizing various aspects of technological development of the nations under review (Table 1), such as:

– industrial production index (ai );

– the share of the production of machinery and equipment in total value added (bi );

– the share in global value added by the economic activity ‘Production of computing, electronic and optical equipment’ (ci );

– the share in global value added by the economic activity ‘Production of machinery and equipment’ (di );

– ICT development index (ei );

– domestic R&D costs, % in GDP (fi ).

For empirical verification, we use official statistics. The frequency of data updating does not allow reflecting the most recent trends that affect economic processes (such as the impact of COVID-19). This is a research limitation. We also need to understand that some trends are short-term in nature, and their impact can be neglected.

[Chart omitted]

The method of Euclidean distances is used to rank the indicators’ values; normalization (Ixi ) is calculated by formula (1). The boundaries of normalized indicators are set in the range from 0 to 1.

[Equation omitted]

where Xi is the actual value of the indicator; Xmin is the minimum value of the indicator for the sample population; Xmax is the maximum value of the indicator for the sample population.

The level of technological effectiveness is calculated using the cumulative method as a weighted mean:

[Equation omitted]

The closer the Index value is to 1, the higher the level of technological effectiveness of economy.

To determine the econometric relationship between economic growth and indicators characterizing the technological factor, a linear multiple regression model was applied.

[Equation omitted]

where X1, X2, X3…, Xn denote factors; ɛ denotes error; β denotes a vector of the parameters under evaluation.

The gross domestic income of the United States and China for the period of 1996–2019 was taken as dependent variables (Table 2).

Table

Description automatically generated

The independent variables were represented by the volume of electronics production (Elc), costs incurred in installation and maintenance of equipment/technologies (CTech), the volume of high technology exports (HTExp), and investment in R&D activities (RD). Data are given in Table 3.

[Table omitted]

Based on the purpose of the study, we put forward two hypotheses about the nature of the patterns observed:

Н1. Growing R&D costs accelerate economic growth. Such an increase is expected to stimulate R&D in industries with comparative advantage. Consequently, this strengthens the country’s exports (foreign trade surplus).

Н2. Arrested technological development adversely affects competitiveness and, as a result, economic growth, since outdated equipment results in higher resource intensity and low labor productivity.

We test the hypotheses and the methodology for assessing the level of technological effectiveness using the sample of 30 countries. The aggregate of research objects embraces several developed countries, developing countries with high GDP, as well as developing countries not included in leading world economies. The selection is due to the need to cover a wide range of economies characterized by a wide variety of development conditions.

5. Results comparing technological effectiveness of economies

5. 1. Leading countries and outsiders in terms of technological innovation

The global economy in the context of Industry 4.0 demonstrates a number of specific features that distinguish it from the previous development stages. Firstly, technological innovation is becoming increasingly expensive, which causes a significant increase in R&D costs [38]. Secondly, the rate of technological change has increased dramatically. The terms of development and implementation of new solutions were reduced in the first place [8]. Technological gap can now be measured exponentially [39].

Look at a range of indicators characterizing the level of technological effectiveness of national economies. The share of domestic R&D costs in GDP is one of them (Fig. 1). The highest level of R&D funding in GDP is observed in the Republic of Korea, Sweden, Japan, Germany, the United States, China and other countries leading in the Global Competitiveness Report.

Analysis of the current changes in the global economy indicates that the importance of the comparative advantages of the lower order – cheap labor, basic production resources and the availability of raw materials – is decreasing [40]. At the same time, advantages of a higher order are gaining in significance, such as the ability of countries to develop high-tech industries, to manufacture and export products with a high intellectual component and in-depth processing [41]. For instance, the United States and China account for 90 % of the market capitalization value of the world’s 70 largest digital platforms, 75 % of all patents related to blockchain technologies, more than 75 % of the world market for public cloud computing, about 50 % of global spending on IoT, 40 % of world data centers, 36 % of the global value of e-commerce [42], and 69 % of supercomputers [43]. These areas are of significant potential and can have a serious impact on economic restructuring. Therefore, a special focus of the analysis is put on such indicator as the share of high-tech production (including computing, electronic and optical technology) (Fig. 2). China, Germany, Italy, the United States and Japan have the largest share in global value added in the production of computing, electronic and optical equipment. Norway, Canada, Australia, Sweden, Romania, Poland, etc. are relatively poorly represented in these world markets.

High-tech industries focusing on domestic production can be viewed as sources of economic growth. Data on the share of machinery and equipment production in GDP show similar trends (Fig. 3). High-tech industries strongly stimulate the economic growth of the leading countries – the Republic of Korea, China, the United States, Germany, and Japan, – while countries with low competitiveness demonstrate poor results.

[Table omitted]

Analysis of the countries indicates that some of them did not demonstrate high values of the indicators reviewed, but the level of their technological effectiveness is much higher (the group of “backward” countries embraced Denmark, the Netherlands, Sweden, Norway, and Canada). To gain a comprehensive picture and rank the countries, we have calculated the integral index of the technological effectiveness that covers financial aspects of development, as well as qualitative characteristics of economic growth. The Index calculation methodology is presented in section 4 of the paper. The countries’ ranking is presented in Table 4.

[Table omitted]

5. 2. Assessment of the dependence of economic growth on the technological factor

As articulated earlier, an increase in GDP can result from various factors. To substantiate the relationship between economic growth and the technological factor, we construct a number of models. The parameters of the regression models for the USA and China are given in Tables 5, 6. The parameters of the multiple regression model were obtained using STATISTICA software.

[Table omitted]

We have obtained a model with good quality characteristics; in this case, the coefficient of determination R2=0.996, normalized R-squared=0.995, multiple R=0.998.

[Table omitted]

The model obtained for China is also characterized by good quality characteristics: the coefficient of determination R2=0.999, normalized R-squared=0.999, and multiple R=0.999. Checking of the model adequacy according to the F-test produced the following results: the calculated value F=10.09 at the level of significance p=0.01.

Having analyzed the models’ data, we can conclude that there are no factors with a high probability of insignificance (t-Statistic for each model are greater than the critical value at a significance level of p=0.01), i.e. all regressions are significant.

To evaluate the degree of adequacy of the constructed trend equation to the real process, the mean approximation error was computed. Its value (3.167 % for China and 1.54 % for the United States) indicates that the degree of the quadratic equation’s adequacy to the real conditions of the relationship between economic growth and the technological factor is high.

Fig. 4 provides a visual distribution of actual and calculated values of the regression models.

Analysis of the models for the United States and China allows us to deduce that R&D costs are significant regressants contributing to economic growth; the factor impact on GDP growth in the United States and China is 31.6% and 41.9%, respectively; export of high-tech products provides an increase in GDP by 2.7% and 4.7%, respectively. It is worth noting that the obtained negative coefficients in the regression models suggest a weak correlation between the effective feature (economic growth through GDP) and some factor variables. For China, the indicator “Costs incurred in installation and maintenance of equipment/technologies” reveals an inverse relationship with GDP. A similar trend is observed in the United States for the indicator “Production of electronics”. Our calculations confirm that the strongest relationship is observed between GDP and development costs, as well as the share of high-tech industries in global value added.

[Chart omitted]

The current research proves that countries with substantial R&D funding and a large share of high-tech products in GDP and total exports are characterized by sustainable economic growth. Thus, the H1 hypothesis was confirmed.

The H2 hypothesis was partially confirmed: countries capable of using their innovative potential effectively are characterized by an elevated level of competitiveness. However, the use of outdated technologies does not always results in a decrease in global competitiveness, since these processes can be influenced by the institutional environment, which was beyond the scope of the present study

6. Discussion of the results comparing technological effectiveness of economies

Testing the approach using the case studies of China and the United States makes it possible to extrapolate their experience to countries with a low level of technological effectiveness. For example, the China and USA lead the global market for technological innovation. The country’s competitiveness in this field is due to the highly dynamic nature of American business, strong institutional underpinnings, finance mechanisms and a powerful innovation ecosystem [1]. Index of the countries’ technological effectiveness (Table 4) confirms this trend. The calculated values of the Index indicate the leading positions of these countries. The rapid growth of the renewable energy sector is a testament to why China will continue to dominate the sectors in which it invests heavily [44]. Currently, the PRC accounts for 90 % of the world’s supply of mobile phones and personal computers. In 2018, the country’s share in global semiconductor consumption was 41 %; by 2024, it is forecasted to increase to 54 % [45]. Significant funds received from low- and medium-tech industries in China are directed to those economic sectors, which enjoy research, development and implementation of hightech solutions.

It is noteworthy that in terms of the level of technological development, Kazakhstan, Brazil and Ukraine lag significantly behind some European nations (Romania, Poland, and Bulgaria), Turkey and Mexico. These countries do not exhibit sufficient potential to introduce innovations independently, but with regard to successful transfer and adaptation of foreign high technologies, they are significantly ahead of other countries with a similar development level. India is among the countries with high technological growth potential. India is now at a stage where machine learning tools are rapidly replacing entry-level programmers in the IT sector. So far, India is ranked 15th, but the situation may change soon. The comparison showed the advantage of the proposed methodological approach. We have been able to analyze the technicality of countries using universal data sets. The Index of the countries’ technological effectiveness can be a good alternative to other methods of assessment.

During the research, we have confirmed the hypotheses put forward. Assessment of the dependence of economic growth on the technological factor showed a strong relationship between GDP and R&D costs (Tables 5, 6). These results prove that sustainable economic growth is explained in most cases by significant funding for R&D (the presence of a large share of high-tech products in the country’s GDP) and the export of high-tech products.

Therefore, technologies determine competitive advantages of states at large. However, qualitative factors of economic growth prevail in a continuous innovation process. What determines additional limitations of our methodological approach. Special focus should be placed on a specific feature of the periods when changes occur, i.e. the periods of the so-called “technological gap” [46]. This is when the foundations of the future economy are set. Technological incentives crucial for growth are based on the ability to deliver better results. If technological inequality is excessively gross, it can jeopardize economic growth. Creating favorable conditions for the use of high technologies will not only support the competitiveness of production and attract investment in the economy, but also help resolve such issues as enhancing the efficiency of resource exploitation.

Hence, scientific and technological progress is the central stimulus for economic development, which in production processes is implemented through investment and innovation. At that, the dynamics of economic growth in the long run is dependent on a wide array of factors forming supply and demand for technological change: the current techno-logical capability of the national economy [19]; the development stage of financial institutions; companies’ awareness of R&D, and the effectiveness of technology transfer within the innovation infrastructure [47]; the nature of the state scientific and technical, scientific and technological, structural, and stabilization policy, and the level of state guarantees for the protection of intellectual property rights [25]; conditions of foreign economic activity, and competitiveness of products and services in the global market [48]. The characteristics of the listed factors vary significantly across countries, but the multicausality of the factors indicates that their combinations at certain time intervals can both reduce and boost the level of technological effectiveness.

#### BUT solely by ensuring lagging incumbents AND potential entrants can compete will lock-in any productivity gains.

James Manyika & Michael Spence 21, Chair and Director of the McKinsey Global Institute; Philip H. Knight Professor and Dean Emeritus at Stanford University's Graduate School of Business, "A Better Boom: How to Capture the Pandemic’s Productivity Potential," Foreign Affairs, Vol. 100, No. 4, August 2021, HeinOnline. language edited.

The pandemic did more than temporarily [freeze] ~~paralyze~~ the global economy, however. It spurred businesses in practically every sector to radically rethink their operations, often accelerating plans for technological and organizational innovation that were already in the works. Overwhelmingly, firms adopted new digital technologies that enabled them to continue doing business even under severe coronavirus restrictions. The result was a profound economic transformation, one that has hastened the potential for productivity gains even in sectors that have historically been slow to change. In health care, for example, telemedicine had long promised new efficiencies and added value, but it was not until the COVID-19 crisis that it took off. In retail, with the exception of e-commerce players, firms had been slow to adopt digital sales strategies, doing so mostly as a way to complement Main Street retailing. That changed rapidly with the pandemic.

Surprising as it may seem, out of the deepest economic crisis since World War II could come a new era of productivity gains and prosperity. Whether that happens will depend largely on the decisions that governments and businesses make as they prepare to exit the pandemic in the coming months. In the short and medium term, the prospects for increased productivity—and prosperity—are encouraging, as the United States and other countries spend heavily on economic recovery and businesses reap the benefits of digitization. But the outlook is less optimistic over the long term, since governments cannot spend indefinitely and consumer and investment spending may not fill the gap.

Governments and businesses must therefore seek to create the conditions for sustained productivity growth and prosperity, in particular by facilitating the diffusion of technological and organizational innovations and bolstering consumer demand. Out of a major global crisis could come a major jolt of productivity growth—but only if policymakers and business leaders make the most of this moment.

THE PRODUCTIVITY PARADOX

The history of productivity growth can be understood as a succession of technological revolutions, from the steam engine to the computer. Each offered the promise of accelerated productivity and economic growth, and each eventually delivered. But there has often been a delay between innovation and adoption, and another between adoption and economic impact. The economist Robert Solow summed up these apparent discrepancies in a 1987 article in The New York Times Book Review, writing, “You can see the computer age everywhere but in the productivity statistics.” His formulation became known as “the Solow paradox.”

But then came the revolution in information and communication technologies between 1995 and 2005, a decade in which the Solow paradox was temporarily resolved. Widespread adoption of these technologies was accompanied by a simultaneous acceleration in productivity, which grew at an annualized rate of 2.5 percent in the United States, a full percentage point faster than the rate between 1970 and 1995. Companies invested heavily in information and communication technologies and reorganized their operations and managerial practices around them. They did so out of the desire to gain a competitive edge, but also because of relatively robust consumer demand for their products.

Productivity growth accelerated in several sectors as a result, driving growth in the U.S. economy as a whole. This period was characterized by an unusual combination of large spurts in productivity growth in a few big sectors employing many workers, such as retail and wholesale, and even larger productivity growth in smaller sectors, such as those that produced computers and electronic products. In both big and small sectors, there was a virtuous cycle of employment growth to meet demand and even faster growth in the value of the output from these sectors. The value of outputs across all sectors of the economy grew by 3.4 percent per year between 1995 and 2005, whereas the total number of hours worked grew by only 0.9 percent per year.

But the boom did not last. Between 2005 and 2019, annual productivity growth in the United States fell by more than half, to 1.0 percent. In the aftermath of the 2008 global financial crisis, from 2010 to 2019, it was even lower, at 0.6 percent. Unlike the United States, European countries had not experienced rapid productivity gains in the 1995–2005 period, but they did experience the postcrisis decline. Between 2010 and 2019, annual productivity growth fell below one percent in France, Germany, and the United Kingdom.

The Solow paradox was back. After a decade of rapid productivity gains, the information technology revolution had reached a point of diminishing returns. But the next wave of technology—the digitization of processes, big data and analytics, cloud computing, the Internet of Things—was not yet ready to fill the gap. Despite early breakthroughs in image recognition and natural language processing, few firms had begun to make use of artificial intelligence technologies, and digitization was proceeding slowly. We estimated, based on a sector-by-sector assessment, that in 2015, the United States had reached only 18 percent of its digital potential and Europe had reached only 12 percent. Moreover, a gap had opened up between the firms that were digital leaders and those that were digital laggards—a gap that other researchers found was correlated with a gap in labor productivity.

This gap in technology adoption was widening at a time of weak consumer demand for goods and services, in large part due to the aftereffects of the financial crisis. Firms scaled back their investments, and fewer new businesses were created. Making matters worse, the share of income that flowed to top earners and the owners of capital increased, while the share that went to labor decreased, further weakening demand.

Across the United States and Europe, the vast majority of sectors experienced declines in productivity growth. Only four percent of all sectors recorded productivity jumps in 2014, compared with an average of 18 percent of sectors that achieved substantial increases in productivity in the previous two decades. Growth in gross value added—a measure of a firm’s or a sector’s contribution to GDP—declined from 3.4 percent annually between 1995 and 2005 to 1.8 percent between 2005 and 2019. Growth in hours worked remained roughly unchanged, at 0.7 percent, throughout both periods.

These two very different periods of economic activity in the United States reveal much about the underpinnings of productivity growth. It stems first and foremost from the widespread adoption of technological innovations, especially general-purpose technologies such as electricity and the Internet. But it also stems from the managerial innovation and reorganization of functions and tasks that occur when firms adopt new technologies. Both of these processes must spur leaps in productivity growth in many sectors, or at least in a few large ones, so that productivity jumps in the economy as a whole. Finally, adoption and reorganization within and across sectors must be driven by competition, which incentivizes firms to innovate and helps spur technological diffusion.

Not all productivity growth is created equal, however. Productivity growth can be achieved through gains in the volume or value of outputs for a given number of hours worked, or it can come about as a result of a reduction in hours worked for a given output. Often both happen at the same time. But it is when the former exceeds the latter that a virtuous cycle is created in which innovation and investment generate growth in employment and wages, which in turn generates demand for increased (or more valuable) output. This is what happened during the period from 1995 to 2005. When the latter source of productivity growth exceeds the former, however, a vicious cycle results in which firms reduce labor costs faster than they grow the volume or value of their outputs, which in turn puts pressure on employment and incomes.

POST-PANDEMIC POTENTIAL

The pandemic has primed advanced economies for another period of rapid productivity growth. It is too early to say for sure whether such growth will be the product of a virtuous or a vicious cycle, but signs point to the former. Despite uncertainty, stress, and plummeting economic activity in the early days of the COVID-19 crisis, many firms boldly deployed and used new general-purpose technology—especially digital technology—in ways that have driven virtuous productivity gains in the past. In October 2020, we surveyed 900 C-suite executives in various sectors and countries and found that many had digitized their business activities 20 to 25 times as fast as they had previously thought possible. Often, this meant shifting their businesses to online channels, since roughly 60 percent of the firms we surveyed experienced a significant increase in customer demand for online goods and services as a result of the pandemic.

Before the pandemic, e-commerce was forecast to account for less than a quarter of all U.S. retail sales by 2024. But during the first two months of the COVID-19 crisis, e-commerce’s share of retail sales more than doubled, from 16 percent to 33 percent. And that growth did not just reflect brick-and-mortar firms setting up shop online for the first time. Firms that were already highly digitized before the pandemic significantly expanded their online capabilities to meet the surge in demand. They also reorganized their operations, including their logistics, to complement what they were doing digitally—for example, by expanding their direct-to-home delivery capabilities.

Businesses also strove to become more efficient and agile. In Europe and North America, nearly half of the respondents to our survey said that they had reduced their operating expenditure as a share of revenue between December 2019 and December 2020. Two-thirds of senior executives said they had increased investment in automation and artificial intelligence, whether to help warehouse and logistics operations cope with higher e-commerce volumes or to enable manufacturing plants to meet surging demand. Many companies used technology to reduce the physical density of their workplaces or to enable contactless service—for instance, by expanding self-checkout in grocery stores and pharmacies and employing online ordering apps for restaurants and hotels. Other businesses, such as meatpacking and poultry plants, accelerated the deployment of robotics to reduce their need for labor. If there was one lesson from the pandemic, it was that digital capability and resilience go hand in hand.

But even as the arrival of vaccines has made it possible to imagine a return to relative normalcy in parts of the developed world, continued digitization and the adoption of other technological innovations promise to deliver still more productivity gains. The largest of these gains—roughly an additional two percentage points per year—could come in the health-care, construction, information technology, retail, pharmaceutical, and banking sectors. In health care, for instance, accelerating the use of telemedicine beyond the pandemic could drive incremental productivity growth for years. According to one recent U.S. poll, 76 percent of patients expressed interest in using telemedicine in the future, and industry experts project that the services for 20 percent of health-care spending could be delivered virtually—up from 11 percent before the pandemic. Other sectors, including automotive, travel, and logistics, show less—but still substantial—potential for productivity growth as a result of more flexible task scheduling, leaner operations, and smarter procurement.

Overall, these innovations and organizational changes could accelerate productivity growth by around one percentage point per year between now and 2024 in the United States and the six large European economies that we analyzed (France, Germany, Italy, Spain, Sweden, and the United Kingdom). This gain would result in a productivity growth rate twice as high as the rate after the 2008 global financial crisis, and in the United States, it would expand per capita GDP by roughly $3,500 by 2024. That would be a stunning outcome, but it will hinge on continued technology adoption by firms and the maintenance of robust demand.

Even more productivity gains could be on the horizon thanks to other advancements. The accelerating revolution in biology, for instance, could transform sectors from health care and agriculture to consumer goods, energy, and materials. Biological innovation has already enabled the rapid development of new vaccines for COVID-19. Equally impressive revolutions in energy could make possible the widespread adoption of solar and wind power, especially in light of recent progress toward better (and cheaper) batteries. Artificial intelligence is also advancing rapidly, but is still a long way from being deployed widely across companies and sectors. When and if that happens, the productivity gains could be enormous.

FOLLOW THE DIGITAL LEADER

Future gains in productivity, even those that boost overall growth, are likely to be uneven. We analyzed metrics that have the potential to unleash future productivity growth—such as research-and-development spending, revenue, capital expenditures (including digital expenses), and mergers and acquisitions—and found that especially in the United States, a small number of large superstar firms accounted for a disproportionately large share of the activity in all these categories. From the third quarter of 2019 to the third quarter of 2020, U.S. superstars (defined as the top ten percent of firms by profit) saw much shallower declines in capital expenditures and revenue than did other companies. During the same period, U.S. superstars spent $2.6 billion more on R & D than they did the previous year, while all other firms spent just $1.4 billion more.

If this investment, innovation, and technology adoption gap between superstars and the rest of the large firms and smaller, less profitable firms persists, any post-pandemic acceleration in productivity growth could fall short of its potential. Small and medium-sized enterprises have been hit disproportionately hard by the COVID-19 crisis. As a result, many of them are unable to make big investments in future productivity and are therefore liable to fall even further behind the superstars. This is what happened in the aftermath of the 2008 global financial crisis, when only a minority of companies achieved productivity growth.

But there is room for cautious optimism about the ability of non-superstars to close some of the gap. Before the pandemic, the superstars tended to be highly digitized and innovative in their managerial approaches, as well as more profitable and resilient. They were therefore better placed to weather and even take advantage of the shock. But as the hardest-hit firms and sectors recover, and as early digital adaptors demonstrate the enormous potential of these technologies, many of the digital laggards could begin to catch up. Indeed, in another survey of executives we conducted in December 2020, about 75 percent of respondents in North America and Europe said they expected investment in new technologies to accelerate substantially between 2020 and 2024, up from 55 percent between 2014 and 2019. This expected uptick was similar across firm sizes.

Another reason for optimism is that in 2020, a year that saw the darkest economic days of the pandemic, 24 percent more new businesses were created in the United States than in 2019. Europe lagged behind the United States on this metric, with new business creation staying roughly flat in 2020 in France, Germany, and the United Kingdom and declining by more than 15 percent in Italy and Spain. If the American increase in business dynamism persists, however, it should contribute to more productivity growth.

Investment, innovation, and technology adoption are only one-half of the virtuous cycle of productivity growth, however. The other half is demand for the expanded output that results—in other words, income growth from increased productivity has to flow to people who will spend that additional money. In the short term, the outlook for demand is good, especially for countries that have made progress toward vaccinating their populations and could be among the first to open up their economies. Pent-up demand and savings from the pandemic could be unleashed all at once, resulting in a strong initial bounce in demand led by consumers. In the United States, President Joe Biden’s $1.9 trillion economic support bill should push demand even higher.

In the medium term, the outlook for demand is also relatively solid, although it will depend on the size, deployment, and longevity of government spending. In the United States, Biden now has set his sights on a large infrastructure package. As his administration shifts its focus from economic relief to investment in productive areas, it could also increase productivity growth by raising demand to match potential supply, creating a high-pressure economy, that is, one with low unemployment and high growth. The outlook in continental Europe, where large-scale government economic support is harder to coordinate, is less certain. Nonetheless, the EU has put in place an unprecedented plan totaling some $900 billion to boost investment in the digital and green energy transitions.

But government spending on this scale will likely be time-limited, making the long-term outlook for demand less rosy. Moreover, long-neglected problems, including the falling share of firms’ income going to workers, rising inequality, and the long-term decline in private investment, could drag down demand. Roughly 60 percent of the post-pandemic productivity gains that we estimate could come from innovations and organizational restructuring—the one percentage point of acceleration per year between now and 2024—would stem from firm-level measures, such as automation, designed to cut labor and other business costs. Unless firms do more to boost the volume or value of their output and help workers transition by acquiring new skills, the drive for efficiency will risk generating productivity gains through a vicious, rather than a virtuous, cycle, undermining wages and jobs and weakening consumption-driven demand and investment.

A NEW AGE OF DYNAMISM?

What can businesses and governments do to capitalize on the positive short- and medium-term outlook for productivity and to improve the long-term outlook? First, they should work to speed up technology adoption and managerial innovation, helping these changes spread within and across sectors. As the recovery begins, firms that have until recently been focused on crisis management and survival should follow the lead of superstar firms by investing in technology and reorganization. The superstars can assist in this process by supporting their broader ecosystems, in particular by doing business with smaller firms that offer complementary products and services. Governments can support the process, as well, by investing in research and development.

Policymakers should also seek to strengthen competition and business dynamism. In a healthy economy, the firms that add the most value prosper and grow, while the firms that add the least value shrink or disappear: so-called creative destruction. Policymakers can revive and reinforce this natural sorting process by revising competition rules, bankruptcy procedures, and product and labor-market regulations.

Governments and businesses should also aim to bolster demand and encourage business investment, the other half of the virtuous productivity cycle. As government spending tapers off, businesses should play their part by creating broad-based revenue growth while also finding efficiencies. Additionally, they should spend more on upgrading the skills of their employees, helping them make the most of technological and organizational innovations while also reducing inequality and unemployment. Governments can incentivize such investments in human capital through tax credits that encourage retraining and by shifting the tax burden away from labor income and toward capital income.

But productivity growth isn’t everything, especially as it is measured and projected today. It does not capture important dimensions of individual and social well-being that may be significantly augmented in the post-pandemic environment. For instance, the spread of digital technologies could foster more inclusive patterns of growth, and telemedicine could deliver timely primary health-care services to millions in the developing world. Nor do measures of productivity growth account for some of the negative externalities associated with modern innovations, which will compound over time and profoundly affect people’s quality of life.

What is perhaps most notable is that productivity as it is currently measured does not account for climate change. To mitigate that risk around the world, significant investment in technologies that make energy greener and more efficient is needed. Some of this investment will increase productivity growth. Electric vehicles, for instance, are not just good for the environment; they also require less labor to produce and so raise productivity. To the extent that energy-efficient investments divert resources and talent away from other, even more potentially productive areas of the economy, they could dampen short-term productivity growth. Over the long term, however, their effect will be positive, since they will prevent a dramatic decline in future productivity, among other catastrophic outcomes. Many of these gains may never be captured by the standard productivity measures, since the gains will represent a downturn that never occurred. But some of the productivity gains could eventually be captured, especially those related to infrastructure designed to help the economy adapt to climate change.

As they prepare to exit the pandemic, governments and businesses alike will have to balance these short- and long-term goals. Yet even now, as COVID-19 continues to exact a human and economic toll, a potential upside appears to be emerging. After years of sluggish productivity and economic growth following the 2008 global financial crisis, COVID-19 has triggered a frenzy of technological and organizational innovation. Whether this frenzy leads to a new age of dynamism will depend on what governments and businesses do to sustain a virtuous cycle of ever-greater productivity.

#### Slow growth collapses the liberal order AND causes global hotspot escalation---it culminates in numerous existential risks.

Michael F. Oppenheimer 21, Clinical Professor, Center for Global Affairs, New York University. Senior Consulting Fellow, Scenario Planning at the International Institute for Strategic Studies. Former Executive Vice President, The Futures Group. Member, Council on Foreign Relations. Member, The Foreign Policy Roundtable at the Carnegie Council on Ethics and International Affairs. Member, The American Council on Germany, "The Turbulent Future of International Relations," in The Future of Global Affairs: Managing Discontinuity, Disruption and Destruction, Chapter 2, 2021, pg. 23-43.

Four structural forces will shape the future of International Relations: globalization (but without liberal rules, institutions, and leadership)1; multipolarity (the end of American hegemony and wider distribution of power among states and non-states2); the strengthening of distinctive, national and subnational identities, as persistent cultural differences are accentuated by the disruptive effects of Western style globalization (what Samuel Huntington called the “non-westernization of IR”3); and secular economic stagnation, a product of longer term global decline in birth rates combined with aging populations.4 These structural forces do not determine everything. Environmental events, global health challenges, internal political developments, policy mistakes, technology breakthroughs or failures, will intersect with structure to define our future. But these four structural forces will impact the way states behave, in the capacity of great powers to manage their differences, and to act collectively to settle, rather than exploit, the inevitable shocks of the next decade.

Some of these structural forces could be managed to promote prosperity and avoid war. Multipolarity (inherently more prone to conflict than other configurations of power, given coordination problems)5 plus globalization can work in a world of prosperity, convergent values, and effective conflict management. The Congress of Vienna system achieved relative peace in Europe over a hundred-year period through informal cooperation among multiple states sharing a fear of populist revolution. It ended decisively in 1914. Contemporary neoliberal institutionalists, such as John Ikenberry, accept multipolarity as our likely future, but are confident that globalization with liberal characteristics can be sustained without American hegemony, arguing that liberal values and practices have been fully accepted by states, global institutions, and private actors as imperative for growth and political legitimacy.6 Divergent values plus multipolarity can work, though at significantly lower levels of economic growth-in an autarchic world of isolated units, a world envisioned by the advocates of decoupling, including the current American president.7 Divergent values plus globalization can be managed by hegemonic power, exemplified by the decade of the 1990s, when the Washington Consensus, imposed by American leverage exerted through the IMF and other U.S. dominated institutions, overrode national differences, but with real costs to those states undergoing “structural adjustment programs,”8 and ultimately at the cost of global growth, as states—especially in Asia—increased their savings to self insure against future financial crises.9

But all four forces operating simultaneously will produce a future of increasing internal polarization and cross border conflict, diminished economic growth and poverty alleviation, weakened global institutions and norms of behavior, and reduced collective capacity to confront emerging challenges of global warming, accelerating technology change, nuclear weapons innovation and proliferation. As in any effective scenario, this future is clearly visible to any keen observer. We have only to abolish wishful thinking and believe our own eyes.10

Secular Stagnation

This unbrave new world has been emerging for some time, as US power has declined relative to other states, especially China, global liberalism has failed to deliver on its promises, and totalitarian capitalism has proven effective in leveraging globalization for economic growth and political legitimacy while exploiting technology and the state’s coercive powers to maintain internal political control. But this new era was jumpstarted by the world financial crisis of 2007, which revealed the bankruptcy of unregulated market capitalism, weakened faith in US leadership, exacerbated economic deprivation and inequality around the world, ignited growing populism, and undermined international liberal institutions. The skewed distribution of wealth experienced in most developed countries, politically tolerated in periods of growth, became intolerable as growth rates declined. A combination of aging populations, accelerating technology, and global populism/nationalism promises to make this growth decline very difficult to reverse. What Larry Summers and other international political economists have come to call “secular stagnation” increases the likelihood that illiberal globalization, multipolarity, and rising nationalism will define our future. Summers11 has argued that the world is entering a long period of diminishing economic growth. He suggests that secular stagnation “may be the defining macroeconomic challenge of our times.” Julius Probst, in his recent assessment of Summers’ ideas, explains:

…rich countries are ageing as birth rates decline and people live longer. This has pushed down real interest rates because investors think these trends will mean they will make lower returns from investing in future, making them more willing to accept a lower return on government debt as a result.

Other factors that make investors similarly pessimistic include rising global inequality and the slowdown in productivity growth…

This decline in real interest rates matters because economists believe that to overcome an economic downturn, a central bank must drive down the real interest rate to a certain level to encourage more spending and investment… Because real interest rates are so low, Summers and his supporters believe that the rate required to reach full employment is so far into negative territory that it is effectively impossible.

…in the long run, more immigration might be a vital part of curing secular stagnation. Summers also heavily prescribes increased government spending, arguing that it might actually be more prudent than cutting back – especially if the money is spent on infrastructure, education and research and development.

Of course, governments in Europe and the US are instead trying to shut their doors to migrants. And austerity policies have taken their toll on infrastructure and public research. This looks set to ensure that the next recession will be particularly nasty when it comes… Unless governments change course radically, we could be in for a sobering period ahead.12

The rise of nationalism/populism is both cause and effect of this economic outlook. Lower growth will make every aspect of the liberal order more difficult to resuscitate post-Trump. Domestic politics will become more polarized and dysfunctional, as competition for diminishing resources intensifies. International collaboration, ad hoc or through institutions, will become politically toxic. Protectionism, in its multiple forms, will make economic recovery from “secular stagnation” a heavy lift, and the liberal hegemonic leadership and strong institutions that limited the damage of previous downturns, will be unavailable. A clear demonstration of this negative feedback loop is the economic damage being inflicted on the world by Trump’s trade war with China, which— despite the so-called phase one agreement—has predictably escalated from negotiating tactic to imbedded reality, with no end in sight. In a world already suffering from inadequate investment, the uncertainties generated by this confrontation will further curb the investments essential for future growth. Another demonstration of the intersection of structural forces is how populist-motivated controls on immigration (always a weakness in the hyper-globalization narrative) deprives developed countries of Summers’ recommended policy response to secular stagnation, which in a more open world would be a win-win for rich and poor countries alike, increasing wage rates and remittance revenues for the developing countries, replenishing the labor supply for rich countries experiencing low birth rates.

Illiberal Globalization

Economic weakness and rising nationalism (along with multipolarity) will not end globalization, but will profoundly alter its character and greatly reduce its economic and political benefits. Liberal global institutions, under American hegemony, have served multiple purposes, enabling states to improve the quality of international relations and more fully satisfy the needs of their citizens, and provide companies with the legal and institutional stability necessary to manage the inherent risks of global investment. But under present and future conditions these institutions will become the battlegrounds—and the victims—of geopolitical competition. The Trump Administration’s frontal attack on multilateralism is but the final nail in the coffin of the Bretton Woods system in trade and finance, which has been in slow but accelerating decline since the end of the Cold War. Future American leadership may embrace renewed collaboration in global trade and finance, macroeconomic management, environmental sustainability and the like, but repairing the damage requires the heroic assumption that America’s own identity has not been fundamentally altered by the Trump era (four years or eight matters here), and by the internal and global forces that enabled his rise. The fact will remain that a sizeable portion of the American electorate, and a monolithically proTrump Republican Party, is committed to an illiberal future. And even if the effects are transitory, the causes of weakening global collaboration are structural, not subject to the efforts of some hypothetical future US liberal leadership. It is clear that the US has lost respect among its rivals, and trust among its allies. While its economic and military capacity is still greatly superior to all others, its political dysfunction has diminished its ability to convert this wealth into effective power.13 It will furthermore operate in a future system of diffusing material power, diverging economic and political governance approaches, and rising nationalism. Trump has promoted these forces, but did not invent them, and future US Administrations will struggle to cope with them.

What will illiberal globalization look like? Consider recent events. The instruments of globalization have been weaponized by strong states in pursuit of their geopolitical objectives. This has turned the liberal argument on behalf of globalization on its head. Instead of interdependence as an unstoppable force pushing states toward collaboration and convergence around market-friendly domestic policies, states are exploiting interdependence to inflict harm on their adversaries, and even on their allies. The increasing interaction across national boundaries that globalization entails, now produces not harmonization and cooperation, but friction and escalating trade and investment disputes.14 The Trump Administration is in the lead here, but it is not alone. Trade and investment friction with China is the most obvious and damaging example, precipitated by China’s long failure to conform to the World Trade Organization (WTO) principles, now escalated by President Trump into a trade and currency war disturbingly reminiscent of the 1930s that Bretton Woods was designed to prevent. Financial sanctions against Iran, in violation of US obligations in the Joint Comprehensive Plan Of Action (JCPOA), is another example of the rule of law succumbing to geopolitical competition. Though more mercantilist in intent than geopolitical, US tariffs on steel and aluminum, and their threatened use in automotives, aimed at the EU, Canada, and Japan,15 are equally destructive of the liberal system and of future economic growth, imposed as they are by the author of that system, and will spread to others. And indeed, Japan has used export controls in its escalating conflict with South Korea16 (as did China in imposing controls on rare earth,17 and as the US has done as part of its trade war with China). Inward foreign direct investment restrictions are spreading. The vitality of the WTO is being sapped by its inability to complete the Doha Round, by the proliferation of bilateral and regional agreements, and now by the Trump Administration’s hold on appointments to WTO judicial panels. It should not surprise anyone if, during a second term, Trump formally withdrew the US from the WTO. At a minimum it will become a “dead letter regime.”18

As such measures gain traction, it will become clear to states—and to companies—that a global trading system more responsive to raw power than to law entails escalating risk and diminishing benefits. This will be the end of economic globalization, and its many benefits, as we know it. It represents nothing less than the subordination of economic globalization, a system which many thought obeyed its own logic, to an international politics of zero-sum power competition among multiple actors with divergent interests and values. The costs will be significant: Bloomberg Economics estimates that the cost in lost US GDP in 2019- dollar terms from the trade war with China has reached $134 billion to date and will rise to a total of $316 billion by the end of 2020.19

Economically, the just-in-time, maximally efficient world of global supply chains, driving down costs, incentivizing innovation, spreading investment, integrating new countries and populations into the global system, is being Balkanized. Bilateral and regional deals are proliferating, while global, nondiscriminatory trade agreements are at an end. Economies of scale will shrink, incentivizing less investment, increasing costs and prices, compromising growth, marginalizing countries whose growth and poverty reduction depended on participation in global supply chains. A world already suffering from excess savings (in the corporate sector, among mostly Asian countries) will respond to heightened risk and uncertainty with further retrenchment. The problem is perfectly captured by Tim Boyle, CEO of Columbia Sportswear, whose supply chain runs through China, reacting to yet another ratcheting up of US tariffs on Chinese imports, most recently on consumer goods:

We move stuff around to take advantage of inexpensive labor. That’s why we’re in Bangladesh. That’s why we’re looking at Africa. We’re putting investment capital to work, to get a return for our shareholders. So, when we make a wager on investment, this is not Vegas. We have to have a reasonable expectation we can get a return. That’s predicated on the rule of law: where can we expect the laws to be enforced, and for the foreseeable future, the rules will be in place? That’s what America used to be.20

The international political effects will be equally damaging. The four structural forces act on each other to produce the more dangerous, less prosperous world projected here. Illiberal globalization represents geopolitical conflict by (at first) physically non-kinetic means. It arises from intensifying competition among powerful states with divergent interests and identities, but in its effects drives down growth and fuels increased nationalism/populism, which further contributes to conflict. Twenty-first-century protectionism represents bottom-up forces arising from economic disruption. But it is also a top-down phenomenon, representing a strategic effort by political leadership to reduce the constraints of interdependence on freedom of geopolitical action, in effect a precursor and enabler of war. This is the disturbing hypothesis of Daniel Drezner, argued in an important May 2019 piece in Reason, titled “Will Today’s Global Trade Wars Lead to World War Three,”21 which examines the preWorld War I period of heightened trade conflict, its contribution to the disaster that followed, and its parallels to the present:

Before the First World War started, powers great and small took a variety of steps to thwart the globalization of the 19th century. Each of these steps made it easier for the key combatants to conceive of a general war.

We are beginning to see a similar approach to the globalization of the 21st century. One by one, the economic constraints on military aggression are eroding. And too many have forgotten—or never knew—how this played out a century ago.

…In many ways, 19th century globalization was a victim of its own success. Reduced tariffs and transport costs flooded Europe with inexpensive grains from Russia and the United States. The incomes of landowners in these countries suffered a serious hit, and the Long Depression that ran from 1873 until 1896 generated pressure on European governments to protect against cheap imports.

…The primary lesson to draw from the years before 1914 is not that economic interdependence was a weak constraint on military conflict. It is that, even in a globalized economy, governments can take protectionist actions to reduce their interdependence in anticipation of future wars.

In retrospect, the 30 years of tariff hikes, trade wars, and currency conflicts that preceded 1914 were harbingers of the devastation to come. European governments did not necessarily want to ignite a war among the great powers. By reducing their interdependence, however, they made that option conceivable.

…the backlash to globalization that preceded the Great War seems to be reprised in the current moment. Indeed, there are ways in which the current moment is scarier than the pre-1914 era. Back then, the world’s hegemon, the United Kingdom, acted as a brake on economic closure. In 2019, the United States is the protectionist with its foot on the accelerator. The constraints of Sino-American interdependence—what economist Larry Summers once called “the financial balance of terror”—no longer look so binding. And there are far too many hot spots—the Korean peninsula, the South China Sea, Taiwan—where the kindling seems awfully dry.

Multipolarity

We can define multipolarity as a wide distribution of power among multiple independent states. Exact equivalence of material power is not implied. What is required is the possession by several states of the capacity to coerce others to act in ways they would otherwise not, through kinetic or other means (economic sanctions, political manipulation, denial of access to essential resources, etc.). Such a distribution of power presents inherently graver challenges to peace and stability than do unipolar or bipolar power configurations,22 though of course none are safe or permanent. In brief, the greater the number of consequential actors, the greater the challenge of coordinating actions to avoid, manage, or de-escalate conflicts. Multipolarity also entails a greater potential for sudden changes in the balance of power, as one state may defect to another coalition or opt out, and as a result, the greater the degree of uncertainty experienced by all states, and the greater the plausibility of downside assumptions about the intentions and capabilities of one’s adversaries. This psychology, always present in international politics but particularly powerful in multipolarity, heightens the potential for escalation of minor conflicts, and of states launching preventive or preemptive wars. In multipolarity, states are always on edge, entertaining worst-case scenarios about actual and potential enemies, and acting on these fears—expanding their armies, introducing new weapon systems, altering doctrine to relax constraints on the use of force—in ways that reinforce the worst fears of others.

The risks inherent in multipolarity are heightened by the attendant weakening of global institutions. Even in a state-centric system, such institutions can facilitate communication and transparency, helping states to manage conflicts by reducing the potential for misperception and escalation toward war. But, as Waheguru Pal Singh Sidhu argues in his chapter on the United Nations, the influence of multilateral institutions as agent and actor is clearly in decline, a result of bottom-up populist/nationalist pressures experienced in many countries, as well as the coordination problems that increase in a system of multiple great powers. As conflict resolution institutions atrophy, great powers will find themselves in “security dilemmas”23 in which verification of a rival’s intentions is unavailable, and worst-case assumptions fill the gap created by uncertainty. And the supply of conflicts will expand as a result of growing nationalism and populism, which are premised on hostility, paranoia, and isolation, with governments seeking political legitimacy through external conflict, producing a siege mentality that deliberately cuts off communication with other states.

Finally, the transition from unipolarity (roughly 1989–2007) to multipolarity is unregulated and hazardous, as the existing superpower fears and resists challenges to its primacy from a rising power or powers, while the rising power entertains new ambitions as entitlements now within its reach. Such a “power transition” and its dangers were identified by Thucydides in explaining the Peloponnesian Wars,24 by Organski (the “rear-end collision”)25 during the Cold War, and recently repopularized and brought up to date by Graham Allison in predicting conflict between the US and China.26

A useful, and consequential illustration of the inherent challenge of conflict management during a power transition toward multipolarity, is the weakening of the arms control regime negotiated by the US and the Soviet Union during the Cold War. Despite the existential, global conflict between two nuclear armed superpowers embracing diametrically opposed world views and operating in economic isolation from each other, the two managed to avoid worst-case outcomes. They accomplished this in part by institutionalizing verifiable limits on testing and deployment of both strategic and intermediate-range nuclear missiles. Yet as diplomatically and technically challenging as these achievements were, the introduction of a third great power, China, into this twocountry calculus has proven to be a deal breaker. Unconstrained by these bilateral agreements, China has been free to build up its capability, and has taken full advantage in ramping up production and deployment of intermediate-range ground-launched cruise missiles, thus challenging the US ability to credibly guarantee the security of its allies in Asia, and greatly increasing the costs of maintaining its Asian regional hegemony. As a result, the Intermediate Nuclear Force treaty is effectively dead, and the New Start Treaty, covering strategic missiles, is due to expire next year, with no indication of any US–Russian consensus to extend it. The US has with logic indicated its interest in making these agreements trilateral; but China, with its growing power and ambition, has also logically rejected these overtures. Thus, all three great powers are entering a period of nuclear weapons competition unconstrained by the major Cold War arms control regimes. In a period of rapid advances in technology and worsening great power relations, the nuclear competition will be a defining characteristic of the next decade and beyond. This dynamic will also complicate nuclear nonproliferation efforts, as both the demand for nuclear weapons (a consequence of rising regional and global insecurity), and supply of nuclear materials and technology (a result of the weakening of the nonproliferation regime and deteriorating great power relations) will increase.

Will deterrence prevent war in a world of several nuclear weapons states, (the current nuclear powers plus South Korea, Iran, Saudi Arabia, Japan, Turkey), as it helped to do during the bipolar Cold War? Some neorealist observers view nuclear weapons proliferation as stabilizing, extending the balance of terror, and the imperative of restraint, to new nuclear weapons states with much to fight over (Saudi Arabia and Iran, for example).27 Others,28 examining issues of command and control of nuclear weapons deployment and use by newly acquiring states, asymmetries in doctrines, force structures, and capabilities between rivals, the perils of variable rates in transition to weapons deployment, problems of communication between states with deep mutual grievances, the heightened risk of transfer of such weapons to non-state actors, have grave doubts about the safety of a multipolar, nuclear-armed world.29 We can at least conclude that prudence dictates heightened efforts to slow the pace of proliferation, while realism requires that we face a proliferated future with eyes wide open.

The current distribution of power is not perfectly multipolar. The US still commands the world’s largest economy, and its military power is unrivaled by any state or combination of states. Its population is still growing, despite a recent decline in birth rates. It enjoys extraordinary geographic advantages over its rivals, who are distant and live in far worse neighborhoods. Its economy is less dependent on foreign markets or resources. Its political system has proven—up to now—to be resilient and adaptable. Its global alliance system greatly extends its capacity to defend itself and shape the world to its liking and is still intact, despite growing doubts about America’s reliability as a security guarantor. Based on these mostly material and historical criteria, continued American primacy would seem to be a good bet, if it chooses to use its power in this way.30

So why multipolarity? The clearest and most frequently cited evidence for a widening distribution of global power away from American unipolarity is the narrowing gap in GDP between the US and China. The IMF’s World Economic Outlook forecasts a $0.9 trillion increase in US GDP for 2019–2020, and a $1.3 trillion increase for China in the same period.31 Many who support the American primacy case argue that GDP is an imperfect measure of power, that Chinese GDP data is inflated, that its growth rates are in decline while Chinese debt is rapidly increasing, and that China does poorly on other factors that contribute to power—its low per capita GDP, its political succession challenges, its environmental crisis, its absence of any external alliance system. Yet GDP is a good place to start, as the single most useful measure and long-term predictor of power. It is from the overall economy that states extract and apply material power to leverage desired behavior from other states. It is true that robust future Chinese growth is not guaranteed, nor is its capacity to convert its wealth to power, which is a function of how well its political system works over time. But this is equally the case for the US, and considering recent political developments is not a given for either country.

As an alternative to measuring inputs—economic size, political legitimacy, technological innovation, population growth—in assessing relative power and the nature of global power distribution, we should consider outputs: what are states doing with their power? The input measures are useful, possibly predictive, but are usually deployed in the course of making a foreign policy argument, sometimes on behalf of a reassertion of American primacy, sometimes on behalf of retrenchment. As such, their objectivity (despite their generous deployment of “data”) is open to question. What is undeniable, to any clear-eyed observer, is a real decline in American influence in the world, and a rise in the influence of other powers, which predates the Trump administration but has accelerated into America’s free fall over the last four years. This has produced a de facto multipolarity, whether explainable in the various measures of power—actual and latent—or not. This decline results in part from policy mistakes: a reckless squandering of material power and legitimacy in Iraq, an overabundance of caution in Syria, and now pure impulsivity. But more fundamentally, it is a product of relative decline in American capacity—political and economic—to which American leadership is adjusting haphazardly, but in the direction of retrenchment/restraint. It is highly revealing that the last two American presidents, polar opposites in intellect, temperament and values, agreed on one fundamental point: the US is overextended, and needs to retrench. The fact that neither Obama nor Trump (up to this point in his presidency) believed they had the power at their disposal to do anything else, tells us far more about the future of American power and policy—and about the emerging shape of international relations—than the power measures and comparisons made by foreign policy advocates.

Observation of recent trends in US versus Russian relative influence prompts another question: do we understand the emerging characteristics of power? Rigorously measuring and comparing the wrong parameters will get us nowhere at best and mislead us into misguided policies at worst. How often have we heard, with puzzlement, that Putin punches far above his weight? Could it be that we misunderstand what constitutes “weight” in the contemporary and emerging world? Putin may be on a high wire, and bound to come crashing down; but the fact is that Russian influence, leveraging sophisticated communications/social media/influence operations, a strong military, an agile (Putin-dominated) decision process, and taking advantage of the egregious mistakes by the West, has been advancing for over a decade, shows no sign of slowing down, and has created additional opportunities for itself in the Middle East, Europe, Asia, Latin America, the Arctic. It has done this with an economy roughly the size of Italy’s. There are few signs of a domestic political challenge to Putin. His external opponents are in disarray, and Russia’s main adversary is politically disabled from confronting the problem. He has established Russia as the Middle East power broker. He has reached into the internal politics of his Western adversaries and influenced their leadership choices. He has invaded and absorbed the territory of neighboring states. His actions have produced deep divisions within NATO. Again, simple observation suggests multipolarity in fact, and a full explanation for this power shift awaiting future historians able to look with more objectivity at twenty-first-century elements of power.

When that history is written, surely it will emphasize the extraordinary polarization in American politics. Was multipolarity a case of others finding leverage in new sources of power, or the US underutilizing its own? The material measures suggest sufficient capacity for sustained American primacy, but with this latent capacity unavailable (as perceived, I believe correctly, by political leadership) by virtue of weakening institutions: two major parties in separate universes; a winnertake-all political mentality; deep polarization between the parties’ popular bases of support; divided government, with the Presidency and the Congress often in separate and antagonistic hands; diminishing trust in the permanent government, and in the knowledge it brings to important decisions, and deepening distrust between the intelligence community and policymakers; and, in Trump’s case, a chaotic policy process that lacks any strategic reference points, mis-communicates the Administration’s intentions, and has proven incapable of sustained, coherent diplomacy on behalf of any explicit and consistent set of policy goals.

Rising Nationalism/Populism/Authoritarianism

The evidence for these trends is clear. Freedom House, the go-to authority on the state of global democracy, just published its annual assessment for 2020, and recorded the fourteenth consecutive year of global democratic decline and advancing authoritarianism. This dramatic deterioration includes both a weakening in democratic practice within states still deemed on balance democratic, and a shift from weak democracies to authoritarianism in others. Commitment to democratic norms and practices—freedom of speech and of the press, independent judiciaries, protection of minority rights—is in decline. The decline is evident across the global system and encompasses all major powers, from India and China, to Europe, to the US. Right-wing populist parties have assumed power, or constitute a politically significant minority, in a lengthening list of democratic states, including both new (Hungary, Poland) and established (India, the US, the UK) democracies. Nationalism, frequently dismissed by liberal globalization advocates as a weak force when confronted by market democracies’ presumed inherent superiority, has experienced a resurgence in Russia, China, the Middle East, and at home. Given the breadth and depth of right-wing populism, the raw power that promotes it—mainly Russian and American—and the disarray of its liberal opponents, this factor will weigh heavily on the future.

The major factors contributing to right-wing populism and its global spread is the subject of much discussion.32 The most straightforward explanation is rising inequality and diminished intergenerational mobility, particularly in developed countries whose labor-intensive manufacturing has been hit hardest by the globalization of capital combined with the immobility of labor. Jobs, wages, economic security, a reasonable hope that one’s offspring has a shot at a better life than one’s own, the erosion of social capital within economically marginalized communities, government failure to provide a decent safety net and job retraining for those battered by globalization: all have contributed to a sense of desperation and raw anger in the hollowed-out communities of formerly prosperous industrial areas. The declining life expectancy numbers33 tell a story of immiseration: drug addition, suicide, poor health care, and gun violence. The political expression of such conditions of life should not be surprising. Simple, extremist “solutions” become irresistible. Sectarian, racial, regional divides are strengthened, and exclusive identities are sharpened. Political entrepreneurs offering to blow up the system blamed for such conditions become credible. Those who are perceived as having benefited from the corrupt system—long-standing institutions of government, foreign countries and populations, immigrants, minorities getting a “free ride,” elites—become targets of recrimination and violence. The simple solutions of course, don’t work, deepening the underlying crisis, but in the process politics is poisoned. If this sounds like the US, it should, but it also describes major European countries (the UK, France, Italy, Germany, Poland, Hungary, the Czech Republic), and could be an indication of things to come for non-Western democracies like India.

We have emphasized throughout this chapter the interaction of four structural forces in shaping the future, and this interaction is evident here as well. Is it merely coincidence that the period of democratic decline documented by Freedom House, coincides precisely with the global financial and economic crisis? Lower growth, increasing joblessness, wage stagnation, superimposed on longer-term widening of inequality and declining mobility, constitute a forbidding stress test for democratic systems, and many continue to fail. And if we are correct about secular stagnation, the stress will continue, and authoritarianism’s fourteen-year run will not be over for some time. The antidemocratic trend will gain additional impetus from the illiberal direction of globalization, with its growth suppressing protectionism, weaponization of global economic exchange, and weakening global economic institutions. Multipolarity also contributes, in several ways. The former hegemon and author of globalization’s liberal structure has lost its appetite, and arguably its capacity, for leadership, and indeed has become part of the problem, succumbing to and promoting the global right-wing populist surge. It is suffering an unprecedented decline in life expectancy, and recently a decline in the birth rate, signaling a degree of rot commonly associated with a collapsing Soviet Union. While American politics may once again cohere around its liberal values and interests, the time when American leadership had the self-confidence to shape the global system in its liberal image is gone. It may build coalitions of the like-minded to launch liberal projects, but there will be too much power outside these coalitions to permit liberal globalization of the sort imagined at the end of the Cold War. In multipolarity, the values around which global politics revolve will reflect the diversity of major powers, their interests, and the norms they embrace. Convergence of norms, practices, policies is out of the question. Global collective action, even in the face of global crises, will be a long shot. To expect anything else is fantasy

Unbrave New World and Future Challenges

At the outset of this chapter we described these structural forces as interacting to produce more conflict and diminished prosperity. We also predicted a world with shrinking collective capacity to address new challenges as they arise. What specifically will such a world look like? We address below three principal challenges to global problem solving over the next decade.

Interstate Conflict

In the world experienced by most readers of this volume, conflict is observed within weak states, sometimes promoted by regional competitors, by terrorist groups, or by great powers, acting through surrogates or by indirect means. Sometimes, as in Syria, this conflict spills over to contiguous states and contributes to regional instability, and challenges other regions to respond effectively, a challenge that Europe has not met. Much of this will continue, but the global significance of such local conflicts will be greatly magnified by increasing great power conflict, which will feed—rather than manage or resolve—local instabilities and will in turn be exacerbated by them. Great powers will jockey for advantage, support their local partners, escalate preemptively. Conflicts initially confined to failing states or unstable regions will be redefined by great powers as global in scope and significance.

This tendency of states to view local conflicts in the context of a zero-sum, global struggle for power is familiar to students of the Cold War, but now with the additional challenges to collective action, expanded uncertainty and worst-case thinking associated with the power transition to multipolarity. We can easily observe increased conflict in US–China relations, as we will in US–Russia relations as future US administrations try to make up for ground lost during the Trump presidency, especially in the Middle East. We can observe it among powerful states with mutual historical grievances, now with a weakening presence of the hegemonic security guarantor and having to consider the renationalization of their defense: Japan-South Korea, Germany-France. We can observe it among historical rivals operating in rapidly changing security landscapes: India-China. We can observe it within the Middle East, as internal rivalries are appropriated by regional powers in a contest for regional dominance. We can observe it clearly in Syria, where the regime’s violent suppression of Arab Spring resistance led to all-out civil war, attracted outside support to proxy forces by aspiring regional hegemons Saudi Arabia and Iran, enabled the rise of ISIS, and eventually to great power intervention, principally by Russia. In a world of effective great power collaboration or American primacy, the Syrian civil war might have been settled through power sharing or partition, or if not, contained within Syria. The collapse of Yugoslavia, occurring during a period of US “unipolarity” and managed effectively, demonstrates the possibilities. Instead, with the US retrenching, Middle East rivals unconstrained by great powers, and great power competition rising, the Syria civil war was fed by outside powers, then metastasized into the region, and—in the form of refugee flows—into Europe, fundamentally altering European politics. Libya may be at the early stages of this scenario.

This is not the end of the Syria story. Russia has established itself as a major player in Syria and the Middle East’s power broker, the indispensable country with leverage throughout the region. China is poised to reap the financial and power benefits of Syrian reconstruction. The US has just demonstrated, in its act of war against the Iranian regime, its willingness, without consultation, to put its allies’ security in further jeopardy, accentuating the risks of security ties with Washington and generating added opportunities for Russia and China. The purpose here is not to critique US policy, but to point out the dramatically shifting power balance in a critical region, toward multipolarity. The dangers of such a shift will become apparent as some future US president attempts to reassert US influence in the region and finds a crowded playing field.

Can a multipolar distribution of power among several states whose interests, values, and political practices are divergent, all experiencing bottom-up nationalist pressures, all seeking advantages in the oversupply of regional instability, be made to work? I think not. Will this more dangerous world descend into direct military confrontation between great powers, and could such confrontation lead to use of nuclear weapons? Here the question becomes, what will this more dangerous world actually look like; what instruments of coercion will be available to states as technology change accelerates; how will states employ these instruments; how will deterrence work (if at all) among several states with large but unequal levels of destructive capacity, weak command, and control, disparate— or opaque—strategies and simmering rivalries; can conflict management work in a world of weak institutions? The collapse of the Cold War era nuclear arms control regime, the threat to the Non-Proliferation Treaty represented by the demise of the JCPOA, and multiple indications of an accelerating nuclear arms race among the three principle powers, augurs badly. Given the structural forces at play, and without predicting the worst, we are indeed entering perilous times.

Global Poverty and Inequality

Despite the challenges of volatility and disruptive change inherent in globalization, the world under American liberal leadership has managed a dramatic reduction of extreme poverty. According to World Bank estimates, in 2015, 10 percent of the world’s population lived on less than $1.90 a day, down from nearly 36 percent in 1990.34 In fact, as of September 2018, half the world is now middle class or wealthier.35 The uneven success of the UN Millennium Development Goals (MDGs) exemplifies this achievement, and demonstrates what is possible when open markets are managed through strong global institutions, effective leadership and interstate collaboration. What this liberal hegemonic system did not achieve, however, was a fair distribution of the gains from globalization within states, and among those states that for various reasons were not full participants in this system.

This record of partial achievement leaves us with a full agenda for the next fifteen years, but without the hegemonic leadership, strong institutions, ascendant liberalism or robust global growth that enabled previous gains. There are powerful reasons to question the sustainability of these poverty reduction gains, leading to doubts about the realization of the Sustainable Development Goals, which have replaced the MDGs as global development targets.36 (See Jens Rudbeck’s chapter and Sidhu’s UN chapter for SDGs). Skeptics have pointed to slowing global growth, specifically in China, whose demand for imported commodities was a major factor in developing country growth and job creation; growing protectionism in developed country markets, fueled by bottom-up forces of nationalism, and from top-down by a weakened global trading regime and increased geopolitical rivalry; the effects of accelerating climate change on agriculture, migration and communal conflict in poor countries; and the growth burst among poor countries from the rapid transition to more efficient use of resources, a transition that is now slowing down.37

Perhaps the greatest concern in this scenario is a general deterioration in the developing country foreign investment climate. Foreign direct investment (FDI) has been a major contributor to growth, job creation, and poverty alleviation among poor countries. It has incentivized growthfriendly policies, reduced corruption, introduced technology and effective management practices, and linked poor countries to foreign markets through global supply chains.38 It has stimulated growth of indigenous manufacturing and service companies to supply new foreign investments.

It has been the major cause of economic convergence between rich and poor countries. From 2000 to 2009, developing economies’ growth rates were more than four percentage points higher than those of rich countries, pushing their share of global output from just over a third to nearly half.39 However, FDI flows into poor countries are imperiled by the structural forces discussed here. Political instability arising from slower growth and environmental stress will increase investors’ perception of higher risk, reinforcing their developed country bias. Protectionism among developed countries will threaten the global market access upon which manufacturing investment in developing countries is premised, causing firms to pare back their global supply chains. As companies retrench from direct investment in poor countries, the appeal to those countries of Chinese debt financed infrastructure projects, under the Belt-Road Initiative with little or no conditionality, but at the risk of “debt traps,” will increase.

Global Warming

The question posed at the beginning of this section is whether the international system, evolving toward multipolarity and rising nationalism, will find the collective political capital to confront challenges as they arise. Global warming is the mother of all challenges, and the weakness in the system’s capacity to respond is clear. With the two major political/economic powers and greenhouse gas emitters locked in deepening geopolitical conflict (and with one of them locked in climate change denial, possibly through 2024), the chances of significantly slowing global warming or even ameliorating its effects are very slim. We are reduced to the default option, nation-specific adaptation to climate change, which will impose rising human, political and economic costs on all, and will widen the gap between rich countries with adaptive capacity (of varying degrees), and the poor, who will suffer deteriorating economic, political, and social conditions. (For a contrary, optimistic view see Michael Shank’s chapter, which credits new actors—like cities—as playing a more constructive role in climate mitigation.) This would bring to a close liberal globalization’s greatest achievement; the raising of 1.1 billion people out of extreme poverty since 1990,40 with all its associated gains in quality of life (in the WHO Africa region, for example, life expectancy rose by 10.3 years between 2000 and 2016, driven mainly by improvements in child survival and expanded access to antiretrovirals for treatment of HIV).41

Several forces are at work here. The problem itself is graver—in magnitude and in rate of worsening—than predicted by climate scientists. The UN Intergovernmental Panel on Climate Change (IPCC), the major source of information on global warming, has consistently underpredicted the rate of climate deterioration. This holds true even for its “worst-case scenarios,” meaning that what was meant as a wake-up call has in fact reinforced complacency.42 (see Michael Shank’s chapter for further discussion of climate change). The IPCC, in its 2019 report, has tried to undo the damage by emphasizing the acceleration in the rate of warming and its effects, the only partially understood dynamic of climate change, and—given wide uncertainty—the possibility of unpleasant surprises yet to come. This strengthens the scientific case for urgency—to both severely limit greenhouse gas emissions, and to increase investment in ameliorating the effects.

Unfortunately, the crisis comes at a moment when the climate for collective action is ice cold. Geopolitical competition incentivizes states to out produce each other, regardless of the environmental effects. Multipolarity complicates collective action. Economic stagnation mandates job creation, making regulation politically toxic. Bottom-up nationalism/populism causes states to pursue “relative gains,” meaning that if the nation is seen as gaining in a no-holds-barred economic competition with others, the negative environmental effects can be tolerated. A post-Trump presidency would help, with the US rejoining the Paris Agreement, and lending its weight to tighter regulation, increased R and D, and stronger economic incentives to reduce carbon emissions. Keep in mind, however, that President Obama was fully behind such efforts, but in a deeply polarized America was unable to implement measures needed to fulfill the Paris obligations through legislation, and his executive orders to do this were swiftly overturned by Trump.

Conclusion

It may be tempting to hope that post-Trump, the US can regain its global leadership and exert its considerable power in a liberal direction, but with enough self-awareness of its relative decline to share responsibility with others. This was, I believe, the broad direction of the Obama strategy, evidenced by the JCPOA and the Trans-Pacific Partnership: liberal, collective solutions to global problems, as US dominance receded.

This would constitute an optimistic scenario, and it confronts two major problems: can US internal politics support it (can, for example, the country legislate controls on carbon, essential for the global credibility and durability of such commitments); and is the world ready to reengage with American leadership, given the damage to its reputation and the structural forces discussed in this chapter?

My educated guess is no, on both counts. The rot within is extensive, the concrete evidence clear in the economic inequality/immobility numbers, the life expectancy numbers, the deep political polarization, between the two major parties, between regions, between cities and rural areas. We are in fact a long way from fitness for global leadership, and the recognition of this by others will accelerate the decline of American influence. The rest of the world is well on its way toward adjusting to post-American hegemony, some by renationalizing their defense, or by cutting deals with adversaries, by building new alliances or by seizing new opportunities for influence in the vacuum left by American retrenchment. The evidence for this will accumulate. Observe the current and emerging Middle East, where all these post-hegemonic strategies are visible.

#### The AFF is key:

#### 1. Kronos effect---unlimited vertical acquisitions destroys lagging and entrant firms.

Kevin A. Bryan & Erik Hovenkamp 20, Assistant Professor, University of Toronto Rotman School of Management; Assistant Professor, University of Southern California Gould School of Law, "Reassessing the Chicago School of Antitrust Law: Startup Acquisitions, Error Costs, and Antitrust Policy," University of Chicago Law Review, Vol. 87, No. 331, March 2020, Lexis.

[\*331] INTRODUCTION

High tech industries are not only lucrative, but also highly innovative and dynamic. Large firms are not their sole source of innovation, however. Many valuable technologies are first developed by startup companies. Innovative startups are frequently acquired by powerful incumbents at an early stage. Well-known examples include acquisitions of WhatsApp and Instagram by Facebook; Waze and DoubleClick by Google; and GitHub and LinkedIn by Microsoft. These cases have drawn very little antitrust scrutiny, leading many commentators to question whether antitrust is in need of reform.

[\*332] This paucity of meaningful oversight is driven by uncertainty about a startup's future impact on the marketplace. Merger enforcement is usually directed at proposed combinations of large, established firms. It largely focuses on the estimated immediate effect of the proposed deal on competition. But startups are new and comparatively small, leaving little data with which to estimate competitive effects. Further, the relevant antitrust concerns relate mainly to more speculative effects on future competition. Rather than taking calculated steps to balance such uncertainties against the potential benefits of enforcement, antitrust policy has maintained a rigid policy of near-universal inaction.

This result is emblematic of a broader principle often associated with the influential Chicago School of antitrust thought, which has had significant influence on the Supreme Court in recent decades. This principle holds that antitrust should err on the side of nonintervention (false negatives), because erroneous condemnations (false positives) are seen as more socially costly. A leading rationale is that competitive entry will discipline anticompetitive behavior organically, whereas the adverse effects of erroneous intervention will persist indefinitely. This view has [\*333] spurred very demanding evidentiary requirements, making it difficult for plaintiffs to prevail in most kinds of antitrust cases.

Thus, considering the uncertainties they present, it is unsurprising that startup acquisitions have received very little antitrust scrutiny. However, a growing body of economic theory and empirics identifies various harmful effects from such acquisitions. Over time, they can expand the technological gap between industry leaders and "laggards" (smaller or less successful rivals). The product market is thus left less competitive and more concentrated. Startups are sometimes acquired by dominant firms solely to exclude rivals from accessing such technologies. In addition, incentives for innovation may also be adversely affected, as they are influenced in part by the prospect of future acquisitions. An innovator's decisions about what lines of research to invest can become skewed. Further, once a habitual acquirer becomes sufficiently dominant, its willingness to pay for new technologies falls, reducing the returns innovators receive for future inventions. Incentives for prospective startups to innovate are thus weakened.

To be sure, in most startup acquisitions, it is probably not possible to precisely predict the transaction's but-for impact on commerce. But it does not follow that society is best served by a policy that permits dominant incumbents to acquire all promising startups soon after they form. These acquisitions may have significant adverse effects in the aggregate, even if it is difficult to [\*334] assess how any particular transaction would influence the marketplace. Consequently, society may benefit from a policy that permits limited intervention based on reasonably ascertainable evidence, even if this carries some risk of false positives.

The traditional argument favoring false negatives is particularly ill-suited to this setting. There is a clear circularity problem. The driving force behind the error cost argument, competitive entry, is directly threatened by the conduct in question. One cannot expect potential entrants to discipline anticompetitive behavior if they are consistently absorbed by powerful incumbents. When the market leader is sufficiently dominant, it is generally most profitable (for both the leader and the startup) for technology rights to be sold exclusively to the leader. This softens competition by increasing the leader's technological advantage over its competitors. Hence there is no reason to expect that the market will self-correct the problem, as it is more profitable than the alternative.

This Essay is organized as follows. In Part I, we address the error cost argument and some subsequent rebuttals. Part II addresses the potential harms from startup acquisitions by dominant incumbents, provides supporting empirical evidence, and explains why current merger policy is unlikely to provide a satisfactory solution. In Part III we argue that expanded (albeit limited) intervention in startup acquisitions is likely to be beneficial, and that the traditional error cost argument holds little weight in such cases. We also discuss reasonable indicia for the likelihood of harm and potential remedies that might be implemented in practice. We conclude by noting that the error cost argument for nonintervention may be inappropriately applied in many other settings as well.

I. ERROR COSTS AND MARKET ENTRY

The complexities of antitrust are often difficult for courts to manage in practice, and it is important for antitrust policy to be mindful of these administrative limitations. Further, the effects [\*335] of an antitrust judgment are often felt by many parties not before the court, including most or all consumers in the relevant market. Thus, a natural question is what costs arise from different types of judicial errors and how they compare. Beginning in the late twentieth century, many scholars set upon this question of "error costs" in antitrust, leading to a widespread view that it is far less harmful to condone an anticompetitive practice than to condemn an efficient one.

The argument's driving premise is that false negatives will generally self-correct over time, while false positives will not. The most commonly given justification for this is that anticompetitive behavior invites new competition, attracted by the prospect of dethroning a high-priced incumbent. After all, higher prices are the easiest to undercut while still turning a profit. Entry is thus more enticing when the relevant market is less competitive, all else being equal.

This argument is complemented by the theory of market "contestability," introduced by Professor William Baumol in the 1980s. This theory describes markets that are concentrated but nevertheless competitive because the prospect of entry deters incumbents from setting supra competitive prices. Such a market [\*336] is said to be "contestable." This theory, when applicable, suggests that market concentration does not necessarily preclude competitive behavior. While arguments about contestability are typically not focused on error costs specifically, they give some theoretical support to the proposition that markets will end up correcting anticompetitive behavior over time, which is the core premise behind the traditional error cost argument.

A. Strategic Behavior and Entry

Arguments about market power and contestability have advanced greatly among economists since the price-theoretic intellectual origins of the Chicago School. In a price-theoretic world, conduct is a function of demand, cost, and technology. Firms are ex ante identical in the technology they can access and can freely enter markets. Excess profits therefore attract entry. Regulatory error that improperly permits anticompetitive behavior by a firm will lead to excess profits, hence pressure to enter. The welfare harms of error, it is argued, are thus less than they appear.

The problem with this argument is that it abstracts away from strategic interactions among the incumbent and the entrant. If a new firm is considering entry, the incumbent may [\*337] have effective deterrence strategies. In particular, the incumbent may be able to invest ex ante in ways that limit entry. Further, the actions necessary to deter entry may harmfully distort socially useful incumbent investments, such as research and development (R&D), product variety, or product diffusion. We treat each of these cases in turn.

To deter entry by inducing potential entrants to believe that competition with the monopolist will be unprofitable requires the post-entry deterrence action to be credible. In the traditional Chicago School framework, entrants have access to the same technology, with the same economies of scale as incumbents. If actions or investments pre-entry do not affect the nature of competition post-entry, then strategic interaction is irrelevant: the extent of entry is affected only by static post-entry profits available to the entrant.

However, there are a number of straightforward strategies by which an incumbent can leverage its market power to deter entry. Incurring sunk costs that lower future marginal costs causes the post-entry price to fall, and can therefore make entry unprofitable. Network effects and other switching costs that require coordination across buyers permit short-term monopoly to turn into long-term deterred entry. Empirically, firms do appear [\*338] to react to the threat of entry in a different manner than they do to realized entry. For instance, when Southwest begins flying from A to B when they are already flying from B to C, the probability they begin flying to A to C thereafter increases as there are economies of scope across routes. The mere threat of those flights induces competitors to drop prices on A to C. When Southwest has already committed to fly from A to C, there is no such preemptive price cutting. Similar entry-deterring actions have been shown in excess advertising of soon-to-expire pharmaceuticals, and in new procedures among potentially competing hospitals. That is, a model in which firms can never credibly deter entry, and hence one in which regulatory error leading to excess profits necessarily attracts entry, reflects neither modern economic theory nor recent empirical results.

Indeed, welfare is harmed not only when entry is successfully deterred, but also from the inefficiencies that arise as firms attempt to deter entry. Efficient markets are not only about price, but also about the amount of innovation, the extent of variety available to consumers, the compatibility of products in a network setting, and the avoidance of "money burning" through zero-sum advertising.

For instance, consolidation in radio permitted by the Telecommunications Act of 1996 led new station entry to fall. Why were entrants not attracted by the more profitable consolidated entry? Newly merged incumbents modestly differentiated their [\*339] station offerings in such a way that, despite consolidation, there were no profitable "holes" in the variety spectrum worth incurring the fixed cost of setting up a station. The particular station variety was chosen for entry deterrence reasons rather than efficiency concerns. Attempts to deter not only reduce welfare by successfully deterring competition, but further harm welfare by pushing incumbents away from the most efficient product variety mix. As we discuss in the following Part, similar dual long-run harms of market power--less entry with its consequent higher prices, and distortions along other margins used by the incumbent to prevent entry--occur in highly innovative industries.

II. STARTUP ACQUISITIONS AND ANTITRUST POLICY

In this Part we consider the potential justifications for limited antitrust intervention in startup acquisitions. Throughout our analysis, we focus on cases where the acquiring firm is highly dominant within a relevant product market, meaning that it has significant market power.

A. Potential Harms

Antitrust usually focuses on potential injuries to competition. Such considerations are certainly relevant in the present context, too, but they do not tell the full story. Startups are typically innovative enterprises, and potential acquisitions may thus play an important role in the innovation and entrepreneurship processes. We address both sets of considerations below. The next Section then discusses empirical findings that shed light on these potential harms.

1. Diminished competition.

An important question in antitrust treatment of mergers and acquisitions is whether the proposed combination is "horizontal" or "vertical." The two types of mergers receive different treatment [\*340] because they involve different theories of harm and different potential justifications. In the horizontal case, the startup is a new or potential competitor of the acquiring incumbent. In this case, the potential antitrust concerns are more salient, for the acquisition necessarily forestalls competitive entry. Indeed, the acquirer may have no interest in actually using the startup's technology; it may simply wish to prevent such technology from reaching the marketplace.

For example, in Federal Trade Commission v Mallinckrodt ARD, Inc, the defendant was initially a monopolist in the market for adrenocorticotropic hormone drugs used for the treatment of infantile spasms. It outbid potential rivals to acquire the domestic rights to the lone competing product, named Acthar, which had not previously been marketed in the United States. The Federal Trade Commission (FTC) brought suit and succeeded in securing a stipulated judgment in which the defendant would be required to license the rights to sell Acthar to a competing US manufacturer, in addition to paying a $ 100 million fine.

However, Mallinckrodt is a somewhat rare case in which the relevant acquisition target was obviously a prospective direct competitor in a clearly defined market that was otherwise utterly dominated by the acquirer. In practice, matters are rarely this clear-cut. First, in many instances, the startup presently offers only a technology that is complementary to the acquirer's product. [\*341] Although it may be quite plausible that the startup would eventually have entered the acquirer's product market (or vice versa) but for the acquisition, this may be impossible to prove as of the acquisition date. Second, the startup's technology may be complementary in some respects but substitutable in others, making it hard to say whether it should be regarded as a competitor. This challenge is particularly salient in high tech sectors, where it is often difficult to define markets.

For these reasons, many startup acquisitions will be presumptively vertical in the sense that they are not provably horizontal. However, in such cases there may still be a material risk of anticompetitive harm if the acquisition prevents the acquirer's rivals from obtaining access to a promising new technology developed by the startup. That is, if the acquirer is dominant in its product market, then its motivation for the acquisition may be (in whole or in part) to exclude its smaller rivals from gaining access to the startup technology. This prevents rivals from improving their own products, thereby extending the acquirer's market advantage relative to a scenario in which several or all incumbents obtain the rights to the startup technology.

Consider a simple example. Suppose there is a dominant leader in the market for smartphones and that the startup technology is an improved processor for mobile devices. The leader maintains an advantage due to the fact that its smartphone is technologically superior in some respects, and/or because it has lower production costs. We can think of either possibility as an advantage in terms of quality. Any smartphone producer can improve its product quality by utilizing the new processor; all else [\*342] being unchanged, this quality improvement would increase demand for this producer's smartphone. Suppose that, before the startup emerges, the leader is using a processor that is at least as good as that used by the laggards (which would at least partially explain why this firm is the leader to begin with). Then, if the leader and laggards all obtain access to the new processor, the rivals will partially catch up to the leader: their own smartphones improve in quality by incrementally more. By leaving the firms on more equal footing, this would make the market more competitive and less concentrated.

By contrast, if the dominant smartphone producer obtains exclusive rights to the startup and then declines to license the processor technology (or to sell the processors wholesale) to its smaller rivals, then it will increase its market dominance. Its own quality level improves, but its rivals' do not. This leaves the market less competitive. Rivals' smartphones now look comparatively worse to consumers, leading these firms to apply less competitive pressure and hence permitting the dominant firm to behave more like a pure monopolist. The result is that static consumer surplus is lower (perhaps significantly so) than if the acquirer's rivals had also obtained access to the startup technology.

As in the pure horizontal case, the acquisition may serve no purpose other than to forestall an increase in competition. That is, the acquirer itself may derive little or no value from using the startup technology itself, perhaps because it is already using a comparable (or superior) alternative technology. However, smaller rivals may still benefit from using it, and the acquirer may purchase the exclusive rights to prevent them from obtaining access to it. In this case, there is no static welfare improvement from the acquisition, since the startup technology is simply not used. Consumers are thus worse off than they would have been if [\*343] rivals had been able to utilize the new processor. First, the market is less competitive than it would have been, and hence output is lower. Second, consumers get comparatively less value (net of price) from rivals' smartphones, as the rivals have been denied access to the improved processor.

Of course, if the dominant firm's rivals would benefit from the startup technology, then they are willing to pay the startup for the right to use it. So how do we know that the dominant firm is willing to outbid them and thereby obtain exclusive rights? The most likely explanation is that it is generally more profitable in the aggregate to soften competition than to invigorate it. By extension, it is more profitable to preserve or increase the dominant firm's market power than to enable smaller rivals to catch up. The result is that a dominant firm is generally willing to pay more to exclude rivals than such rivals would pay to gain ground on the dominant firm.

2. Innovation incentives.

The concern about competition, therefore, is that startups will sell their technology to industry leaders rather than to lagging incumbents even when the sale to laggards benefits consumers by increasing the competitiveness of the product market. A traditional error cost argument does not sit well here: the anticompetitive action is one that simultaneously limits the emergence of innovative new firms.

There are further harms beyond reduced competition. Consider the decision problem of an innovator (a prospective startup, for instance) in deciding what kind of new technology to invest in. Some technologies improve the quality of all incumbents' products, such as flexible or unbreakable smartphone glass. Others merely reduce the technological gap between leaders and laggards, like giving smartphone manufacturers an alternative to an [\*344] otherwise patented technology held by the market leader. Both types of invention improve consumer welfare: the former from directly improving the quality of all products, and the latter from inducing more competitive pricing behavior by reducing vertical differentiation.

Nonetheless, the startup who can license freely is always biased against producing inventions that only help the laggards catch up. The purchase price or licensing fee charged by the startup depends on how the use of such invention would influence competition and industry profits, since these things determine an incumbent's willingness to pay to use a new technology. Inventions that improve all firms' technologies, when bought exclusively by the industry leader, directly benefit consumers while also increasing differentiation between the leader and laggard. The second effect can be strong enough that industry joint profits are highest when only the leader possesses the new technology. However, an invention that only helps the laggard catch up increases competition without directly pushing the quality frontier forward. The industry leader may buy this technology solely to prevent this greater competition, but the purchase price will be lower than that of inventions that also increase the leader's product quality.

Unrestricted startup acquisition, therefore, both makes it harder to compete against strong incumbents and distorts the direction of invention. Things get worse dynamically. As the market leader ingests startups and startups shift their research effort toward technology that helps the leader pull away from its competitors, lagging incumbents will exit. As the number of competing firms falls, the purchase price for startups also falls: The threat to sell to a firm's competitor improves the startup's bargaining power, and such bargaining power diminishes as there are fewer competitors. This fall in purchase price therefore decreases the [\*345] incentive of startups to innovate, directly reducing productivity growth.

B. Recent Empirical Research

Numerous recent articles cast light on the potential adverse effects that may result from a laissez-faire policy toward startup acquisitions. The antitrust concerns are perhaps most salient when an acquisition is motivated purely by the desire to forestall new competition. To that end, one recent paper finds that, in the pharmaceutical industry, numerous innovative new firms are effectively terminated through "killer acquisitions" by incumbent firms. In these acquisitions, the acquirer does not utilize or further develop the target's innovation, but instead merely prevents such innovation from entering into competition with the incumbent's own product.

Although a startup is typically small, the economic effects of startup acquisitions may accumulate over time. To that end, another study finds that, following a statutory amendment that weakened the reporting requirements for prospective mergers, there was an increase in "anticompetitive deals whose individual size enables them to escape regulatory scrutiny but whose cumulative effect is large." Indeed, between 1994 and 2011, "submarine" acquisitions of firms below the Hart-Scott-Rodino Act reporting limit cumulatively consolidated $407 billion in annual US output, equivalent to a 30 percent increase in four firm industry concentration.

[\*346] One important concern is that dominant incumbents will acquire promising new technologies and then decline to license the relevant technologies to smaller rivals. Such a pattern would act to strengthen the market power of the dominant acquirer over time as its technological advantage grows. In this vein, several recent studies find that there is a widening gap between market leaders and laggards. Additionally, a number of articles have purported to find evidence that markets are generally growing more concentrated over time, although this result has not been causally linked to acquisitions or any other specific practices.

The preceding studies bear principally on potential harms to price competition, taking firms' technologies as given. But one may also be concerned about the potential impact of acquisitions on innovative activity, particularly in cases where the startup is an innovative enterprise. One recent study, relying on data from the pharmaceutical industry, finds that mergers generally lead to diminished R&D activity by both the merged firm and incumbent rivals. An additional study finds that incumbents may rely on acquisitions of innovative startups as a substitute for conducting R&D internally. Another study finds that, while innovative new [\*347] entrants have historically played an important role in market productivity and growth, this trend has started to diminish.

C. Limitations of Contemporary Merger Enforcement

The current state of antitrust merger enforcement makes it very difficult to bring a viable challenge against a startup acquisition, even if the acquirer is highly dominant. Here we briefly discuss some of the principal reasons for this. First, merger enforcement usually relies chiefly on estimates of the price effects that would result immediately following a merger. Put differently, merger analysis is static, generally declining to form predictions about how today's transaction will affect competition tomorrow. Static price effects are estimated using established data on firm characteristics and behavior, such as market shares and pricing activity. But a startup is a new player that usually does not presently have a significant market share. Thus, a static analysis will typically suggest that there is no potential harm, but this may only be because the relevant anticompetitive threat involves diminished future competition.

Relatedly, evaluating potential effects on future competition is necessarily more speculative than the analysis of mergers between established firms, where one can reasonably focus on static effects. This makes it much harder (if not impossible) to rely on rigorous empirical methods to estimate anticompetitive effects. Antitrust facially recognizes the elimination of "potential competition" as a basis for intervention, but in practice this kind of [\*348] claim is quite difficult to bring successfully and is rarely attempted. (The Mallinckrodt case was a rare exception. )

Second, as noted above, many acquisitions will not be provably horizontal, even if it is quite plausible that the startup would have gone on to enter the acquirer's product market (or vice versa). In that case, an antitrust plaintiff must attack the acquisition as a vertical merger. However, antitrust has grown increasingly hostile toward vertical merger challenges, leaving very little chance of success even if both parties to the merger are large, established incumbents. Combining this with the general dearth of useful data in startup acquisition cases, it is hard to see a viable path to enforcement without some departure from current judicial treatment of vertical mergers.

In sum, current enforcement policy demands more precise economic evidence than can typically be supplied in cases involving startup acquisitions. While there is good reason to believe that persistent acquisitions by dominant incumbents may produce harmful effects in the aggregate, it is often difficult to establish this in any individual case under the existing standards of merger review.

III. ADMINISTRATION UNDER UNCERTAINTY

The traditional error cost argument implicitly treats the prospect of competitive entry as an external, immutable force that persists independently of a defendant's conduct. But, at minimum, this characterization is inapt when the relevant conduct involves persistent acquisitions of newly formed firms with promising new [\*349] technologies. Innovative new entrants will not challenge dominant incumbents--or aid smaller rivals in doing so--if they can always reap larger profits by simply being acquired by market leaders. Thus, nonintervention in startup acquisitions cannot be justified by allusions to the prospect of competitive entry. On the contrary, the more importance one places on entry as a mechanism by which markets self-correct, the more uneasy one should feel about a pattern in which dominant incumbents regularly acquire the most promising startups that come along.

It is instructive to consider Professor Joseph Schumpeter's well-known discussion of "creative destruction," the dynamic process by which new technologies and new rivals persistently upturn the status quo over time. Schumpeter states that

[e]conomists are at long last emerging from the stage in which price competition was all they saw. As soon as quality competition and sales effort are admitted into the sacred precincts of theory, the price variable is ousted from its dominant position. . . . [I]n capitalist reality as distinguished from its textbook picture, it is not [price] competition which counts but the competition from the new commodity, the new technology, the new source of supply, the new type of organization (the largest-scale unit of control for instance)--competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits and the outputs of the existing firms but at their foundations and their very lives.

Schumpeter thus emphasized the prospect of new competition and innovation as playing a key role in fostering economic efficiency over time. In the same spirit as those who believe antitrust should persistently err on the side of nonintervention, he argued that an apparent deficiency in static competition does not imply that the market will perform poorly over the long run. But, as he stressed, this requires that a leading incumbent view the prospect of innovative new firms as an existential threat--not a transactional opportunity with which to extend its lead over [\*350] smaller rivals. However, we have little reason to expect this if such acquisitions are virtually never subjected to meaningful antitrust scrutiny.

As emphasized in the last Section, a serious challenge is that startup acquisitions present significant uncertainties and are therefore less amenable to empirical forecasting than conventional mergers. This means that hypothetical intervention would have to be predicated on less precise economic evidence than courts usually demand, creating some risk of false positives. But that does not mean that such a policy could not improve upon on the status quo.

Importantly, the uncertainties cut in both directions. The current policy, which permits virtually all startup acquisitions by dominant incumbents, is also making errors: some portion of these transactions will inevitably lead to a but-for reduction in future competition, even if this result was not rigorously quantifiable ex ante. As such, it would be naïve to suggest that antitrust is currently avoiding errors simply because almost none of the relevant acquisitions are expressly litigated. If anything, this merely signals that current antitrust standards are too onerous to be administered in practice. This reflects a failure to acknowledge the distinct economic and practical difficulties that distinguish startup acquisitions from more conventional mergers between established incumbents. Current policy simultaneously (a) makes no effort to confront the uncertainties in a practicable way, and (b) gives no weight to the broader incentive problems that may arise if leading incumbents can rely on persistent acquisitions to modulate the future course of competition.

Realistically, any antitrust policy toward startup acquisitions (including one of inaction) is bound to make errors in some percentage of cases. But, as we have argued above, in this context there is no good reason to maintain the traditional view that false positives are more problematic than false negatives. A better approach is to acknowledge that this area involves unavoidable uncertainties, but also significant potential harms, and to develop [\*351] standards that strike a reasonable balance between administrability and the risk of judicial error. Consistent with this, the courts have occasionally cautioned that antitrust standards ought not to demand such a degree of economic precision that they become impracticable. For instance, in Barry Wright Corp v ITT Grinnell Corp, the court noted that "[r]ules that seek to embody every economic complexity and qualification may well, through the vagaries of administration, prove counter-productive, undercutting the very economic ends they seek to serve."

#### The structure of market dictates the ability to innovate.

Mark A. Lemley & Andrew McCreary 21, William H. Neukom Professor of Law, Stanford Law School. Partner, Durie Tangri LLP; J.D./M.B.A., Stanford Law School and Stanford Graduate School of Business, "Exit Strategy," Boston University Law Review, Vol. 101, Issue 1, January 2021, Lexis.

While these may be valid points in particular cases, they neither disprove nor help solve the problems of concentration caused by the norm of selling startups to incumbents.

First, market structure matters. Markets that are not competitive not only distort prices but also reduce innovation. Further, incumbent acquisitions prevent potential competitors from combining to form a company that can credibly threaten entry at scale. So reducing the possibility of Schumpeterian competition is likely to discourage innovation in the long run. And precisely because incumbency does bring some real advantages, we may need to create incentives to support Schumpeterian competition and avoid perpetual incumbency.

[\*68] And second, in any event, the incumbent will put the innovation in the hands of more consumers only if it actually deploys that product. As we have seen, incumbents often buy startups and then kill them, either deliberately or by dissipating the team and not focusing on the acquired product. Incumbents have less incentive to deploy new technologies than startups do. That's because incumbents that replace their existing product with a new one are mostly stealing customers from themselves. And incumbents don't need to innovate to stay alive if they can buy any entrant that looks like a threat.

Finally, the value of scale is similarly not a persuasive reason for most incumbent mergers. There may be markets where network effects are so strong that merger is inevitable. But we should be reluctant to assume that just because scale has value, the incumbent will always make a better product. History is full of cases where that turned out not to be true. Sometimes it just means we need a new dominant firm. And we won't see those leapfrog products if the incumbent buys the potential disruptor. Even in the relatively rare case of technologies that can reach their full potential only when deployed to the entire market, there are alternatives to allowing incumbents to buy up all new technologies. We could, for example, require that certain AI training databases be open to all AIs or that companies allow access by competitors seeking to make their products interoperable with the de facto standard.

#### 2. Regional inequality---tech monopolies create clusters that result in disparate financial outcomes.

Dr. Maryann Feldman et al. 21, Heninger Distinguished Professor in the Department of Public Policy at the University of North Carolina. Adjunct Professor of Finance at Kenan-Flagler Business School. Research Director at UNC Kenan Institute of Private Enterprise; Dr. Frederick Guy, Senior Lecturer, Management, Birkbeck University of London; Simona Iammarino, Professor, Economic Geography, London School of Economics and Political Science; Carolin Ioramashvili, PhD, Economic Geography, London School of Economics and Political Science, "Gathering Round Big Tech: How the Market for Acquisitions Reinforces Regional Inequalities in the US," London School of Economic and Political Science: Geography and Environment Discussion Paper Series, Paper No. 28, May 2021, pg. 1-17.

1. Introduction

Digital start-ups are often seen as a promising vehicle for local economic development, yet as they become established many are pulled away to relocate in major tech clusters. This locational choice is often treated as being explained by agglomeration economies which boost the productivity of tech workers (Duranton and Puga 2004; Nathan 2015; Adler et al. 2019; Gazel and Schwienbacher 2020); others have pointed, additionally, to proximity to venture capital financing (Kenney 2011; Florida and Mellander 2016). It contributes to the growing gulf between prosperous tech clusters and “left behind” places – a gulf seen in incomes, housing prices, age profiles, education levels and voting behavior – which has become central to our understanding of social and political polarization (Rodriguez-Pose 2018; Iammarino, RodríguezPose, and Storper 2019; Rodríguez-Pose and Storper 2020).

We argue that one important factor in many digital start-ups’ choice to locate in major tech hubs – and the Silicon Valley in particular – is that proximity to the largest digital firms makes it more likely that the start-up will be acquired by one of the latter. We study acquisitions by the seven largest American digital platforms – Amazon, Alphabet [Google], Apple, Microsoft, Facebook, Oracle and Adobe (we refer to these collectively as Big Tech). As others have shown (see e.g. Rikap and Lundvall 2020), acquisition of smaller companies is, for Big Tech, a major source of new technologies, ideas and talent. The ability of Big Tech to exploit the monopoly power conferred by positions in control of critical digital platforms has given these companies extraordinary financial resources; continuing acquisition of start-ups and, with them, new features and capabilities, both makes use of these financial resources, and extends and cements Big Tech’s control of its markets.

The other side of the acquisition market – being acquired – is a critical part of the business model of digital start-ups. This is due to the scale economies – low marginal costs and network properties – of many digital products. Success – even survival – in markets for these products requires rapid scaling in order to secure a decisive, often winner-take-all, first mover advantage. Growth therefore requires substantial infusions of equity investment, which typically comes in two distinct stages. In the first, an investor such as a venture capitalist or business angel will take a large, often controlling, interest in the firm. In the second, part or all of the first-round stakes will be sold, either through a stock market flotation (initial public offering, or IPO), or through acquisition by another firm. Either of these routes can represent financial success for a start-up, but acquisition is far more common. Since many start-ups fail to reach the second round of financing, making it happen is an important business objective.

In an earlier paper, we argued that the divide between wealthy and poor places is deepened by monopoly power, and in particular by the power of the large digital platform companies (Feldman, Guy, and Iammarino 2020). Among the reasons we gave for this is that the power of these companies amplifies the centripetal pull of agglomeration economies. It does so for two reasons. One is that the productivity of labor is augmented by monopoly rents, some portion of which is shared in the form of high remuneration for a stratum of skilled workers. The other is that, for smaller digital firms, part of the attraction of major tech hubs is the market for acquisitions: proximity makes it more likely that a particular start-up will be chosen as the latest augmentation of a major digital platform, with the result that some sliver of the platform’s rents will be shared among the owners and key employees of the startup. In this paper, we develop the second of these claims, about the market for acquisitions.

Why does it matter if the productivity-enhancing characteristics of a tech cluster – the familiar Marshallian properties – are being amplified by monopoly power? Simply, because the welfare implications of the two sources of agglomeration are different. If an agglomeration grows large and wealthy purely for Marshallian productivity reasons, then any efforts to redistribute benefits from that agglomeration to places left behind, must face the question of whether a spatially-targeted economic development policy may kill the goose that is laying golden eggs. But if the agglomeration has grown still larger and wealthier due to the monopoly power of firms based there, then familiar remedies such as anti-trust enforcement and regulation may mitigate both the problems typically associated with monopoly, and regional disparities in income and opportunity.

In this paper, we develop evidence for the market for the acquisition of start-ups. Section 2 places our main arguments in the literature. Section 3 provides data sources and descriptive evidence on the acquisitions. Section 4 describes our reference groups of companies in the same sectors. Section 5 presents regression analysis that supports the finding from the US descriptive statistics. Section 6 discusses implications and concludes with future directions for research.

2. Start-ups and the geography of Big Tech acquisitions

Theories of cluster dynamics have often addressed the relationships between large and small firms (e.g. Feldman, Francis, and Bercovitz 2005; Feldman and Lowe 2015). It has been common to see large firms as anchors to a cluster, with smaller firms present to serve them, typically as suppliers, and with large firms connecting the cluster to distant markets and sources of knowledge. Such have variously been labelled hub-and-spoke cluster, solar cluster, or industrial complex serving the needs of larger client firms (Markusen 1996; Piore and Sabel 1984; Gordon and McCann 2000). Small firms may also be clustered close to a large one because the small ones were founded by former employees (Klepper 2011; 2015).

We are interested here in a different spatial dynamic between large and small firms: not a stable set of small suppliers to the large firms, nor small firms founded on exit from large ones, but a marketplace for acquisition of the small firms by the large firms.

Start-ups may be drawn to tech clusters by productivity-enhancing agglomeration economies; by better access to financing for growth (we include being acquired as financing); or a combination of the two. Against these are factors encouraging digital start-ups to locate elsewhere. As Dahl and Sorenson (2012) show, entrepreneurs, even more than workers, prefer to grow their business in a place they already know – in other words, they would rather stay home (see also Florida and Hathaway 2018; Sorenson 2018). The product of the digital start-up seems well suited to this: it is weightless, something that in technical terms can be shipped anywhere in the world as easily as across the street; the infrastructure requirements for operating such a firm are easily satisfied in a developed country – a reliable broadband connection and electric power; state and local governments offer incentives to stay; and, while most places do not have the large pool of talent available in a major tech hub, skills of programming and software engineering are taught universally. Notably, programming skills are taught and researched at a high level in places such as Urbana-Champaign, Illinois, and Ann Arbor, Michigan, and dozens of other rustbelt or southern universities with fine engineering schools. Some of these places figure prominently in the world ranking of start-up hubs (Florida and Hathaway, 2018); yet, as we see below, these places are negligible as home bases for start-ups at the time of acquisition by platform giants.

Digital start-ups1 are likewise producing highly scalable products – sometimes tools or features to augment existing platforms, sometimes platforms on their own (e.g. Ruggieri et al. 2018). The commercial logic of proprietary digital products is that they must be scaled, because a successful first mover will dominate a market segment (Schilling 2002). This winner-take-all market structure means that scaling must occur quickly, depriving digital firms of the option of organic growth from retained earnings: external financing is needed. Moreover, starting up in a winner-take-all market is very risky, so the external financing must be equity financing, from investors with pockets deep enough to sustain some years of losses at best and to lose all at worst. The same technological features which make unregulated proprietary platforms into monopolies thus drive small digital companies to prioritize relations with equity investors. In what we can call the venture capital stage, this investment typically comes in the form of private equity, often from a firm specializing in venture capital; sometimes, it will come from an individual “angel” investor (Kenney 2011; see also Table 4, below). Such equity investors habitually acquire controlling stakes in the company, and usually intend to sell the company at a later stage. Sale may either entail the start-up being floated on the stock market in an initial public offering, or its acquisition by another firm. The IPO route comes with the prospect of wealth and fame for the founders as principals in an independent company; it is, however, by far the less common of the two routes. An acquisition by Big Tech is among the most lucrative outcomes (Dwoskin 2020).

Guzman (2019) considers two competing explanations of US start-ups’ location – agglomeration economies and social embeddedness and concludes that agglomeration economies are far more important than social networks. He finds selection of high-quality start-ups, particularly from lower agglomeration regions, into Silicon Valley, and shows that moving results – among other financial and market benefits – in a higher likelihood of being either acquired and offered an IPO or, in other words, of achieving “extreme” growth (see also Andrews et al. 2020). Guzman is silent on any role the market power of the acquiring firms might have in raising the likelihood of takeover, simply treating the greater likelihood of financing for extreme growth as one of the benefits of agglomeration. We see the same in Kerr and Robert-Nicoud’s (2020) otherwise exhaustive review of the character of tech clusters, and of the frantic attempts of places to brand themselves as “Silicon Something”: consideration of market power is absent.

There is now considerable public scrutiny2 of acquisitions by tech giants. There is also a growing body of research. This literature does not ignore market power, but it does ignore geography, location in space. “Space” needs the qualifier because the literature is full of references to “space”, “cluster” or “zone”, but these refer to “spaces” of products or technologies

The secular growth of market power in the US since 1980 is now widely recognized (De Loecker and Eeckhout 2017; Eggertsson, Robbins, and Wold 2018; Hsieh and Rossi-Hansberg 2019), and has been particularly great in digital sectors (Calligaris, Criscuolo, and Marcolin 2018). It is an era which dawned with the de-regulation of old network industries, and singularly failed to come to terms with the opportunities for market power presented by proprietary digital platforms. The market power of giant technology firms has become, belatedly, a matter of concern for, among others, the US Congress (anti-trust hearings), the government of Australia (clamping down on Google and Facebook’s free-riding on newspapers), the DG Competition of the European Commission (new proposal for the Digital Market Act) and legislators in states like North Dakota (taking aim at Google and Apple’s monopsonies on apps using their respective phone platforms). Meanwhile, a handful of Silicon Valley and Seattle companies operating digital platforms have far and away the highest market valuations of any corporations on Earth.

Platforms connect users as a network. Networks benefit from increasing returns, which creates economies of scale and results in lock-in as a source of their advantage (Kenney and Zysman 2016; Rikap 2020). Our Big Tech companies control particularly large platforms, and often control more than one network (Ducci 2020; Stallkamp and Schotter 2021). For example, Alphabet and Apple both control phone operating systems, which are platforms for apps – most phone apps in the world go through the online “stores” of these two companies; Adobe, Microsoft and Oracle all sell general purpose software which produces user files in proprietary formats, making users dependent on the platform in order to ensure full inter-operability; Alphabet and Facebook dominate online advertising in much of the world; Amazon’s “marketplace” connects hundreds of thousands of vendors with hundreds of millions of customers; Facebook and Microsoft operate social media platforms; and so on.

Digital platform business models serve networks, but they are unlike the relatively static networks of the 19th and 20th centuries. Owners of an electric power grid, once it is in place, have little to fear from competitors; for a digital platform, maintaining market position, and monetizing that position, demand ongoing innovation, refinement of features, addition of services (Hindman 2018). The major platforms have considerable internal research capabilities but, compared with the industrial giants of the previous century – General Electric, IBM, and such – they source more innovation through acquisition (Lazonick 2009; Rikap and Lundvall 2020). And, because of the financial resources which their market power confers, the major platforms are always in a position to acquire.

Gautier and Lamesch (2020) review five tech giants’ (Google, Amazon, Facebook, Amazon and Microsoft, or GAFAM) acquisitions during 2015-17. They find that acquisitions mostly fall in the firms’ core markets segments or product spaces. Similarly, Argentesi, Buccirossi, Calvano, Duso, Marrazzo and Nava (2019), in a study of acquisitions by Amazon, Facebook and Google (AFG) in the decade 2008-2018, conclude that acquired products and services are largely complementary to those already supplied by the three companies. This is supported by Lopez Giron and Vialle (2017) in their study of Microsoft’s acquisitions in the period 1992-2016, focussing on acquired resources and competences: the largest share of acquisitions complements (rather than diversifies) Microsoft’s core businesses. Gautier and Lamesch (2020) find that most of the acquired products are discontinued post-acquisition, implying that the acquisitions are largely motivated to gain intangible assets such as intellectual property rights and talent.

This pattern of acquisitions has implications for innovation. Bryan and Hovenkamp (2020) find that start-ups which aim to be acquired are biased toward inventions that improve the leader’s technology, rather than offering an alternative to it.

Tech giants may acquire companies to suppress competition. This is well documented in other information-based sectors, such as pharma and microprocessors (Feldman et al. 2020). Cunningham, Ma and Ederer (2021), for instance estimate (conservatively, they say) that between 5.3% and 7.4% of US acquisitions between 1989 and 2010 were “acquire to kill”, and thus harmful for both innovation and competition.

Moreover, even without any deliberate suppression of a competing product, a tech giant’s acquisition or development of a product tends to create a “kill zone”, in which competing projects struggle to get both users and capital (Kamepalli, Rajan, and Zingales 2020). Rival startups offering substitutes for an acquired product find themselves starved of both capital and customers; notice that this describes a winner-take-all market, in which the winner is the first to be acquired by a tech giant. Wen and Zhu (2019) find that when Google appears likely to develop a new app or capability for Android internally, its smaller competitors reduce innovation and raise prices.

Argentesi et al. (2019), Bryan and Hovenkamp (2020), and Kamepalli, Rajan and Zingales (2020), Motta and Peitz (2020), Katz (2020) and Cabral (2020) all draw conclusions about implications for competition and innovation policies. Abstracting from the considerable differences in methods and disciplines there are common themes. One is the deleterious effect platform acquisitions can have on innovation. Another is the fact that the vast majority of acquisitions by platform-based tech giants evades investigation, often because the turnover of the acquired digital start-ups falls below the threshold required to trigger government intervention. The consensus is that legal restraints on merger activity are not doing the job, and that tightening these restrictions could improve both innovation and competition. Kamepalli, Rajan and Zingales (2020) stress, however, that more effective measures would be ones which directly attack the exclusivity of platforms: open standards; controls on data ownership; and – for countries outside the US – restrictions on global, US-based platforms to make room for national alternatives (policies in China and India are offered as examples here).

Acquisition is usually a financially desirable outcome for the shareholders of a small firm; for firms in the winner-take-all digital platform economy, it may be the only route to survival. The home bases of the platform giants are geographically concentrated: the seven Big Tech we study are located in just two metropolitan areas, San Francisco-San Jose (the Silicon Valley), and Seattle. If physical proximity to the acquirer makes a successful acquisition more likely, this provides a motivation for start-ups to move to a major tech hub, and to the Silicon Valley in particular. This centripetal pull of the tech clusters for precisely those start-ups which have the greatest growth potential, can impose a low glass ceiling on local economic development initiatives based on developing digital technologies: the product may be weightless, but the market for the company is elsewhere.

3. Data and descriptive analysis

Our empirical investigation relies on all acquisitions made in the US by the seven largest digital platforms in terms of market capitalization3 : Alphabet (Google), Adobe, Apple, Facebook, and Oracle headquartered in Silicon Valley, Amazon and Microsoft located in Seattle, from their inception to the current time.4 In choosing these firms we are excluding a number of neighboring categories: non-US platforms (China has at least two that would qualify on the basis of market valuation, Alibaba and Tencent Holdings); payment platforms (had we included American payment platforms with market valuations in the neighborhood of the companies that we did include, it would have changed the geographical picture little, with two in the Silicon Valley (Visa and PayPal), and one in New York (Mastercard)); hardware (the standards of Intel or Nvidia can be regarded as platforms); telecommunications (mobile phone networks); and entertainment (Netflix, for instance, is a digital platform).

Each of our seven firms began as an entrepreneurial start-up. Most received venture capital financing (Oracle leveraged federal procurement contracts), grew rapidly, and went public. The phenomenal growth of these firms, together with that of hundreds of smaller digital platform companies, contributed to the belief that “tech”, specifically, digital technology, offers an attractive building block for local economic development.

Table 1 provides an accounting of acquisitions, by acquiring company, from their earliest acquisition through 2020. We include all full (100%) acquisitions made by Big Tech and their subsidiaries. Data are drawn from three databases, Zephyr-Bureau van Dijk, Capital IQ-S&P and SDC Platinum-Refinitiv, and then cross-checked and verified through manual searches. Acquisitions which cannot be verified through either mention on the company website or news articles are not included in our sample on the assumption that these are likely “acqui-hires” with no substantial start-up company involved. The Appendix details the laborious procedure used to construct this list, and shows statistics for the various data sources.

In total, the seven Big Tech acquired 940 firms worldwide, with 674 acquisitions based in the US. Notably, all of the tech giant firms started acquiring of other firms in the years following their own IPO. Alphabet completed the largest number of acquisitions to date (237), closely followed by Microsoft (235). Acquisitions occur consistently over time, with the annual average number of acquisitions ranging from 12 a year for Alphabet, to two for Adobe.

[Table omitted]

For all of the Big Tech, the majority of acquisitions were sourced from the United States; 266 (28%) were acquisitions of firms based outside the US (Table 2). The largest number of acquisitions are from the UK (50), Canada (42) and Israel (32). Acquisitions within countries are also geographically concentrated. This pattern broadly reflects the distribution of major high tech clusters, and their linguistic and political affinities with the United States, identified in Arora and Gambardella (2005). However, out of all of Asia and the developing world the only country with more than two of the takeover targets was India (6) while China, Japan and Korea – all major locations for digital technology – are under-represented.

[Table omitted]

Caution is required in comparing countries. “Acquiring a company” does not mean the same thing everywhere; in some countries it comes with much greater obligations to employees and other stakeholders, than it does in others (Hall and Soskice 2001). The tax implications may also differ. The Big Tech acquisitions are largely about acquiring skilled employees and intangible assets, rather than operations and physical assets. Hiring the employees, or purchasing the intangible assets, can be alternatives to acquiring the firm as a whole, and differences in national institutions may weigh in the choice.

This caution does not, however, hinder within-country comparisons. In most countries for which the numbers are large enough to generalize, we see an overwhelming concentration in the country’s financial capital – London, Tel Aviv, Paris, Stockholm, Dublin; however, in Canada and Germany – both federal, polycentric states – the pattern is more geographically dispersed.

We limit our further analysis to the acquisition targets with known locations in the United States for the two decades 2001-2020.5 Table 3 gives a breakdown of the Big Tech acquisitions by metropolitan statistical areas (MSAs) and the percentage that received venture capital investment. Investor (vendor, to Zephyr) data are from Zephyr and includes the majority owners at the time of acquisition, which are similar to the ownership information that would be provided in an IPO prospectus: reliable data are only available after 2001. While VC is the largest source of financing, businesses’ founders, angel investors and other entities such as banks, and wealth and investment management firms were mentioned. There was only one public-private equity investment from the New York City Investment Fund LLC, now known as Partnership Fund for NYC. Individual investors include founders and angel investors. Two universities were listed: Stanford (in the Silicon Valley) and the University of Washington (in Seattle).

Of the 603 US acquisitions about half of them received venture capital investment. For non-US acquisitions only one quarter received venture capital investment. Note the extreme concentration in the Silicon Valley, which we define as the combined San Jose-Sunnyvale-Santa Clara and San Francisco-Oakland-Berkeley MSAs. The largest number of firms (291 or almost half of the US Big Tech acquisitions) were located in Silicon Valley at the time of acquisition, with 54% of the companies receiving VC investment. Four other MSAs form a distinct second tier for acquisitions: New York, the financial capital of the US; Boston and Los Angeles, both of which are important centers of both technology and private equity finance; and Seattle, the home of Microsoft and Amazon. In all, there were 18 cities with two to five acquisitions. There were 25 MSA’s with one acquisition each and one third of these companies received VC investment. Overall, 49% of the acquired companies received VC investment: the percentage is slightly higher, among the main hubs, in Silicon Valley, New York and Boston.

[Table omitted]

We compare the spatial distribution of Big Tech acquisition targets to the distribution of four different sets of firms seeking finance in relevant industries.

Three-quarters of the seven Big Tech acquisitions are attribute to three SIC codes: 7371 – computer programming services, 7372 – prepacked software, and 7374 – computer processing and data preparation and processing services. The remainder of the acquisitions were dispersed across many sectors. Our four comparison groups are limited to these three SICs, using the sources and definitions described in Table 4.

[Table omitted]

The broadest of the four comparison samples consists of 6,213 firms in the three SIC codes that received Small Business Administration (SBA) 7(a) loan guarantees. This is a government guaranteed loan that is made to firms who are seeking investment funding and have demonstrated their credit worthiness. A narrower sample consists of the 3,005 firms listed in the Zephyr database as having been fully acquired, but with a purchaser other than one of our seven Big Tech (All Other Acquisitions). The narrowest sample consists of the 196 firms which had Nasdaq IPOs, Nasdaq being the leading exchange for tech company stocks (IPOs). Finally, we also compare with a subset of the All Other Acquisitions sample: 1,030 firms sold by VC firms which also sold firms to Big Tech. Just as the map of international Big Tech acquisitions is not the global map of digital technology, neither is the US acquisitions map the same as the map of the sectors involved.

[Table omitted]

Table 5 examines the geographic distribution of the comparison data sets. Silicon Valley has the highest count of firms for all the categories except SBA loans, which has a larger geographic reach. Firms that receive SBA loans have the human capital and organizational capabilities required to establish start-ups in these industries and are widely geographically distributed. Indeed, the counts of firms applying for SBA loans encompass a much larger set of cities and suggests that public financing may be an alternative substitute when VC funding is not available. New York City, the largest metropolitan area is more heavily represented by SBA loans.

[Table omitted]

The maps in Figure 1 show the distribution of the first three comparison groups, and their number relative to Big Tech acquisition targets. The size of circles indicates the number of firms in the comparison group, while the shading indicates the number of Big Tech acquisitions relative to that. Circles in the darkest shade of red indicate that there were a larger number of Big Tech acquisitions in that place than companies in the comparison group. Empty (white) circles indicate an MSA with cases from the relevant comparison group but no Big Tech targets.

[Figure omitted]

The top panel shows the distribution of firms from the selected tech industries that received SBA 7(a) loans. Most large cities had a substantial number of SBA loan recipients but small places also have firms that received these loans. This does not mean that the latter are evenly geographically distributed: cities known to be tech hubs had a larger share, as we would expect. In most MSAs, the number of Big Tech acquisitions is less than 25% of the number of firms receiving SBA loans. The number exceeds this level only in a few small cities that are secondary tech hubs (the most substantial being Boulder, Colorado), and the Silicon Valley, where the number of Big Tech acquisitions exceeds the number of SBA loans.

The comparison group for the second map is all other acquisitions of firms in relevant industries. An acquisition is a relatively rare event, so there are overall fewer circles than in the SBA case. Compared with SBA loans, acquisitions are more concentrated on the West Coast and in the Northeast; still, they remain far more widely distributed than Big Tech acquisitions. Of large cities, the ones with the higher ratios of Big Tech acquisitions to other acquisitions are all on the West Coast: Seattle, Silicon Valley, Los Angeles, San Diego.

The third map compares Nasdaq IPOs. In most locations that have IPOs, there are only a few, and the number of Big Tech acquisitions is at a comparable, or even higher level.

Finally, we compare the growth in geographical concentration of Big Tech Acquisition to that observed in the four comparison samples, using the share of targets located in Silicon Valley and the Herfindahl-Hirschman index by MSA (Figure 2). The extent and increase in concentration of tech giants’ targets is striking. In the first period, other businesses sold by investors that sold to Big Tech were actually more concentrated on both measures. However, this changed by 2011-2020, where almost 30% of Big Tech US-based targets were located in the Silicon Valley, compared to less than 10% of all acquisition targets in the relevant industries.

[Figure omitted]

4. Differential Outcomes: Probit Analysis

We estimate probit models to provide additional descriptive analysis. Table 6 presents definitions for the variables we use in our regression along with summary statistics.

[Table omitted]

How are the odds of being acquired by a Big Tech affected by the location of the acquired firm? If being acquired is an objective for a start-up and if location affects the likelihood of being acquired, this would be a factor in drawing such firms to certain locations, above and beyond any productivity advantages. Moreover, the digital tech sector is not homogeneous; an acquisition is the outcome of a matching process. Some start-ups have products, intangible assets or personnel that offer better potential matches than others for our tech giant firms, and for that reason would be more likely to locate in a place which makes acquisition by one of the Big Tech firms more likely.

The likelihood of being acquired by Big Tech may also be affected by the source of external financing in the first stage, however. Proximity to venture capital, in particular, is often claimed as an advantage to locating in the Silicon Valley (Saxenian 1994; Kenney and Florida 2000; but see Lerner 2009). A variant of this claim stresses the role of certain specialist investors – most but not all of them VCs, and many but not all of them located in the Silicon Valley – which have strong ties to the tech giant companies dating back to the first-round financing of the latter: initial backing from a member of this small group may bring a start-up into the right networks, and improve its chance of being acquired by one of the Big Tech. These early investors in Big Tech are listed in the Appendix (Table A.3); they include such well-known firms as Venrock, Sequoia Capital, and Greylock Partners.

Finally, it is possible, from what we have seen in our description of the data, that Big Tech targets are disproportionately – that is, compared with other acquisitions in the relevant industries – not just from the Silicon Valley, but from major digital tech clusters (e.g. Boston, Seattle, New York, Los Angeles) overall. To check this, and also to control for this agglomeration effect in our estimate of the Silicon Valley effect, we include a variable for the number of SBA loans in the three focal industries and in the MSA in which the target is located.

Any of these same factors – Silicon Valley location, VC funding, the location of the VC, and a backer in the group of initial Big Tech funders, size of digital tech agglomeration – could also affect the likelihood that second stage financing will take the form of IPO rather than acquisition. In the IPO case we do not have strong priors on what these effects would be, but are interested in what the comparison will tell us.

We estimate two sets of probit models. In the first set (Table 7), the data consists of all Big Tech US acquisition targets in the years 2001 to 2020, together with all other acquisition targets in our three focal industries. In the second set (Table 8), the data consists of firms in the focal industries that went to IPO, versus all acquisition targets other than the Big Tech targets. All models include Target in Silicon Valley (binary), SBA Loans (continuous), and a dummy for the year. Other variables, depending on the model, are Investor Venture Capital, Investor in Silicon Valley, and Investor Early Investor in Big Tech (all binary).

We see in Table 7 that a Silicon Valley location for the target has a positive and statistically significant effect; with all covariates included, the marginal effect of a Silicon Valley location is 0.17, which is to say a 17 percentage point increase in the probability of being acquired by a Big Tech. The investor being a venture capitalist actually makes it more likely that the firm will be acquired by a non-Big Tech firm; in contrast to the target’s location, the investor’s location in the Silicon Valley has no discernable effect on whether the acquirer is a Big Tech. On the other hand, the investor being an early backer of one of the Big Tech companies has a positive and statistically significant effect on the likelihood of a Big Tech matchup; with a marginal effect of 0.16, it is essentially the same as the effect of a Silicon Valley location for the target.

SBA loan numbers have a positive and statistically significant effect, indicating that in MSAs with fewer digital tech SMEs the ratio of Big Tech acquisitions to acquisitions by other firms is lower.

[Table omitted]

Table 8 considers the same factors for the IPO outcome. Again, we see a strong positive effect from Silicon Valley location – at 0.19, about the same as for Big Tech acquisitions. If an IPO is the big prize for founders of a start-up – the attraction of growing while staying independent – this suggests one more attraction of the Silicon Valley. Having a venture capitalist for an investor has (again) a negative effect; having an investor located in the Silicon Valley is now significantly negative; having an investor who was one of the early Big Tech backers has a statistically insignificant effect after controlling for other variables. As with Big Tech acquisitions, more SBA loans – which is to say, greater size of the digital tech agglomeration in the MSA where the target is located – raises the likelihood that second round financing will come in the form of an IPO rather than acquisition.

[Table omitted]

5. Conclusion

The conditions which link digital platform monopoly and the pull of start-ups to the major tech hubs, may be summarized as follows. Digital products are scalable; some can be scaled as platforms, which connect users as a network, creating lock in. For the platform giants this has been the basis for monopoly power. The monopoly power of the Big Tech (as for other giants, in other IPR-, network-based industries) is never secure. They are Schumpeterian (Schumpeter 1942), innovating to maintain and extend their market power. However, unlike the manufacturing giants of the twentieth century, much of the platform giants’ innovation is essentially outsourced to start-ups, which the Big Tech may then choose to acquire. The latter compete in a sort of tournament in which being acquired is the prize. The tech giants have vast financial resources, putting them in a position to make an offer for any smaller firm they might find useful.

Many who start new firms might prefer independence, rather than being acquired. Digital start-ups are, however, producing something which is scalable, with very low marginal costs. This puts them in a winner-take-all market, where the first mover into a particular platform function or a new technical standard can have an overwhelming advantage. Start-ups thus require infusions of equity, which we can think of as coming in two phases. It is common, in the first stage, for this to come from investors who expect to sell the company on if its product proves successful. The second stage is either and IPO or acquisition by a larger company. We regard either of these second stage outcomes as infusions of capital for the start-up, though in the acquisition case the start-up may lose its identity altogether. From the standpoint of shareholders in the start-up, completion of the second phase represents success.

For the start-up, there is no certainty in this path to being acquired: even if the start-up’s product (“product” here might be a new platform, but it can also simply be some IPR, or a team’s demonstrated ability to solve a particular kind of problem) is a good one, another start-up may have something similar, or the large firm may develop something internally. Start-ups will therefore be motivated to position themselves ways that improve the likelihood of being acquired.

Is moving to Silicon Valley one of those ways? Although our seven Big Tech do acquire firms throughout the US and in many other countries, some places see far more than their share of acquisitions. Certain foreign countries (Canada, UK, Israel), certain foreign cities (London, Paris, Bangalore, Tel Aviv), certain cities in the US (New York, Boston, Seattle, Los Angeles) and, far above even those, the Silicon Valley itself. Within the US, we are able to compare this with the distribution of all acquisitions, and with the distribution of SBA loans, in the relevant industries. By both measures, Big Tech acquisitions lean far more heavily to the major tech clusters and, again, far far more to the Silicon Valley. Big Tech may be able to source its weightless acquisitions globally, but it tends to make most of its purchases in a few very familiar shopping malls.

To what extent are these acquisitions driven by the market power of Big Tech? Market power gives Big Tech the means, in the form of piles of cash. It also gives Big Tech the motive: just as Schumpeter described, the monopoly is maintained, extended and renewed through innovation – albeit, now, innovation is to a large extent initiated outside the monopoly firms themselves.

Actions on both sides of acquisition transaction – the need for the start-up to scale up, and the means and motive of Big Tech – grow out of the proprietary control of access points to digital networks.

The problem of network monopoly has been faced before. In the late 19th and early 20th centuries, for instance, then-new network industries such as electric power, telephones and railways developed huge power, and were subsequently brought either under public ownership, or public regulation, almost everywhere in the world. The various modes of regulation are beyond the scope of this paper – suffice it to say that in technical terms this is not an unknown problem. In political terms it is perhaps a bigger problem than that faced with the old network industries, because the geography of digital platforms is different. An electric power network or a railway has assets and employees distributed around the limited territory it serves; those who are harmed by the monopoly are in roughly the same place as the monopoly’s assets and employees, which makes the regulation of the monopoly a distributional matter within a welldefined polity. Twentieth century American regulation of public utilities and banks actually enforced this by keeping the companies within state lines.

The geography of a digital platform firm is much different, and that different geography makes for a different politics of regulation. The platform firm typically has assets and employment concentrated in a few locations. For those locations, it is an important export industry – that is precisely why state and local governments seek to foster tech clusters. Big platform firms exercise market power nationally in the United States, and internationally. Within the US, the economic interests of the major tech clusters are in conflict with those of the places left behind; internationally, the maintenance of Big Tech’s monopolies has become a central pillar of US trade policy (Guy 2007; Rodrik 2018).

Should this situation change – following, perhaps, the sorts of measures outlined by Kamepalli et al. (2020) – the consequent decline in the acquisition market should make it more feasible to foster the growth of digital start-ups in what are now left-behind places. It would also remove one factor which drives the seemingly endless growth in size and housing costs in the major technology clusters.

What we have observed here, in the case of seven large digital platform companies, raises a bigger question about the geography of market power and of acquisitions. What goes for digital platforms may, or may not, go for other types of information-based product with extreme increasing returns and wide geographical reach, such as pharma, biotech, and digital media. Moreover, with digital platforms and with others, how much of the acquisition market is held by giant firms, as opposed to merely large? In the first instance, both questions could be addressed through a mapping of takeover relationships – locations, distances – in relevant industries.

#### That decks growth.

Valerie Cerra et al. 21, Assistant Director and Division Chief, European and Middle Eastern Division, IMF’s Institute for Capacity Development. PhD, Economics, University of Washington; Ruy Lama, Senior Economist, IMF. PhD, Economics, University of California at Los Angeles; Norman V. Loayza, Director, Global Indicators Group, Development Economics Vice-Presidency, World Bank. PhD, Economics, Harvard University, "Links Between Growth, Inequality, and Poverty: A Survey," Development Economics Development Research Group, Working Paper 9603, March 2021, pg. 23-26.

The empirical evidence shows that poverty is detrimental to long-term economic growth. Using panel data of 85 countries covering 1960 to 2000, López and Servén (2015) find that a 10 percentage-point increase in the poverty rate reduces the GDP per capita growth rate by 1 percentage point. In particular, an increase in the poverty rate reduces the investment rate for countries with low levels of financial development. There is also evidence that the negative impact of poverty on growth depends on the initial level of poverty. In a sample of 156 countries covering 1960 to 2010, Marrero and Servén (2018) find that for low levels of poverty (below the median), poverty has an insignificant impact on growth (Figure 8). In contrast, when the poverty rate is high, a 10 percentage-point decrease in headcount poverty is associated with an increase in economic growth ranging from 1 to 2 percent per year.

Related evidence comes from the observation that despite the global reduction in poverty rates, cross-country evidence indicates a lack of convergence in poverty rates. Studying 90 developing countries during the 1991–2004 period, Ravallion (2012) finds that two distinctive effects prevented the convergence of poverty rates. First, poverty reduces growth, consistent with the results from López and Servén (2015). Second, high initial poverty dulls the impact of growth in reducing poverty. The combination of these two channels makes it more difficult for the poorest countries to reduce their poverty rates.

Figure 8. Growth in GDP per capita vs Initial Poverty, 1960–2010

A higher poverty rate is associated with lower growth in subsequent decades.

Chart, scatter chart

Description automatically generated

4.1.2. From Inequality to Growth

As an illustration of the relationship from inequality to growth, Bénabou (1996) compares the growth outcomes of East Asian and Latin American economies conditional on the initial levels of income inequality. According to Bénabou (1996), the conventional wisdom among development economists is that the relatively equal distribution of income and land in East Asian economies contributed to their observed high economic growth rates. By the same token, the lack of a similar economic dynamism in Latin America has been attributed to the consequences of high concentration of wealth and income in that region.13

The panel a of Figure 9 reports the correlation between income inequality in 1980 and the average GDP per capita growth in the subsequent 30 years for selected Latin American and Asian economies. Consistent with Bénabou (1996), on average countries that exhibited lower levels of initial inequality also experienced higher rates of economic growth. While there are many other factors that might explain the economic dynamism of these Asian economies, such as the quality of institutions and high rates of saving and investment (Collins and Bosworth 1996), this figure illustrates that income distribution might be one key element for understanding differences in economic performance. An extended sample of advanced and developing countries (Figure 9, panel b) confirms the relationship between initial income inequality and subsequent growth.14

Figure 9. Growth in GDP per capita vs Initial Inequality

Countries that started with lower levels of inequality experienced higher rates of economic growth.

Chart, scatter chart

Description automatically generated

The empirical relationship between inequality and growth has been investigated formally in a number of cross-country growth studies, following Barro and Sala-i-Martin (1995). Many of these studies find that inequality, typically measured by a Gini coefficient, enters with a negative and statistically significant sign in cross-country growth regressions, indicating that an increase in inequality leads to lower economic growth. In a survey of 23 different empirical studies on inequality and growth, for instance, Bénabou (1996) finds that despite differences in data sets, sample periods, and measures of income distribution, the studies consistently find that initial inequality is negatively associated with growth. In particular, the quantitative effects of inequality are quite robust across studies: a one-standard-deviation decrease inequality raises the annual growth of GDP in the range of 0.5 percentage points to 0.8 percentage points.

Various studies examine different dimensions of the relationship. An early work by Alesina and Rodrik (1994) finds that income and land inequality are statistically significant variables that decrease long-term growth in a sample of 70 advanced and developing countries. Perotti (1996) finds a negative and robust association between inequality, inversely related to the share of the middle class (third and fourth quantiles of the income distribution), and growth. He finds that social political instability and fertility rates could be driving the relationship between inequality and growth.

The impact of inequality on growth can also depend on the initial level of development. Barro (2000) estimates the impact of inequality on growth by splitting a sample of 100 countries into high- and low-income samples. In that specification, there is a negative relationship between inequality and growth for poor countries, similar to previous studies, while the relationship is positive for richer countries. The empirical results suggest that in the presence of credit constraints, inequality prevents low-income households from accumulating human and physical capital, resulting in lower growth in poor countries. On the other hand, the positive relationship observed in richer economies is consistent with the traditional growth-enhancing effects of inequality emphasized by Kaldor (1957).

The effects of inequality on output might also differ across economic sectors. For instance, Erman and te Kaat (2019) identify the effect of inequality on industry-level value added growth. The authors use a data set that includes 22 industries in 86 countries for the period between 1980 and 2012. They find that that higher income inequality increases the growth rates of industries that use physical capital intensively, while it decreases the growth rates of industries that use skilled labor intensively. Thus, the lower human capital stock associated with inequality drives its negative effect on growth. At the country level, these results are consistent with the theoretical predictions by Galor and Moav (2004).

#### AND regional inequality foments domestic terrorism.

Peter Lawrence 21, Emeritus Professor, Development Economics, Keele University, "The Global and National Inequality Faultlines: The Economic Dimensions of (In)security," Journal of Global Faultlines, Vol. 8, No. 1, pg. 28-29, June 2021, JSTOR.

A concern for ruling classes throughout the world must be the potentially destabilising effects of inequality, resulting in more than simply their governing parties losing elections. Civil wars, revolutions and terrorist incidents are possible outcomes of inequality. There are many studies that have looked for a relationship between inequality and social and political instability as evidenced, for example, by revolts and terror attacks. Survey data has been used to show that, while controlling for other possible causal variables, inequality has a positive relationship with a preference for revolution (Macculloch, 2005), although what people say and what they would actually do when faced with a potential revolutionary situation is another matter. Indeed, when the evidence is collected for revolutions that have occurred, the relation with inequality is not always in the same direction and, as Macculloch observes, some analyses have found a U-shaped relationship where revolts take place in high and low inequality countries.

Studies into the relationship between inequality and terrorism have shown inequality is positively related to the number of terrorists acts. One study found that controlling for various factors that could cause such attacks, such as types of regime (no relationship to terrorism), civil war (a positive relationship), population size (positive) and levels of economic development (no relationship), an increase in the Gini coefficient by one unit results in a 7.4% increase in the number of domestic terrorist attacks (Krieger and Meierrieks, 2018). The same study tested for endogeneity, that is, that terrorism could lead to income inequality as more state resources were spent on suppressing terrorism and, thus, less available for redistributive measures, and they found that this did not affect the robustness of their findings. Their results suggested mechanisms through which inequality is transmitted to terrorism, the most important being the quality of institutions. For example, they find that inequality results in poor institutional quality and, therefore, for example, more corruption and a poorer level of human rights. They also include a horizontal equality control variable, ethnic discrimination, and find that it is associated with increased domestic terrorism. They suggest that, added to these indirect ways in which inequality affects terrorism, there may be direct ways, as the relative deprivation of the lower income groups results in increasing discontent spilling over into terrorist attacks.

Not surprisingly, they find that using the ratio of the Gini coefficients of gross to net income as the redistribution variable, more equal distribution has the opposite effects, such that a one-unit increase in the Gini ratio results in an 85.3% reduction in terrorist acts. They also find that greater equality is associated with better institutions but negatively associated with investment, as redistribution from higher earners, who save more, to lower earners, who save less, will result in less investment and lower rates of growth. Redistributive policies resulting in higher public expenditure may, according to this argument, crowd out private investment leading to lower growth outcomes. This is a classic neoliberal conclusion and contradicts other findings associated with structuralist economic approaches cited above that suggest redistribution is good for growth as it increases consumption demand and, so, stimulates investment.

Distributions of income and wealth across countries’ populations may be very different across regions, such that interregional inequality may be the key variable determining the amount of terrorism a country suffers. One study found that, after controlling for other possible explanatory variables such as GDP per head, population size, interpersonal inequality, economic decentralisation and openness to international trade, higher interregional inequality increased the incidence of domestic terrorism (Ezcurra and Palacios, 2016).

#### HVEs (homegrown violent extremists) decimate critical infrastructure.

Maggie O’Connell 20, Regulatory Affairs Specialist, American Fuel & Petrochemical Manufacturers Association, "Critical Infrastructure & the Emerging Market for Domestic Terrorism," The Infragard Journal, Vol. 3, Issue 1, Winter 2020, pg. 1-6.

CRITICAL INFRASTRUCTURE ASSET OWNERS and operators are no strangers to the threat landscape in which their facilities reside. 2 Companies invest millions of dollars annually to secure their fencelines, patch their cybersecurity vulnerabilities, and develop protocols and procedures to comply with regulations seeking to mitigate and prevent terrorist activity. The safety and security of facility operations, personnel, and the surrounding communities are the goals of these investments, and critical infrastructure companies cannot afford to shortchange in any of these areas. Nevertheless, the threat landscape often evolves so quickly that the legislative and regulatory frameworks fall vastly behind the curve. The nation becomes the victim of a divisive political climate, a lack of understanding, and a general distrust for big corporations. What remains is a door wide open to nefarious actors and emerging threats, the vast majority of which are not yet completely understood nor accounted for in the law.

Although guns, guards, and gates are still the first lines of defense along a fenceline, they can provide a false sense of security. In February 2016, two airport employees in Somalia facilitated the transfer of a sophisticated bomb built into a laptop through airport security, where it was carried onto plane and detonated (Kriel and Cruickshank, 2016). Malware stormed the Ukrainian industrial control systems (ICSs) in 2015, 2016, and again in 2017, the latter effectively shutting down the government and key critical functions, including the radiation monitoring system at the Chernobyl nuclear power plant (Greenberg, 2018). In September 2019, a series of drone attacks at Saudi Aramco oil processing facilities in Abqaiq and Khurais in eastern Saudi Arabia forced the country to shut down half of its oil production capacity, to the tune of 5.7 million barrels per day (Li, 2019). This attack was a major blow to both the country and the global markets. Meanwhile, extremist environmental justice activists are turning pipeline valves, potentially causing harmful spills, injuries, and catastrophic explosions that could cripple the very communities they seek to protect (Williams, 2016).

Although some of these examples did not occur in the United States by homegrown actors, they all could have, and these are real cases of activities that skirt existing U.S. laws and regulations intended to safeguard critical infrastructure from domestic terrorist activity. Currently, U.S. critical infrastructure owners and operators have little-to-no recourse when it comes to such threats. The nation must understand that homegrown violent extremists (HVEs) are capable of carrying out attacks previously perpetrated by foreign-born terrorists. A reactive approach to terrorism does little to thwart the threat, and near misses can very quickly turn into hits without the proper statutes in place to give relevant authorities the license to enhance their pre-attack intelligence gathering and investigative efforts. There must be accountability in the legal and regulatory framework to investigate and prosecute those persons who tamper with, attack, and threaten any critical infrastructure operation in the United States.

Understanding Domestic Terrorism

The concept of domestic terrorism in the United States is complex. Domestic terrorist activities are often understood as mass shootings at soft targets and crowded places, and certainly qualify as terrorism. However, this only represents one aspect. As the lead agency for investigating terrorist activity, the Federal Bureau of Investigation (FBI) classifies domestic terrorism as U.S. persons who commit criminal acts based on their “political, religious, social, racial, or environmental” ideologies, rather than for monetary purposes (FBI, 2019). This definition illuminates an important component of domestic terrorism that is often understated in the context of threat analysis: a criminal act does not need to result in mass casualties to be investigated as an act of terrorism. From a prosecutorial standpoint, the Department of Justice (DOJ) views domestic terrorism as activities that:

(A) involve acts dangerous to human life that are a violation of the criminal laws of the United States or of any State;

(B) appear to be intended –

(i) to intimidate or coerce a civilian population;

(ii) to influence the policy of a government by intimidation or coercion; or

(iii) to affect the conduct of a government by mass destruction, assassination, or kidnapping; and

(C) occur primarily within the territorial jurisdiction of the United States [emphasis added] (18 U.S. Code § 2331.5)

Given this broad definition, there is no single crime of “domestic terrorism.” Gaps exist in the federal statute and are ripe for exploitation, particularly via emerging threats. Although a person or group involved in a targeted attack will not escape prosecution, the charge may not fall under the federal terrorism statute. State criminal laws offer some avenues for penalizing terrorist activity, but punishments vary greatly state-to-state.

Chapter 113B of Title 18 of the U.S. Code remains the federal statutory guide for terrorism, but the crimes described therein are somewhat limited to foreign terrorism in that many require a transnational or foreign element in the fact pattern. Although a few can apply to domestic terrorist activity – most notably through an interstate commerce element – the language in these statutes do not effectively capture emerging threats. They cover the use of traditional explosives or assume that lone actors and small, very loosely coalesced groups cannot easily rise to the level of active terrorists. In today’s threat landscape, such assumptions create an environment in which terrorists can use emerging technologies and techniques to maximize devastation with limited effort and at minimal cost.

Emerging Threats

Over the last several decades, domestic terrorists have targeted critical infrastructure to advance political or social justice agendas. However, some maintain that domestic terrorists lack the organizational capacity and technical wherewithal to accomplish any meaningful attack on critical infrastructure (Riedman, 2017). This argument is myopic. Although it is true that the number of successful attacks against critical infrastructure in the United States is historically small, it is not prudent to rest on our laurels with respect to the nation’s critical functions. 3 A successful attack need not be from a formally organized group. In fact, the FBI notes that current threat actors are typically “autonomous and lone offenders, and small cells pose the greatest threat” (McGarrity and Brzozowski, 2019). Evolution and access to technologies, open markets, and the dark web make homegrown terrorists just as organized as the international groups traditionally associated with terrorist activity. Earlier this year, Assistant Director of the Counterterrorism Division at FBI, Michael McGarrity, noted domestic terrorism is “on the rise” (Levine, 2019). Indeed, three factors actively contribute to the growth and evolution of this threat landscape according to the FBI: the internet, use of social media, and HVEs.4

In 2018, the FBI investigated 50 reported incidents, threats, or suspicious activity at pipelines alone (McGarrity and Brzozowski, 2019). In that same year, the FBI investigated 87 reported threats to refineries (McGarrity and Brzozowski, 2019). These are real, credible threats. In fact, on September 19, 2019, a federal grand jury returned an indictment charging two women with knowingly and willfully damaging and attempting to damage the Dakota Access Pipeline, causing “a significant interruption and impairment of a function of an energy facility” (DOJ, 2019a). The women made no secret of their efforts to sabotage the pipeline to advance a political agenda, by burning exposed valve sites and attempting to pierce portions of empty pipeline with torches (Schiano, 2019). These acts are dangerous to human life, in violation of the laws of the country and state, and are intended to influence the policy of the government. However, there is no mention of domestic terrorism in the indictment because the federal terrorism statute does not account for these types of attacks. Prosecutors must instead rely on other criminal laws to charge offenders - most of which carry far more lenient sentences.

Although these women operated as part of a larger extremist environmental justice movement, lone or small cell insider threats are another increasingly significant concern in the security space. Apart from a disgruntled employee carrying out a devastating, mass casualty attack on facility property, insider threats can take the form of economic espionage, cyber hacks, and – for these purposes – actions that cause or can cause catastrophic physical damage to facility assets and systems. In December 2013, a former avionics technician entered a secure gate at the Wichita airport using a valid employee access card and attempted to detonate a car bomb. He later pled guilty to one count of use of a weapon of mass destruction – a federal terrorism charge – only because his well-documented ties to violent jihad very clearly revealed a transnational element in the fact pattern of the case (DOJ, 2015).

Now imagine the following scenario: the largest refinery in the United States at Port Arthur, TX undertakes a turnaround to repair a major process unit. In these instances, the regular permanent staff of approximately 1,450 might double to include contractors, temporary help, and outsourcers. A turnaround is a major security risk in and of itself. The regulatory framework exists to help mitigate these risks through certain measures within the Department of Homeland Security’s (DHS) Chemical Facility Antiterrorism Standards (CFATS) and the Transportation Safety Administration’s (TSA) Transportation Worker Identification Card (TWIC) system, but there are significant flaws in both of these programs. 5 Even so, one contractor with a hidden political agenda whose background check failed to detect any terrorist affiliations or criminal history, with authorized temporary access to a refinery or chemical facility’s operational technology (OT) could, with relative ease, facilitate a devastating event that does not involve an explosive device like a car bomb. Such a situation may seem unrealistic, but it is not. Although asset owners and operators do their best in good faith to comply with regulations to manage these security and safety risks, the sheer volume of employees in large turnaround-type situations creates a statistical advantage for the lone assailant. Nonetheless, depending upon the type of attack, such as disabling OT controls to cause a noxious chemical leak versus use of an improvised explosive device, the crime itself may not be prosecuted under federal terrorism laws.

As illustrated, just one person can cause an incident at a critical infrastructure facility, and this is particularly true for cyberattacks. The Department of Energy (DOE) and DHS are actively conducting outreach to educate stakeholders on the ramifications of cyber intrusions on physical security and safety. It is well known that hackers can steal sensitive data with relative ease (Yahoo, 2013 and 2014; Target, 2013; Marriot, 2018; Facebook, 2019; Capital One, 2019; Ecuador, 2019, among others), and remarkably, these reported data breach incidents are occurring with increasing frequency. In addition to sensitive data, a hacker can also access and manipulate ICSs in the United States. A couple of young computer scientists readily demonstrated this by hacking into a 2014 Jeep from more than 10 miles away and remotely taking over the controls (Greenberg, 2015). Certainly, a nuanced hacker with targeted phishing campaigns to a facility’s third party suppliers knows how to capitalize on that access to open gates, shut valves, and bridge both the information technology (IT) and OT systems so it becomes possible not to recognize that an attack is occurring until an explosion happens.

Cyberattacks are such a powerful tool that they are increasingly used as part of the United States’ military strategy (Schneider, 2019). Still, given that the U.S. is inextricably tied to the global communications infrastructure, the nation’s own vulnerabilities are countless. In the critical infrastructure world, companies work tirelessly to uncover and correct these vulnerabilities before they are exploited. Even so, cybersecurity cannot really be guarded by prescriptive regulations because innovation is stifled and adaptation in response to new attack vectors is limited. Technologies spread quickly. Within the last 40 years, the world shifted from mainframes to desktops to laptops to mobile devices, then the cloud, and now the “Internet of Things” (Danzig, 2014, 2018). Terrorists, domestic and foreign, exploit these rapid advancements. Acting Director of National Intelligence (DNI) Joseph Maguire recently called attention to this very issue, observing that: “At one point in time, you had to be a sovereign nation to have this kind of technology, but with the proliferation of technology and with the global economy, much of it is now easy to acquire and simple to use” (Cruickshank and Dodwell, 2019, 2011).

Not only is cyberterrorism challenging to attribute, but it is not defined anywhere in the federal criminal code. This is a recognized problem, both for purposes of understanding what constitutes cyberterrorism and for prosecuting offenders. Acting DNI Director Maguire continues, “[Terrorists] use the internet and encryption to a great extent. They understand technology. We are a technological nation, and we have to make sure we understand the problem set and not be reactive but be anticipatory to what they’re going to do” (Cruickshank and Dodwell, 2019, 12). As cyberattacks continue to increase and become more sophisticated, safeguarding against cyberterrorism is no longer exclusively a matter of ensuring that private industry is equipped with the appropriate tools to mitigate the risks. The government must assure its citizens, as well as the owners and operators of critical infrastructure assets, that those who perpetrate cyberterrorism are held accountable.

One final emerging threat to critical infrastructure that is increasingly worrisome is drones. The potential safety hazards and security threats presented by errant or malicious unmanned aircraft systems (UAS) activity and the evolving tactics used by hostile operators are provoking a growing number of efforts by public and private sector entities to address these risks. Not only can drones drop explosives and hazardous substances, but they can also be equipped with weapons, conduct unauthorized surveillance, aid hackers in overcoming physical barriers, and act as kamikaze agents for nefarious actors.

The potential for UAS activity to inhibit or halt operations at critical infrastructure facilities is known, as evidenced by recent disruptions to operations at Gatwick Airport in the United Kingdom (December, 2018), Newark Liberty International Airport (January, 2019), and most recently the attacks on Saudi Arabian oil and natural gas infrastructure (September, 2019). At the root of the challenges with UAS activity is the absence of a meaningful regulatory and legal framework.6 While a catastrophic act performed by a drone could potentially warrant a terrorism charge against the operator, critical infrastructure owners and operators are severely restricted in their ability to defend against these emerging technologies, creating an enormous security and safety risk for assets and personnel.

#### Infrastructure disruptions ripple---extinction.

Dennis Pamlin & Stuart Armstrong 15. Dennis Pamlin, Executive Project Manager Global Risks, Global Challenges Foundation, and Stuart Armstrong, James Martin Research Fellow, Future of Humanity Institute, Oxford Martin School, University of Oxford. February 2015. “Global Challenges: 12 Risks that threaten human civilization: The case for a new risk category,” Global Challenges Foundation, https://api.globalchallenges.org/static/wp-content/uploads/12-Risks-with-infinite-impact.pdf

Global Challenges – Twelve risks that threaten human civilisation – The case for a new category of risks 89 3.1 Current risks System Collapse 3.1.5 Global Global system collapse is defined here as either an economic or societal collapse on the global scale. There is no precise definition of a system collapse. The term has been used to describe a broad range of bad economic conditions, ranging from a severe, prolonged depression with high bankruptcy rates and high unemployment, to a breakdown in normal commerce caused by hyperinflation, or even an economically-caused sharp increase in the death rate and perhaps even a decline in population. 310 Often economic collapse is accompanied by social chaos, civil unrest and sometimes a breakdown of law and order. Societal collapse usually refers to the fall or disintegration of human societies, often along with their life support systems. It broadly includes both quite abrupt societal failures typified by collapses, and more extended gradual declines of superpowers. Here only the former is included. 3.1.5.1 Expected impact The world economic and political system is made up of many actors with many objectives and many links between them. Such intricate, interconnected systems are subject to unexpected system-wide failures due to the structure of the network311 – even if each component of the network is reliable. This gives rise to systemic risk: systemic risk occurs when parts that individually may function well become vulnerable when connected as a system to a self-reinforcing joint risk that can spread from part to part (contagion), potentially affecting the entire system and possibly spilling over to related outside systems.312 Such effects have been observed in such diverse areas as ecology,313 finance314 and critical infrastructure315 (such as power grids). They are characterised by the possibility that a small internal or external disruption could cause a highly non-linear effect,316 including a cascading failure that infects the whole system,317 as in the 2008-2009 financial crisis. The possibility of collapse becomes more acute when several independent networks depend on each other, as is increasingly the case (water supply, transport, fuel and power stations are strongly coupled, for instance).318 This dependence links social and technological systems as well.319 This trend is likely to be intensified by continuing globalisation,320 while global governance and regulatory mechanisms seem inadequate to address the issue.321 This is possibly because the tension between resilience and efficiency322 can even exacerbate the problem.323 Many triggers could start such a failure cascade, such as the infrastructure damage wrought by a coronal mass ejection,324 an ongoing cyber conflict, or a milder form of some of the risks presented in the rest of the paper. Indeed the main risk factor with global systems collapse is as something which may exacerbate some of the other risks in this paper, or as a trigger. But a simple global systems collapse still poses risks on its own. The productivity of modern societies is largely dependent on the careful matching of different types of capital325 (social, technological, natural...) with each other. If this matching is disrupted, this could trigger a “social collapse” far out of proportion to the initial disruption.326 States and institutions have collapsed in the past for seemingly minor systemic reasons.327 And institutional collapses can create knock-on effects, such as the descent of formerly prosperous states to much more impoverished and destabilising entities.328 Such processes could trigger damage on a large scale if they weaken global political and economic systems to such an extent that secondary effects (such as conflict or starvation) could cause great death and suffering. 3.1.5.2 Probability disaggregation Five important factors in estimating the probabilities of various impacts: 1. Whether global system collapse will trigger subsequent collapses or fragility in other areas. 2. What the true trade-off is between efficiency and resilience. 3. Whether effective regulation and resilience can be developed. 4. Whether an external disruption will trigger a collapse. 5. Whether an internal event will trigger a collapse. 1. Increased global coordination and cooperation may allow effective regulatory responses, but it also causes the integration of many different aspects of today’s world, likely increasing systemic risk. 2. Systemic risk is only gradually becoming understood, and further research is needed, especially when it comes to actually reducing systemic risk. 3. Since systemic risk is risk in the entire system, rather than in any individual component of it, only institutions with overall views and effects can tackle it. But regulating systemic risk is a new and uncertain task. 4. Building resilience – the ability of system components to survive shocks – should reduce systemic risk. 5. Fragile systems are often built because they are more efficient than robust systems, and hence more profitable. 6. General mitigation efforts should involve features that are disconnected from the standard system, and thus should remain able to continue being of use if the main system collapses 7. A system collapse could spread to other areas, infecting previously untouched systems (as the subprime mortgage crisis affected the world financial system, economy, and ultimately its political system). 8. The system collapse may lead to increased fragility in areas that it does not directly damage, making them vulnerable to subsequent shocks. 9. A collapse that spread to government institutions would undermine the possibilities of combating the collapse. 10. A natural ecosystem collapse could be a cause or consequence of a collapse in humanity’s institutions. 11. Economic collapse is an obvious and visible way in which system collapse could cause a lot of damage. 12. In order to cause mass casualties, a system collapse would need to cause major disruptions to the world’s political and economic system. 13. If the current world system collapses, there is a risk of casualties through loss of trade, poverty, wars and increased fragility. 14. It is not obvious that the world’s institutions and systems can be put together again after a collapse; they may be stuck in a suboptimal equilibrium. 15. Power grids are often analysed as possible candidates for system collapse, and they are becoming more integrated. 16. The world’s financial systems have already caused a system collapse, and they are still growing more integrated. 17. The world’s economies are also getting integrated, spreading recessions across national boundaries. 18. The world’s political and legal systems are becoming more closely integrated as well. Any risk has not been extensively researched yet, and there remain strong obstacles (mainly at the nation state level) slowing down this form of integration. 19. The politics of the post-system collapse world will be important in formulating an effective response instead of an indifferent or counterproductive one. 20. System collapses can be triggered internally by very small events, without an apparent cause. 21. External disruptions can trigger the collapse of an already fragile system. 22. The trade-off between efficiency and resilience is a key source of fragility in a world economy built around maximising efficiency. 23. Climate change, mass movements of animals and agricultural mono-cultures are interlinking ecosystems with each other and with human institutions. 24. There is a lot of uncertainty about systemic risk, especially in the interactions between different fragilities that would not be sufficient to cause a collapse on their own.

### 1AC---Democracy ADV

#### Contention 2 is Democracy.

#### Unchallenged Big Tech enables the erosion of democratic norms and rise of autocratic tendencies through numerous vectors.

Guy Schleffer & Benjamin Miller 21, PhD, Political Science, University of Haifa; Professor, International Relations, University of Haifa. Director, National Security Center, University of Haifa, "The Political Effects of Social Media Platforms on Different Regime Types," Texas National Security Review, Summer 2021, <https://tnsr.org/2021/07/the-political-effects-of-social-media-platforms-on-different-regime-types/>. language edited.

Today, for the first time since 2001, there are more autocracies than democracies in the world. The number of electoral and liberal democracies dropped from 55 percent of all countries at its peak in 2010 to only 48 percent in 2019.1 This decline in the number of liberal democracies is [shredding] ~~crippling~~ the U.S-led liberal world order, weakening America’s post-Cold War hegemony, and shifting power to authoritarian regimes such as China and Russia, leading Fareed Zakaria to claim that “American hegemony died.”2 It seems that illiberal democracies and authoritarian regimes are on the rise all across the globe. Even in the United States, President Donald Trump favored a new kind of hegemony — an illiberal one.3

This phenomenon can be explained by many factors, such as the rise of xenophobic populist movements in reaction to immigration, cultural change, the decline in job and economic security after the 2008 financial crisis, the opposition to globalization, and the loss of sovereignty.4 For authoritarian regimes, the growth of national populist movements in Europe and America is proof that “the liberal idea has outlived its purpose,” as the public has turned against immigration, open borders, and multiculturalism.5 Recent elections worldwide reflect a deep groundswell of anti-establishment sentiment that can easily be mobilized by extremist political parties and candidates.6

In this essay, we highlight the role of U.S. social media platforms in the decline of liberal democracies and the rise of illiberal democracies and autocratic regimes across the world. Previous studies have examined different aspects of the impact of social media platforms on states and regimes but have usually done this in a siloed way — for example, narrowly focusing on the correlation between social media and the rise of populism, or on the influence of social media in Latin America and in the U.S. 2016 elections.7 Sarah Kreps posed a broader research question regarding the different effects of social media platforms in democracies and non-democracies.8 This essay contributes to this body of research by taking an even wider view. We describe the multifaceted impact of American social media platforms in different countries and seek to generalize and categorize this impact by offering an innovative explanation for it. These generalizations and categorizations help to explain why different countries experience the impact of social media in different ways and with varying degrees of intensity. They can also help policymakers to predict what may occur in similar countries where the impact is not yet clear and help them to take preliminary steps to avoid the same phenomenon.

We begin by reviewing social media platforms, their knowledge power (the power they gained by accumulating vast amounts of data about people and turning it into knowledge), and their potential as a liberating mechanism, before examining what went wrong in the last decade. This is followed by a discussion of the variable impacts of social media on the political systems in different states. We explore why the effect varies between different democratic states — for example, in Brazil compared to the United States — as well as between different authoritarian states — i.e., why it has bolstered Russia’s regime while fueling a revolution in Egypt. We suggest that social media’s impact depends mainly on three political actors: domestic opposition, external forces, and the governing regime. By examining how these three actors use social media while considering variations in state capacity and political regime type, a causal model emerges in which there are four different effects of social media: weakening, intensifying, radicalizing, and destabilizing.

We further examine these disparate effects by examining four case studies, each representing a different combination of state capacity (weak vs. strong) and regime type (democratic vs. authoritarian). We argue that, contrary to the optimistic promise of social media platforms at the beginning of the millennium, it seems that they are having a weakening effect on strong democratic regimes, an intensifying effect on strong authoritarian regimes, a radicalizing effect on weak democratic regimes, and a destabilizing effect on weak authoritarian ones. We conclude by presenting the implications of our analysis for the future of the international system.

The Promise of Social Media

Of the seven most popular social media platforms, Facebook owns four: Facebook, Messenger, WhatsApp, and Instagram. Together with YouTube, which is owned by Google, these are the five leading social media platforms not based in China. The most successful social media platform in “grabbing, holding, and processing human attention” is WeChat,9 a China-based application that “encompasses almost every aspect of human life.”10 It is a “one-stop-shop” model that led Facebook to try to consolidate its sub-companies (Facebook, Messenger, WhatsApp, and Instagram) into one giant application.11

Optimists have seen social media platforms as an expression of the liberalizing ethos of the internet: tools for empowering citizens, enabling economic opportunities, increasing freedom of expression, spreading liberal ideas, and providing an alternative communication platform for dissidents.12 This positive view was espoused by some of the founders of U.S. social media platforms and can be traced back to John Perry Barlow’s “Declaration of the Independence of Cyberspace,”13 which was popular at the time when these companies were established in Silicon Valley. Although these corporations started out politically neutral, some have moved in recent years toward publicly challenging governments.14 For example, Facebook founder Mark Zuckerberg has talked about replacing the “old” social infrastructure of the state, “which opposes the flow of knowledge, trade and immigration,”15 with a new global community.16 Zuckerberg also stated: “In many ways, Facebook is more like a government than a traditional company.”17 Google’s Jared Cohen and Eric Schmidt wrote about the game-changing implications of the internet for politics. They predicted that governments “will be caught off-guard when large numbers of their citizens, armed with virtually nothing but cell phones, take part in mini-rebellions that challenge their authority.”18

Social media platforms have the power to strengthen democracies by echoing public opinion. Clay Shirky argues that social media can help to increase freedom and change people’s political views by exposing them to other opinions echoed by friends, family members, and colleagues.19 At the beginning of the millennium, social media platforms were credited with shifting power from authoritarian regimes to ordinary people seeking freedom and social justice.20 Peter Singer and Emerson Brooking wrote in 2018 that social media platforms “illuminated the shadowy crimes through which dictators had long clung to power and offered up a powerful new means of grassroots mobilization.”21 Manuel Castells describes social media as “a mobilizing force” that can “topple an entrenched regime if everybody would come together.”22 These platforms can compensate for the disadvantages of undisciplined groups by reducing coordination costs while increasing shared awareness.23 Indeed, social media platforms played a role in the 2009 civil revolt in Moldova, dubbed “the first Facebook revolution”; the 2009 unrest in Iran, called “the first Twitter revolution”; the 2011 Russian “almost-revolution”; and the first wave of Arab social unrest in 2011, when “the Facebook-armed youth of Tunisia and Egypt” demonstrated “the liberating power of social media.”24 However, according to Singer and Brooking, these internet-enabled democratic movements “represented a high-water mark” that was followed by “a countering wave of authoritarianism using social media itself, woven into a pushback of repression, censorship and even violence.”25

The use of social media has no single preordained outcome. These platforms cannot “bring the world closer together,” as Facebook’s mission states,26 and help connect only democracy-loving people. As Zeynep Tufekci notes, they are also “connecting white supremacists, who can now assemble far more effectively or radical Buddhist monks in Myanmar, who now have much more potent tools for spreading incitement to ethnic cleansing.”27 Social media can be used to support incumbent politicians within a country or to help external authoritarian powers to disseminate propaganda and disrupt the democratic transfer of power through elections in other countries.28 It is also used by populists who pose a fundamental challenge to neoliberal ideology, spreading untruth and stirring outrage that affects voters’ judgment and fuels partisanship.29 The different actors using social media platforms, whether for good purposes or bad, are exploiting the unprecedented concentration of knowledge power that these platforms have amassed over the past few years.

The Knowledge Power of Social Media

In recent years, social media platforms have gained “knowledge power” derived from the vast amounts of data that they have collected and marshaled.30 According to Susan Strange, such power includes “what is believed or known and the channels by which these beliefs, ideas and knowledge are communicated, or confined.” This kind of power lies as much in the capacity to deny knowledge as in the power to convey it.31

The knowledge power of social media platforms may take many forms. Facebook, for example, knows more about a person than the government does.32 In 2002, Google discovered it could use the collateral data that it collects to profile users based on their characteristics and interests and then match advertisements to individual users.33 Over the years, Google and Facebook have sold more ads by reducing user privacy and gaining more access to a person’s data.34 In the competition for what Shoshana Zuboff called “surveillance capitalism” revenue,35 the advantage goes to firms that can acquire vast and varied data streams. Therefore, social media platforms are expanding both the scope of surveillance (migrating from the virtual world into the real world of automobile dashboard) and the depth of the surveillance (accumulating data on individuals’ personalities, moods, and emotions).

In addition to using knowledge power to profile and micro-target their users to sell more ads, Facebook also uses its algorithms to anticipate human behavior and create “prediction products” that make people easier to manipulate.36 This power was allegedly harnessed to reshape popular perceptions of the 2016 U.S. presidential election and the United Kingdom’s referendum on membership of the European Union.37

Another aspect of social media’s knowledge power is reflected in its significant role in today’s media industry. The perceived trustworthiness of the news media in democratic states has given these states advantages over non-democratic ones. Lucie Greene calls Facebook, Twitter, and Google “the Fifth Estate” because they have replaced the traditional news outlets as the main places where people go to get their news.38 They now have the power to shape public life, including what content is produced, where audiences go, and what news and information citizens see.39

In 2012, Facebook declared that its mission is to expand and strengthen relationships between people and to help expose people to a greater number of diverse perspectives.40 Instead, only a few years later, the opposite has happened. Facebook has became one of the sources for divisions among people.41 This can be attributed to two main factors: the “filter bubble phenomenon” and the rise of fake news.

Facebook’s algorithms tend to reinforce a “filter bubble” that shields people from dissenting information and only delivers content that confirms their views.42 Social media platforms are part of the digital “attention economy,” which focuses on the interplay between money and attention. The more people are engaged with the content on social media and are exposed to commercial ads, the more it generates income for these platforms. In order to keep people engaged, Facebook tends to expose them to the most popular posts and to confrontational and inflammatory news items that tend to make people more extreme in their views.43 Facebook encourages society to self-segregate into like-minded communities, which increases the distance between groups with opposing views, causing more polarization.44 YouTube’s recommendation algorithm typically recommends videos that echo the political bias of its viewers and what they choose to view, and feeds them videos containing viewpoints that are more extreme than the ones they currently hold.45

Fake news has gained prevalence in recent years due to the rising role of social media platforms as news outlets, where content can be produced and relayed among users with no significant third-party filtering, fact-checking, or editorial judgment. Hunt Allcott and Matthew Gentzkow define fake news as “news articles that are intentionally and verifiably false, and could mislead readers.”46 This type of news is widespread because it is cheaper to produce than precise reporting and because consumers enjoy partisan news. A fake story shared by millions becomes “real” because people believe that if it’s going viral, it must be true. The most inflammatory materials will travel the farthest and fastest. False stories on Twitter, for example, spread significantly faster and more broadly than true ones, and wider distribution of false stories also makes them more profitable for social media platforms.47

Fake news finds fertile ground in a divided electorate that has clear in-groups and out-groups, where people are ready to accept any statement as long it is consistent with what they already believe.48 Extreme examples of fake news spread by social media platforms can be found in Myanmar and Sri Lanka, where the dissemination of hate speech contributed to the ethnic cleansing of Rohingya Muslims and anti-Muslim riots, respectively.49

In the last several years, political actors have begun to use the knowledge power of social media to their advantage. A 2016 Rand study discusses the “firehose of falsehood” — a high-intensity stream of lies, partial truths, and complete fictions that impacted several democratic elections, including in Ukraine, Italy, France, Germany, and the United States.50 The “firehosing” that took place in America, for example, included attempts to influence public opinion and promote political protests.51 According to some scholars, authoritarian and illiberal states started using social media platforms to spread fake information to exercise their “sharp power.”52 This sharp power can stifle productive discussion in democracies, deepen domestic polarization, exacerbate ethnic tensions, rekindle nationalism, weaken public confidence in both journalism and elections, and diminish the overall influence of the Western-led international system.53 Authoritarian and illiberal regimes also use social media knowledge power, together with artificial intelligence, as a monitoring tool, allowing them to collect and analyze vast amounts of data on entire populations. Such regimes also undercut the credibility of valid information sources by using “bot-fueled campaigns of trolling and distraction, or piecemeal leaks of hacked materials, meant to swamp the attention of traditional media.”54 Once citizens learn to assume that the regime’s fake information is true, they alter their behavior without the regime having to resort to physical repression.55

But it is not only authoritarian and illiberal states that use fake news to deepen domestic polarization, radicalize people’s politics, and rekindle nationalism. This also occurs in democracies. Some democratic countries are experiencing a rise in populist leaders, fueling a drift toward national-populism, illiberalism, and even autocracy. According to Adrian Shahbaz and Allie Funk, populists and far-right extremists exploit social media platforms to “build large audiences around similar interests, lace their political messaging with false or inflammatory content, and coordinate its dissemination across multiple platforms.”56 Paolo Gerbaudo argues that social media is attractive to populists because it provides an outlet for countering the perceived pro-establishment bias of mainstream news media. The filter bubble helps individuals who are politically disgruntled to congregate online and mobilize militant support for anti-establishment candidates.57 The unregulated social media platforms are thus converted into instruments for political distortion and societal control.58

In conclusion, social media can play a positive or a negative role: It can be a liberalizing tool, used to spread information and knowledge, but it can also be a tool of suppression, used to disseminate distorted information and fake news. Grassroots movements and freedom fighters can make use of social media platforms, but so can authoritarian regimes.

The Variable Impact of Social Media Platforms

Although social media platforms are used throughout the world, they seem to have a different impact on the political system in democratic regimes as opposed to authoritarian regimes. What accounts for this difference? Why do they spark revolutions in some states while supporting the rise of populist candidates in others? Why do they disrupt democratic elections in one country, but support the regime’s anti-democratic measures in another? In order to understand the varying effects of social media, it’s important to distinguish clearly between liberal and illiberal regimes. In addition, it is necessary to consider whether those regimes are strong or weak.

Regime Types and Capacity

We used the “freedom score” calculated by Freedom House and the Regimes of the World typology based on V-Dem data to distinguish between democratic and illiberal or authoritarian states.59 Countries are classified as democratic if they de facto hold free and fair multiparty elections and also guarantee freedom of speech and expression.60 A liberal democracy is also characterized by having effective legislative and judicial oversight of the executive and protecting individual and minority liberties and the rule of law.61 In contrast, an authoritarian regime is characterized by a government that permits people only a limited degree of political freedom. In such regimes, the government controls the political process and determines individual freedoms without any constitutional accountability.62

The “freedom score” is based on two main parameters — political rights and civil liberties — and it can show trends in a state’s score over the last several years. Democratic states (“free” in Freedom House’s terminology) are those with a freedom score of at least 70, while authoritarian states (“not free”) are those with scores of 35 or less. Since there are some anomalies around the transition points (70 and 35) for the case studies in this article, we chose democratic states with scores over 74 — Brazil (75), the United States (86), and the United Kingdom (94) — and authoritarian states with scores under 25 — Russia (20), China (10), Egypt (21), and Iran (17). The only exception is Mexico, whose freedom score of 62 represents a deteriorating democratic state. The Regimes of the World typology confirms the above classification of the United States and the United Kingdom as liberal democracies; Brazil and Mexico as electoral democracies; Russia, Iran, and Egypt as electoral autocracies; and China as a closed autocracy.

While the distinction between democratic and authoritarian regimes can help to explain the variable impact of social media on the political system, a more powerful explanation emerges when we add state capacity to the equation. State capacity refers to the institutions and resources available to states for governing the polity. In this article, we distinguish between strong or capable states and weak or malfunctioning ones. In contrast to weak states, strong states possess well-functioning institutions and sufficient resources to carry out policies and major functions. Especially important is a state’s justice system and its monopoly over the means of violence (i.e., its coercive capacity). Thus, a strong state can supply security in its sovereign territory. Such a sense of security sustained over time can facilitate business and commerce.63 Weak states, on the other hand, lack the effective political institutions and resources to implement their policies, protect their populations from violent conflict, and deliver political goods.64

No single index can distinguish strong states from weak ones. State strength or capacity can be measured, among other things, by a state’s capacity to provide services to the whole population, the extent to which infrastructure and communication networks cover the state territory, and the level of the state’s control over its sovereign territory. It can also be measured through gross domestic product (GDP) per capita, GDP growth, and trade indicators because states with higher economic development enjoy larger pools of resources from which to extract taxes.

For this paper, we used three different indicators to determine whether a state is weak or strong: the Fund for Peace’s Fragile States Index,65 which takes into account political stability indicators such as the security apparatus, group grievances, uneven economic development, state legitimacy, and external interventions (which reflects the state’s coercive capacity);66 the World Bank data on Gross National Income, because, in an economically developed state, taxes constitute a significant portion of the annual income; and total military expenditures (as calculated by the International Institute for Strategic Studies), which indicate to what extent a given state is capable of mobilizing manpower for military service. Any one of these indices is not enough on its own to determine whether a state is strong or weak. China, for example, would not be considered a strong state according to the Fragile State Index or GDP per capita, but if one considers Gross National Income and total military expenditures, it is. Similarly, Brazil is not a typical weak state by the Fragile State Index, but when examining its military expenditure, which is much lower than that of strong states, it cannot be considered a strong state.

A Causal Explanation of Social Media’s Impact

In our analysis of the variations in the effects of social media platforms on states, we consider only American social media platforms that wield knowledge power, usually acquired through data collection and processing. We do not consider other social media platforms, such as the Russian platform VK or China’s WeChat, because they are more culture-dependent and less common outside their home countries. U.S. social media platforms are used in different ways by three main types of political actors: domestic opposition (dissidents or populist candidates), external forces (other countries or multinational corporations), and the governing regime. These actors use social media for political purposes in an effort to influence a state’s political system. But the outcome is beyond their reach because it is determined by the combined effect of the state’s capacity and the state’s regime. This essay will analyze these actors, together with the differences in state capacity and regime, to deduce a causal model that indicates four potential effects of social media platforms on states: destabilization, radicalization, intensification, and weakening.

Destabilization takes place in weak authoritarian regimes when social media platforms facilitate the coordination and mobilization of dissidents and grassroots movements (which represent domestic opposition) in resisting the government’s tyranny. The governing regime in these cases usually lacks the coercive capacity to maintain internal order and stop well-coordinated resistance. This destabilization effect can ultimately lead to regime change or, in extreme cases, a failed state scenario. A recent example of this destabilizing effect, which resulted in regime change, is the 2011 revolution in Egypt that culminated in President Hosni Mubarak’s ouster.

The use of social media platforms as a disseminator of fake news and disinformation during an election process leads to a radicalizing effect when it occurs in weak democratic countries. Social media can be used by populist and anti-establishment candidates (domestic opposition) and may diminish democratic institutions and processes. Some of these practices continue even after these candidates are elected to office. Social media platforms may also be malignly exploited by external forces — Russia, for example — to disturb elections in democratic countries. The use of social media platforms is considered an integral part of the democratic election process because the platforms are vehicles for exercising free speech. This limits the governing regime’s capacity to restrict and counter the malign use of these social media platforms. This radicalization effect could steer a state that lacks sufficient checks and balances and a strong democratic tradition toward becoming an illiberal or authoritarian regime. Recent examples of this radicalizing effect are the rise of Jair Bolsonaro in Brazil, Andrés Manuel López Obrador in Mexico, Viktor Orbán in Hungary, and Rodrigo Duterte in the Philippines.

The intensifying effect takes place in strong authoritarian countries when social media platforms become a surveillance tool of the governing regime, which uses them to intensify the government’s coercive capacity, suppress civil rights, and counter the domestic opposition. These platforms also enhance the state’s sharp power with regard to liberal-democratic states. Social media platforms do not usually have a substantial liberalizing effect in strong authoritarian countries because the regimes control the internet in their territory. Recent examples of the intensifying effect include laws in China and Russia that give the states the power to surveil their citizens’ activities on social media, and Russia’s intervention in democratic elections in several liberal-democratic states by using these platforms as disseminators of fake news and disinformation.67

The use of social media platforms by domestic populist forces or external malign forces can weaken domestic authority in strong democratic regimes, support the rise of populism, and diminish democratic institutions and ideas, such as multilateralism and globalization. Democratic norms in strong liberal regimes constrain these states from countering the malign use of social media, despite having the capacity to do so. In this scenario, the domestic political system becomes polarized, but the state’s established system of checks and balances and strong democratic tradition preserve its liberal character. Recent examples include the 2016 U.S. presidential elections and the U.K. Brexit referendum. Table 1 summarizes the four types of political effects that U.S. social media platforms have on states.

[Chart omitted]

In the following sections, we present four primary case studies to illustrate this model and describe the effects that social media platforms have on different states. The case selection is based on variations in the type of regime and the state’s capacity. The detailed case studies described in each classification are only the prominent ones in which the literature and the availability of empirical material regarding social media effects are more prevalent than others. The effect of social media on states is a relatively new research topic and the publications in this area are mainly concentrated on the limited case studies that we chose to present in this article. The case studies are the leading instances that represent the political effects of social media platforms on states: destabilizing, radicalizing, intensifying, and weakening.

For the destabilizing effect, we look at Egypt, a prominent country in the Arab world with a stable leader for three decades who was nevertheless quickly overthrown after less than three weeks of protests. To demonstrate the radicalizing effect, we chose Brazil, a leading economy in Latin America, a region that underwent a rapid democratization process in the last 20 years of the 20th century but that, in the last several years, has reverted to illiberalism and even to authoritarianism. For the intensifying effect, we considered the malign use of social media in the most powerful authoritarian states, Russia and China. For the weakening effect, we chose the external intervention in the 2016 U.S. presidential elections, which is considered to be a leading case of authoritarian intervention because of America’s status as the most powerful liberal democracy in the world. In each of these sections, we discuss additional examples, but it should be noted that not all countries in the same classification are necessarily affected by social media to the same extent as the anchor case.

The Destabilizing Effect in Weak Authoritarian Regimes

Social media platforms can help to create and mobilize domestic opposition to the governing regime by making it easier for opposition members to connect, organize, and circumvent the regime’s restrictions. Organized resistance to the government’s tyranny may lead to regime change in weak authoritarian regimes, but some governments put in place surveillance systems that monitor social media platforms, helping to counter dissidents early in their organization stage and reducing the risk of regime change.

In just over a year, a wave of unrest that first began in Tunisia in December 2010 swept through the Arab region, leading to the overthrow of four Arab heads of state.68 Power seemed to be shifting from authoritarian regimes to citizens. Social media platforms were credited with helping to cause this shift. Philip Howard and Muzammil Hussain explain that digital media (including social media platforms) helped to shape events and outcomes by spreading protest messages, connecting frustrated citizens, and helping them to realize that they shared grievances and could act together to do something about their situation.69 The protestors succeeded in building and mobilizing a spontaneous domestic opposition. Other researchers contend that the “Internet may be the only avenue left for citizens in authoritarian regimes to influence government, fight corruption, or defend their rights.”70 Laura Stein outlines six different ways in which the internet and social media platforms may help social movements: by providing information, assisting people in mobilizing and taking an action, promoting interaction and dialogue, helping to connect different networks of people, serving as an outlet for creative expression, and promoting fundraising and resource generation.71

The use of social media platforms is part of each stage of any uprising in the internet era. In the preparation phase, activists use social media platforms to find each other, build solidarity around shared grievances, and identify collective political goals. In the ignition phase, which involves some inciting incident, social media helps to publicize that incident and enrage the public. Take, for example, the pictures of Khaled Mohamed Saeed, who was beaten to death by police in Egypt, or Mohamed Bouazizi, who set himself on fire in Tunisia.72 In the street protests phase, the call for protests and the protest locations are coordinated online, while in the international buy-in phase, pictures, tweets, and videos from the uprising gain international interest and support. Usually, this pressures the rulers of the state to enter the climax phase, in which the state either cracks down and protesters are forced to go home (as in the case of Iran), rulers concede and meet public demands (as in Egypt and Tunisia), or the groups reach a protracted stalemate (as in Syria). The denouement largely depends on the state’s coercive power.73

Egypt serves as a good case study for the destabilizing effect of social media platforms. Hosni Mubarak ruled Egypt from 1981 until 2011, resigning only 18 days after the beginning of the Egyptian uprising, which started in January 2011. Scholars are divided on the role that social media platforms played in the Egyptian revolution. Killian Clarke and Korhan Kocak argue that Facebook and Twitter contributed meaningfully to mobilizing the “first movers” (the demonstrators who participated in the protest on Jan. 25, 2011) to form an ad hoc domestic opposition. These platforms helped to produce this outcome by recruiting people, planning and coordinating a leaderless protest, and providing live updates. The success across these three dimensions helped to convince many other Egyptians to join in subsequent protests, thus setting in motion a revolutionary cascade that resulted in Mubarak’s ouster.74 Juby John Eipe also underlines the significant role that Twitter played in initiating, organizing, and executing a powerful political movement in Egypt, including mobilizing people with no political background.75 Philip Howard and Muzammil Hussain note that YouTube and other video archiving platforms allowed citizen journalists, using mobile phone cameras and consumer electronics, to broadcast stories that the mainstream media could not or did not want to cover.76

Others are not so fast to give all the praise to social media. Tarak Barkawi is critical of the credit given to “Western technology” rather than to the “ordinary Egyptians, mothers and fathers, daughters and sons, who toppled the regime.”77 He argues that the West imagines itself as the real agent in the uprisings and he denounces “fantastically Eurocentric” narratives. “To listen to the hype about social networking websites and the Egyptian revolution, one would think it was Silicon Valley and not the Egyptian people who overthrew Mubarak,” Barkawi writes. Mason agrees that social networks allow people to assemble and protest but insists that the revolutions in the Arab world “have been social, political and real — not virtual.”78 According to Mohamed Ben Moussa, social media platforms were only effective because they operated in synergy with a huge array of “other more conventional media and offline societal networks.”79 Regina Salanova agrees that, in the end, Al-Jazeera and other international media “amplified the message, attracted the majority of the population to join the revolts and put pressure on the authoritarian states by engaging international audiences.”80 Ian Black notes that state surveillance of social media platforms compelled activists to use alternative media and communication tools.81

The Egyptian uprising is not the only example of social media playing a prominent role in political upheaval. Mark Pfeifle, a former U.S. national security adviser in the George W. Bush administration, wrote regarding the 2009 Green Revolution in Iran that “[w]ithout Twitter, the people of Iran would not have felt empowered and confident to stand up for freedom and democracy.”82 He also called for Twitter to be nominated for the Nobel Peace Prize. In the uprising in Tunisia, the blogosphere provided a forum for open political dialogue on regime corruption and the potential for political change.83 In Sudan’s 2019 uprising, social media platforms (Twitter, Instagram, Telegram, and Facebook) gave people an alternative source of information and an opportunity to organize and rebel against their government.84 This enabled dissent to spread from regional cities, such as Atbara, to Khartoum and elsewhere much faster.85 Social media platforms also helped diaspora communities to stay updated about events in Sudan and play an invaluable role in the uprisings by sharing updates and fostering solidarity.86

Again, some scholars have downplayed the impact of social media on these events. Golnaz Esfandiari wrote regarding the Green Movement after the 2009 elections in Iran: “Simply put: There was no Twitter Revolution inside Iran.”87 Bruce Etling and his coauthors agree that Twitter did not necessarily play a role in organizing the Iranian protests.88 Others attribute less importance to social media’s role in rallying local audiences and focus instead on the “bridging function” of social media platforms, which allows them to inform international audiences and mainstream media.89

Social media platforms can also be effective in bolstering authoritarian regimes,90 which may help explain why, in the years after the Arab Spring, there were fewer revolutions in weak authoritarian regimes. Authoritarian regimes, even those that lack financial resources, can now use social media surveillance tools that weren’t available to Egypt during its revolution to monitor and control society.91 This is due to the availability of various low-cost surveillance tools exported by China and Russia.92

The Iranian regime, which learned from the 2009 unrest, is an example of a weak authoritarian regime that took the use of social media platforms for surveillance to the next level. It combined using surveillance tools with exerting strict control over the domestic internet infrastructure. The development of Iran’s state-controlled National Internet Network significantly enhanced the government’s ability to restrict, block, and monitor internet use in Iran,93 providing it with one of the world’s most sophisticated mechanisms for controlling and censoring the internet and allowing it to examine the content of individual online communications on a massive scale.94 In 2009, mass surveillance operations significantly aided the authorities’ ability to identify, track, arrest, and imprison protesters.95 During the unrest that swept through Iran at the end of 2017, the authorities implemented major disruptions to internet access by slowing it down, blocking social media platforms (such as Instagram and Telegram) that were used by the protesters to mobilize street protests, and briefly cutting off Iranians’ access to the global internet. Some weak authoritarian governments have also learned to control the networked public sphere through “surveillance and repression, using fear, blocking of information, mobilizing armies of supporters or paid employees who muddy the online waters with misinformation, doubt, confusion and distraction.” This makes it hard for ordinary people to “navigate the networked public sphere and sort facts from fiction.”96 Instead of denying internet access to dissidents, which is sometimes difficult to do, these governments prefer to “deny attention, focus, and credibility.”97

Clarke and Koçak claim that social media platforms were, and still are, relevant because dissidents in authoritarian environments have simply switched to new social media platforms that the government hasn’t started monitoring yet.98 The new generation of dissidents uses messaging apps like WhatsApp and Telegram. Activists used these apps instead of Twitter in the 2018 revolt in Armenia and have used “Facebook live” for real-time coverage of anti-governmental protest activities in Nicaragua.99

In summary, social media can have a destabilizing political effect in weak authoritarian regimes. Social media (a liberalizing force) can help to create and mobilize domestic opposition by making it easier to alert and connect people who have shared interests. It also helps people to organize more easily and lets protesters know that they are not alone. Facebook’s filter bubbles may help to convince people that there is more support for their position than there really is, thus generating a self-fulfilling prophecy that drives people to the streets. Although social media may not be the only reason why a revolution takes place, it can certainly play a significant role. However, in the last several years, many weak authoritarian regimes have been able to afford the purchase of surveillance systems from China and Russia that monitor social media platforms and assist the repressive governing regime, deterring and countering dissidents during organization stage.100 The governing regime also uses social media platforms to spread misinformation, leading people to doubt what they read on these platforms and perhaps deterring them from joining a potential uprising.

The Radicalizing Effect in Weak Democratic Regimes

Social media is a low-cost and convenient communication tool that can be used by opposition populists to reach their supporters, by the governing regime to engage directly with the electorate, and by malign external forces to spread fake news. These platforms can be exploited to spread fake news and narratives that are polarizing, divisive, and anti-liberal because they lack the fact-checking found in traditional media outlets. Social media helps populists (both as candidates and as part of the governing regime) to aggregate and unify people to promote a shared cause against the liberal establishment and liberal freedoms and to erode democratic pillars. Malign external actors use social media to intervene in democratic elections in weak democratic countries to cause further erosion of trust in the democratic system. These combined actions create a radicalizing effect in weak liberal democracies that can potentially turn a liberal-democratic regime into an illiberal regime, or even an autocratic one.

There are several examples of weak liberal-democratic regimes worldwide, namely new democracies in Eastern Europe, East Asia, and Latin America. Freedom House’s freedom score has declined in some of these countries over the last several years: From 2016 to 2020 Brazil’s score dropped six points, from 81 to 75; the Philippines fell six points, from 65 to 59; and Mexico lost three points, going from 65 to 62. This decrease in freedom scores may be explained by the rise of populist leaders in these countries and the erosion of democratic pillars, such as free and unbiased elections.101 According to the Democracy Report 2020, Latin America has regressed to a level of democracy last recorded around 1992.102

In the last decade, Latin American presidents and candidates started using social media to engage directly with the electorate. Social media is perceived as the voice of the people and more authentic than the mainstream media, “which responds to the agenda of their super-rich owners and their political allies, rather than to the real needs and interests of the public.”103 By 2014, the region had the world’s highest use of social media by politicians.104

According to Emarketer, people in Latin America are the most avid social media users in the world.105 The vast majority of them get their news straight from social media services because they place less trust in traditional media. For example, WhatsApp has 120 million users in Brazil, a country with a population of 200 million. Thirty-five percent of these users regularly rely on the messaging platform for their news consumption, which makes WhatsApp networks “fertile for planting false information that can spread quickly from group to group until it is out of control.”106 These countries are therefore more susceptible to efforts to promote divisive and anti-liberal narratives, whether by domestic opposition or malign external forces, via online platforms. Moreover, polarization is a significant characteristic of Latin American politics, and the use of fake news communicated via social media platforms has proved to be more effective within polarized societies.107 This environment is primed for the rise of populist candidates who are inclined to promote an illiberal regime and can further foster radicalization and change toward a national-populist, illiberal, and even autocratic regime. This is especially true given that the checks and balances in states with only a short democratic history are less effective than in long-established liberal states.108

Populism is an ideology that views society as divided into two homogenous and antagonistic groups: the pure people and the corrupt elite — “us” versus “them.”109 Populists portray themselves as anti-elitist, anti-pluralist, supporters of moralism, and the exclusive legitimate representatives of the people in defiance of the unresponsive political elites. The rhetoric used by populist leaders generally focuses on the perception of a state in a crisis that needs to be resolved. Populists often use a dramatized and discursive repertoire that creates tension between antagonistic blocks.110 Pippa Norris notes that populism undercuts the legitimacy of the checks and balances on executive power that were “protecting citizens from strong leaders advocating authoritarian values attacking the heart of liberal freedoms, social tolerance, and cosmopolitanism.”111

Populism exists across the political spectrum. According to Paolo Gerbaudo, the populist right “tends to take highly exclusionary and xenophobic forms, whereby the people are constructed in opposition to the Other, and in particular migrants and ethnic and religious minorities.” Left-wing populism, on the contrary, opposes “immoral privilege, as embodied by greedy bankers, rogue entrepreneurs, and corrupt politicians accused of exploiting the people.”112 Postill also discusses centrist populists — opportunistic technocrats who borrow populist rhetoric and blend it with a pro-market language of job flexibility, entrepreneurship, and economic growth.113

Gerbaudo, who studies the relationship between populism and social media, explains that social media provides a platform for populists to gain people’s support against a liberal establishment that is supposedly victimizing them.114 Populists are able to unify otherwise dispersed and divided people to promote a shared cause, exploiting the platforms’ “economy of attention” and filter bubble effect. In this way, populists develop online followers of like-minded individuals and “siloed communities that experience their own reality and operate with their own facts.”115

In Brazil’s 2018 elections, Bolsonaro, a far-right candidate, was elected president with 55 percent of the vote, putting an end to the social-democratic pact that had been established after the end of the military dictatorship in the 1980s. In these elections, the far-right opposition movement “Brazil over Everything, God above Everyone” overtly used the spread of misinformation and fake news through social media to advance its discourse. This included attacks against the Workers’ Party, the group’s main competitor, associating them with child abuse, female nudity, and more. Bolsonaro’s campaign also used social media platforms to attack feminists and minority groups, including the LGBT community, blacks, and indigenous people.116

Up until the 2018 election, political television advertising was the primary means of reaching out to Brazil’s electorate. Bolsonaro’s low-budget campaign, however, relied heavily on political micro-targeting via social media to directly engage with his electoral base.117 His early supporters distrusted mainstream media and assumed that social media is more genuine “because it’s filled with friends and family.”118 The campaign focused on professionalizing a fake news industry by using WhatsApp, YouTube, Twitter, and Facebook.119 Bolsonaro’s entire campaign was built upon “exploiting a political behavior tied into a sense of fear — fear of being shot, of crime, of unemployment — that ends up creating space for the acceptance of authoritarian feelings latent in society.”120

Angered by violence, scandals, and a deep recession, voters were ready for Bolsonaro’s messages about crime, corruption, and family values. He energized voters who disliked the ruling party and who detested all the other candidates.121 He benefited from the high levels of cognitive dissonance that some voters were experiencing — between their image of the country and the world as it is.

Over the years, Bolsonaro has repeatedly called for Congress to be closed and has said that he would “start a dictatorship right away if elected president.”122 During the 2018 election, he continued his attacks on the idea of liberal democracy and the legitimacy of the media as well as on other political opponents.123 His final speech before election day was a direct attack on several democratic norms, reiterating the central themes of his campaign: diluting the power of minorities, closing down non-government organizations, and promising to imprison his opponent in the race, Fernando Haddad.124 Democratic principles further erode when candidates such as Bolsonaro, who use social media manipulation as part of their campaign strategy, continue with these tactics after assuming power and becoming part of the governing regime.125 For example, since assuming power, Bolsonaro has used different means, including fake news, to discredit Brazil’s electoral processes (such as when he questioned the integrity of the 2020 municipal elections), lash out at the Brazilian Supreme Court and Supreme Electoral Tribunal, and fight federal police on investigating him and his allies.126

Support for democracy in Brazil dropped from a peak level of 55 percent in 2009 to 34 percent in 2018.127 By 2020, only 15 percent of Brazilians said they were satisfied with democracy, a drop of 35 points compared to 2014.128 Bolsonaro’s election thus marks the intensification of a process of decay that has affected Brazil’s democratic system for some time. Brian Winter suggests that Bolsonaro, when faced with resistance as a president, “will ignore or trample democratic practices and norms to get his way.”129 Since the checks and balances in Brazil, as in many other Latin countries, are weak and insufficient, such action by the president could further weaken democratic institutions and lead to regime change.

In addition to the domestic attack on democratic institutions, Brazil is also experiencing external intervention as part of Russia’s propaganda operations in Latin America aimed at promoting divisive narratives through online platforms.130 According to Brian Fonseca, Russia’s objective is “to erode confidence in Western institutions such as democracy and free trade, as well as Western-dominated sources of information.”131 Moscow has been using social media platforms to exaggerate, distort, and fabricate falsehoods regarding U.S. and Western activities in the region.132 As a liberal-democratic regime, Brazil has only a limited set of tools to defend itself from the malign use of social media both from outside and within the country.

Brazil is not the only country in which democracy has been deteriorating. In Mexico, internal support for democracy dropped from 48 percent in 2015 to 38 percent in 2018,133 and a far-left populist, anti-establishment candidate, Andrés Manuel López Obrador, was elected in 2018 in a campaign that made heavy use of social media. Since his election, democratic institutions in Mexico have been under attack. For example, Obrador has stated that the National Electoral Institute and the Electoral Tribunal of the Federal Judiciary “were created to prevent democracy.”134 Another example is the actions taken by Obrador to weaken the autonomy of the judiciary and to replace some judges with Obrador’s close allies. Mexico has also been the target of Russian propaganda operations.135 Other Latin American countries with characteristics similar to those of Brazil and Mexico are at risk of following the same path of eroding democratic norms.136

In summary, social media platforms have a radicalizing effect in weak liberal democracies, facilitating the rise of populist candidates who erode the country’s democratic norms and institutions and may lead to regime change. Social media is a playground for spreading fake news and narratives that are polarizing, divisive and anti-liberal — without the fact-checking filter of the traditional media. It helps populists to aggregate and unify people to promote a shared cause against the liberal establishment or the corrupt elites, positioning themselves as worthy alternatives to the existing governments. Malign external forces also try to use social media platforms to intervene in these countries’ democratic elections.137

The Intensifying Effect in Strong Authoritarian Regimes

Social media platforms can intensify the power of strong authoritarian regimes by helping them, directly or indirectly, to become “digital dictatorships.”138 These regimes use the knowledge power of social media platforms as part of their surveillance machine. They can monitor and block social media platforms to hinder the ability of dissidents and domestic opposition groups to organize and mobilize. Authoritarian regimes also use social media platforms as tools to apply sharp power against liberal-democratic countries worldwide and as a way to spread fake news in democratic elections.

Between 2000 and 2017, 60 percent of all dictatorships faced at least one anti-government protest of 50 participants or more. Ten authoritarian regimes fell during this period and 19 were replaced through elections, many of which came in the wake of mass protests.139 According to Democracy Report 2020, pro-democracy protests reached an all-time high in 2019 as people took to the streets to protest the erosion of democracies and to challenge dictators.140 The leaderless nature of 2019 Hong Kong protests against China, for example, was made possible by social media. Protesters took their cues from more than 100 groups on the instant messaging app Telegram, dozens of Instagram pages, and online forums like LIHKG. These groups were used to post everything from news on upcoming protests and tips on defending oneself from tear gas canisters fired by the police to the identities of suspected undercover police and the access codes to buildings in Hong Kong where protesters could hide.141 Overseas Chinese dissidents and activists played a crucial role by assisting and even guiding activists in Hong Kong. Chinese expatriates connected with those in Hong Kong via social media to get information about what was going on to journalists, non-governmental organizations, and activists in other countries.142

In the last 20 years, the more durable authoritarian regimes have been those that have implemented digital repression.143 In order to avoid regime change, strong authoritarian regimes have used their economic strength and coercive power to embrace technology and become “digital autocracies.” That is, they restrict their citizens’ use of the internet and social media while harnessing a new arsenal of digital tools to deal with mass anti-government protests.

China has long maintained strict regulations that determine which websites and social media platforms are accessible in the country and which are blocked behind its “Great Firewall” of internet censorship, which is part of the country’s “cyber sovereignty” model.144 In 2003, Debora Spar claimed that “if people in China want to get information from sites in Silicon Valley, even the most omnipotent of governments will be hard-pressed to stop them.”145 But recent years have proven her wrong. China blocked YouTube in March 2008, the same month that a significant wave of protests-turned-riots swept Tibet. It blocked Facebook and Twitter the next year, soon after an outbreak of ethnic unrest rocked Xinjiang in July 2009.146

China employs advanced technology to censor its citizens on social media (and access their private information). This technology, combined with laws, regulations, and ramped up enforcement, is increasingly being used to repress dissidents and domestic opposition voices and shape the online conversation.”147 Many of the state’s censorship tactics operate with a “light touch,” so that Chinese internet users do not necessarily detect the filtering and deletion of material that is going on behind the scenes. There are seven topics that social media content shouldn’t contravene according to the Chinese government: “China’s rules and laws, the socialist system, the country’s national interests, the legitimate interests of citizens, public order, morality, and authentic information.”148 Chinese social media platforms such as WeChat and Sina Weibo have no choice but to actively participate in the monitoring and censorship of their users in order to stay in business.149

U.S. social media platforms, which could potentially act as a liberalizing external force, cannot operate in China without becoming active partners in the government’s efforts to silence dissent through censorship, mass surveillance, and the use of criminal charges.150 In December 2017, an official from China’s Cyberspace Administration stated: “If they [foreign social media] want to come back, we welcome [them]. The condition is that they have to abide by Chinese law and regulations and that they also would not do any harm to Chinese national security and national consumers’ interests.”151 Collaboration with the Chinese government contradicts the liberal agenda of most of these corporations, which see themselves as champions of free expression. However, some of them, including Google, are directly and indirectly helping China to enhance its internet surveillance capabilities and censorship technology.152

Russia, another strong authoritarian regime, lives in constant fear of U.S. efforts to interfere with the Russian regime. After witnessing the role that social media played in the Arab Spring, Russia became increasingly concerned that America had “found a truly magic tool that could bring people to the streets without any organizing structure.”153 This fear was amplified by several statements made by Alec Ross, then-Secretary of State Hillary Clinton’s senior adviser for innovation at the U.S. State Department, including his comment in 2011 that “the Che Guevara of the twenty-first century is the network” and “dictatorships are now more vulnerable than they have ever been before … because of the devolution of power from the nation-state to the individual.”154 This fear became a reality when protests erupted over irregularities in the 2011 Russian legislative elections — protests that were facilitated by Facebook and Twitter.155

Sergei Smirnov, director of the FSB (Russia’s security agency), stated in 2012: “New technologies are being used by Western special services to create and maintain a level of continual tension in society with serious intentions extending even to regime change.” He emphasized that Russia needed to develop ways to respond to such technologies.156 In June 2012, legislation was introduced in the Duma, the lower house of parliament, to impose a nationwide filtering system on the internet. The legislation was approved a month later. In 2013, a system for social media monitoring — Mediaimpuls — was introduced. Russian law allows the authorities to block online content, including social media websites whose activities are deemed “undesirable” or “extremist,” and to prevent users of social media and communications platforms from remaining anonymous.157 Under its 2019 Sovereign Internet Law, Russia is centralizing internet traffic in the country and creating chokepoints (similar to China’s Great Firewall). The Federal Service for Supervision of Communications, Information Technology and Mass Media (or Roskomnadzor) is exercising its authority inside Russia and outside its borders to silence protesters and anti-Russian voices.158

China and Russia have started to proliferate their models of digital authoritarianism across the globe. China is exporting its digital tools for domestic censorship and surveillance to different countries such as Malaysia, Singapore, Ethiopia, Zimbabwe, and Venezuela. Russia is disseminating its tightened information control model coupled with intimidation of internet service providers, telecom providers, private companies, and civil society groups.159 Russia’s model may be an appealing, relatively low-tech, and inexpensive alternative to the Chinese model because it does not require information filtration capabilities and can be implemented without a pre-existing government firewall.160

In addition, some of these tools that were initially developed for domestic use are now being used as part of the “sharp power” campaign that Russia is waging against liberal-democratic regimes. This includes using automated accounts (“bots”) on social media to manipulate and “amplify influence campaigns and produce a flurry of distracting or misleading posts,” sowing confusion and uncertainty through the dissemination of alternative narratives.161 Another tool, the use of internet trolls, involves paying people to disrupt online discussions by deliberately posting inflammatory or off-topic messages over social media platforms in order to provoke and intimidate. Russia conducted a massive troll attack against Ukraine and other countries after annexing Crimea.162 Both trolls and bots have been used by Russia through social media platforms in democratic elections across the world in the last five years.163 They helped to elect populist nominees or promote their agendas, deepening domestic polarization, ethnic tensions, and anti-migrant and anti-minority sentiments while eroding democratic institutions.

To summarize, American social media platforms may intensify the power of strong authoritarian regimes by helping them, directly and indirectly, to become digital dictatorships. They use the knowledge power of compliant platforms as part of their surveillance machine while blocking those platforms that refuse to play by their rules.

China and Russia export their restrictive practices to other authoritarian states, helping them to adopt similar practices in their countries. Russia also uses social media platforms as tools to apply sharp power against liberal-democratic countries around the world.

The Weakening Effect in Strong Democratic Regimes

The spread of fake news, disinformation, misleading information, and falsehoods through social media platforms as part of malign “perception management”’ orchestrated by domestic populists and external forces (such as Russia) may weaken strong liberal-democratic regimes. Such social media campaigns amplify extreme views, polarization, conspiracy theories, and doubts about democratic institutions and processes and weaken people’s trust and confidence in these institutions and processes. In strong democracies, the erosion of democratic pillars is less dramatic than in weak democracies because the checks and balances of strong democratic regimes remain a solid protection against domestic populist opponents and malign external actors.

The use of social media for political campaigning is not new to liberal-democratic regimes. American President Barack Obama used big data and individual marketing to drive people to the voting booths in both the 2008 and 2012 presidential elections.164 What is new is the malign use of these platforms and their mobilization by external forces and populist domestic opponents to change people’s perceptions. The dependence of democracies on free and open political discourse provides opportunities for external forces to infiltrate their information ecosystems.165 Researchers identify the 2016 U.S. presidential elections as a watershed moment in terms of the impact of fake news on social media platforms on presidential elections.166

A U.S. national intelligence report claims that Russia’s Internet Research Agency, an army of social media trolls created in 2014, was part of Russia’s efforts to interfere in the 2016 U.S. elections. This interference included propaganda campaigns in the media and a troll campaign on social media aimed at undermining public faith in the American democratic process.167 The Internet Research Agency spent more than $100,000 on Facebook political ads between June 2015 and May 2017, using 470 fake accounts.168 Facebook reported to the U.S. Senate that Russian trolls created events on Facebook that were seen by more than 300,000 users between 2015 and 2017 and that around 62,500 people planned to attend these events. Russian accounts used Facebook to promote pro-Trump rallies, such as “Florida Goes Trump” in August 2016, as well as events in May 2016 protesting the opening of an Islamic Center library.169 Facebook acknowledged that 146 million users might have viewed Russian misinformation on its platform during the election campaign, while YouTube identified 1,108 Russian-linked videos and Twitter acknowledged 36,746 Russian-linked accounts.170

Similarly, researchers discovered massive Russian meddling on Twitter in the lead up to the Brexit referendum in the United Kingdom in 2016.171 More than 150,000 Russian-language Twitter accounts posted tens of thousands of messages in English urging British people to vote to leave the European Union in the days prior to the referendum. Most of the messages sought to inflame fears about Muslims and immigrants and to intensify the polarization of the electorate.172 British Prime Minister Theresa May even publicly accused Moscow of seeking to “weaponise information” and “sow discord in the West and undermine our institutions.”173 She added that Russia’s cyber activities included “deploying its state-run media organizations to plant fake stories and photo-shopped images.”174 This phenomenon seems to be spreading and intensifying. In 2017, for example, one year after the U.S. presidential elections, at least 18 other national elections were targeted by social media manipulation and disinformation tactics.175

According to the U.S. Justice Department, the Internet Research Agency used Facebook’s own tools to ensure that its propaganda was as effective as possible. These tools allowed the agency to receive real-time feedback about which ad campaigns were reaching their target audience and which posts were generating the most engagement with viewers.176 These “active measures” of media manipulation and disinformation, using social media campaigns, fake news, and troll armies, are designed to exploit political division and subvert the democratic process in the United States and Europe, “destabilizing the society and the state.”177

Russian “perception management” during liberal-democratic elections is based on the art of disinformation, or “using false or misleading information and injecting it or getting it credited by legitimate and credible sources.”178 Russia typically manipulates information using social media platforms to sow confusion and disruption. The aim is to create the impression that truth does not exist, thus undermining trust and authority in democracies. Russian manipulators on social media amplify extreme views, conspiracy theories, and doubts about democratic institutions.179 Russian intervention has found a receptive audience of people who believe that all truths are partial and that there are many legitimate ways to understand or represent an event.180 Using disinformation and fake news in the public sphere may diminish the role of facts in public life and lead to what Jennifer Kavanagh and Michael Rich call “truth decay.”181 Truth is a cornerstone of democracies and what distinguishes them from autocracies. The decaying of truth is dangerous for American democracy.182

But not everyone agrees that Russian intervention actually affected the 2016 election process. Hunt Allcott and Matthew Gentzkow show in their research that exposure to fake news was insufficient to make a difference in the 2016 U.S. presidential election and that the effect of fake news was smaller than Trump’s margin of victory in key states.183 Another group of researchers also claim that the advertisements that Russia reportedly bought on social media were not targeted effectively on battlegrounds states and that the money it spent was dwarfed by the money spent by Trump and Clinton. Although Russia bought thousands of ads, they constituted only a fraction of the overall posts and tweets circulated on social media in the months leading up to the election. Moreover, even if people engaged with Russian-sponsored content, there is still the question of whether and how it affected their voting behavior.184 Even if Russian influence was not the main reason for Trump’s victory in the 2016 election, the spread of fake news via social media platforms deepened liberal societies’ distrust of political institutions and distrust of the media in particular.

It can be hard to distinguish fake news and misinformation originating from external forces, such as Russia, from information coming from anti-establishment populist candidates in strong liberal democracies, such as Trump and Sen. Bernie Sanders in the United States and Boris Johnson and Jeremy Corbyn in the United Kingdom.185 These populists promote divisive narratives of “them” versus “us,” the “pure” against the “corrupt,” “the masses” against “the elite,” “the people” versus “politicians/parliament/judges.”186 These narratives deepen the polarization that already exists in these countries. The Fragile States Index shows that the Group Grievance and Factionalized Elites indicators in America and the United Kingdom doubled in the last 12 years.187 This polarization may lead to brinksmanship or gridlock, making the country less functional. According to Nate Haken, in this situation of fragmentation “the usual brokers (media, state institutions, opinion leaders, religious and community leaders) lose relevance and legitimacy, making consensus-building difficult with no shared vision or context to build from and organize around.”188

When populists mix divisive speech with fake news and disinformation, they further erode people’s trust in democratic institutions, processes, and the media. However, strong liberal democracies are still unlikely to undergo the same democratic decay experienced by some weak liberal regimes. While they are constrained in the means they can employ to counter the malign use of social media, their checks and balances are more stable and robust than in weak liberal democracies. Their greater resilience is grounded in a long-standing democratic tradition, and their checks and balances are less affected by recent actions, especially when the support for democracy among citizens is still high.189

Social media platforms can be used to weaken strong liberal-democratic regimes. These regimes derive their power from liberal-democratic institutions, which need constant attention and reinforcement in order to serve as effective bulwarks of democracy. It is also important to keep the media free, unbiased, and devoid of fake news and disinformation. The spread of fake news and disinformation on social media as part of malign “perception management” orchestrated by domestic populists and external forces may weaken liberal-democratic regimes. Liberal democracies are restricted in the means they can employ to counter the malign use of social media. For now, the checks and balances of liberal-democratic regimes remain a solid protection against domestic populist opponents and malign external forces.

Counterfactual Reasoning

The causal model presented in the article proposes that social media platforms are used by three types of actors — domestic opposition, external forces, and the governing regime — for political purposes, and that their overall influence on the political system is determined both by the state’s capacity and the type of regime. One may ask whether these same outcomes — destabilizing, radicalizing, intensifying, and weakening — would still take place without social media. The problem is that exposure to social media is so extensive that it is almost impossible to find examples where an uprising took place without social media. Nevertheless, we will mention some counterfactual reasoning related to weak authoritarian and weak democratic countries found in our research.

According to one study, during the Arab spring in 2011, internet penetration was higher in the countries that were experiencing unrest than in those that weren’t.190 Internet users made up just 1.1 percent of Iraqis and 3.4 percent of Afghans in 2010, for example, compared to over 21 percent of the population in Egypt, 34 percent in Tunisia, and 88 percent in Bahrain. In the first three months of 2011, the number of Facebook users in the Arab world increased by 30 percent compared to an 18 percent growth over the same period in 2010.191 Countries where major civil movements have occurred have shown exponential growth in social media use during and after those movements. Still, it is hard to say that social media’s absence decreased the probability of online mobilization campaigns against the governing regime taking form. Some of the unrest was also fueled by traditional media outlets, as discussed earlier.192 In addition, some protest movements were constrained by the ability of governments to block internet access, as was the case in Iran.

When it comes to Brazil, the percentage of people using the internet grew from 40 percent in 2009 to 70 percent in 2018, while support for democracy dropped from a peak level of 55 percent to 34 percent during the same time period.193 Only 15 percent of Brazilians said they were satisfied with democracy in 2018 (a drop of 35 points compared to 2014).194 A survey conducted in August 2018 by the Brazilian Institute of Public Opinion and Statistics showed that only 25 percent of citizens trusted the federal government and only 18 percent trusted Congress.195 These numbers may indicate that Brazil had been radicalizing and drifting away from democracy for several years in relation to the growth in use of social media. But several researchers question the assertion that social media platforms are solely responsible for the deterioration of Brazil’s democracy, pointing to the many years of government corruption, rising crime, and economic recession.196

The above examples show that social media platforms did affect the political system in several countries, but that there may have been other causes as well. Nevertheless, as Patrícia Campos Mello claims, “the use of WhatsApp and other internet platforms amplifies whatever a political group says in an exponential way … If you find the right conditions in a country, they are really dangerous tools to undermine democracy and manipulate public debate.”197

The Way Ahead

This article looked at four case studies to illustrate the varying effects of social media platforms depending on who is using them — domestic opposition, external actors, or the governing regime — the regime type of the country, and the state’s capacity. In weak authoritarian states, social media can help dissidents to communicate and organize more easily (the destabilizing effect), while strong authoritarian states can use it as a suppressive tool to exploit the knowledge aggregated on the different social media platforms (the intensifying effect). When used to disseminate distorted information and fake news in strong liberal democracies, social media platforms can erode democratic institutions (the weakening effect). These platforms can facilitate populist leaders’ rise in weak liberal states (the radicalizing effect), making them more susceptible to turning into an illiberal or even authoritarian regime.

The malign use of various inherent characteristics of social media platforms — such as filter bubbles, echo chambers, a low entry bar, aggregate knowledge about people, the lack of fact-checking, information cascades, and the automatic recommendation algorithm — may lead to the erosion of democratic principles and institutions in liberal democracies across the world.

Although the Russian intervention during the 2016 U.S. elections caught the attention of American policymakers and the American public, what has been less discussed is social media’s impact on other liberal democracies and what that might mean for U.S. national security and the liberal international order. The malign use of social media platforms is only one reason for the disruption of that order. Other factors include the 2008 financial crisis, job losses related to changes in trade and technology, and the increased flow of migrants and refugees, among others.198 But when it comes to the abuse of social media platforms, America should not wait for social media companies to fix the problem themselves.199

In order to prevent liberal democracies from becoming illiberal or autocratic regimes and potentially drifting into the Russian or Chinese spheres of influence, it is crucial that the United States and other democracies take action today. Maintaining the current liberal international order requires keeping the internet an American project led by private companies. This means countering Russia’s and China’s efforts to gain a greater voice in internet governance and to promote their agenda of cyber sovereignty, under which government control and internet regulations would replace a global and open internet.200

U.S. policymakers cannot rely solely on social media companies to implement policies and technological means to decrease the flow of hate speech and fake news on their platforms.201 There are a number of possible approaches to this problem that have been suggested. One possible solution that the U.S. government can pursue is to reintroduce competition into this sector of the market by passing antimonopoly legislation in order to dilute the concentrated power of social media platforms.202 This approach views social media platforms as essential infrastructure (like public utilities) needing specific regulation.203 These regulatory tools would ensure that the infrastructure “serves the public’s needs — rather than incentivizing exploitative or exclusionary uses for private profit.”204 Relatedly, policymakers could consider breaking up or decentralizing these corporations.205 Other options are more concentrated on users and the data that these companies own. One possibility is to make the companies declare “platform bankruptcy,” whereby social media platforms would reset their entire user and group follower counts to zero and rebuild communities from the ground up, with the platforms’ current rules in place.206 Another approach is to make these corporations collect less data and adopt practices to treat that data in a “manner commensurate with its value.”207 The problem is that none of these solutions is a “silver bullet” and American legislators may risk trampling on the constitutional right to free speech that they are trying to preserve.208

#### The erosion of US democratic norms causes imminent global backsliding.

Larry Diamond 21, Senior Fellow at the Hoover Institution and the Freeman Spogli Institute for International Studies at Stanford University, "A World Without American Democracy?" Foreign Affairs, 07/02/2021, https://www.foreignaffairs.com/articles/americas/2021-07-02/world-without-american-democracy.

A prolonged global democratic recession has, in recent years, morphed into something even more troubling: the “third reverse wave” of democratic breakdowns that the political scientist Samuel Huntington warned could follow the remarkable burst of “third wave” democratic progress in the 1980s and the 1990s. Every year for the past 15 years, according to Freedom House, significantly more countries have seen declines in political rights and civil liberties than have seen gains. But since 2015, that already ominous trend has turned sharply worse: 2015–19 was the first five-year period since the beginning of the third wave in 1974 when more countries abandoned democracy—twelve—than transitioned to it—seven.

And the trend continues. Illiberal populist leaders are degrading democracy in countries including Brazil, India, Mexico, and Poland, and creeping authoritarianism has already moved Hungary, the Philippines, Turkey, and Venezuela out of the category of democracies altogether. In Georgia, the dominance of the Georgian Dream Party has led to the steady decline of electoral processes and a breakdown in the rule of law. In Myanmar, the military overthrew the elected government of Aung San Suu Kyi, ending an experiment in partial democracy. In El Salvador, president Nayib Bukele staged an executive coup by removing the attorney general and Supreme Court justices who were obstacles to his consolidation of power. In Peru, democracy hangs from a thread as the right-wing autocrat Keiko Fujimori advances vague claims of election fraud in a bid to overturn her narrow electoral defeat to left-wing opponent Pedro Castillo.

What is especially striking about this last case is that Fujimori’s gambit bears a grim resemblance to the lie perpetuated by former U.S. President Donald Trump and his followers about the 2020 presidential election. This is no coincidence. As the journalist and historian Anne Applebaum has observed, fictitious claims of fraud and “stop the steal” tactics are becoming a common means by which autocratic populists try to obstruct democracy. Such tactics have long been a source of instability in countries struggling to develop democracy. But the fact that the most recent iteration of the antidemocrat’s playbook draws heavily on precedents in the world’s most important and powerful democracy marks the start of a dangerous new era.

Today, the United States confronts a growing antidemocratic movement, not just from the ranks of fringe extremists but also from a substantial group of officeholders—a movement that is challenging the very foundations of electoral democracy. Should this effort succeed, the United States could become the first ever advanced industrial democracy to fail—that is, to no longer meet the minimum conditions for free and fair elections as political scientists and other scholars of democracy define them.

The failure of American democracy would be catastrophic not only for the United States; it would also have profound global consequences at a time when freedom and democracy are already under siege. As Huntington noted, the diffusion of democratic movements and ideas from one country to another has helped drive positive democratic change. Antidemocratic norms and practices can spread in a similar fashion—especially when they emanate from powerful countries. That is why the acceleration of a democratic recession into a democratic depression happened largely on Trump’s watch. And it is why no development would more gravely damage the global democratic cause than the democratic backsliding of its most important champion.

THE DEMOCRATIC TRIAD

A democratic system of government stands on three legs. The first leg is popular sovereignty—rule by the people. Democracy demands that people are able to choose and replace their leaders in regular, free, and fair elections; that all adult citizens are able to vote free from intimidation and obstruction; and that candidates and parties are free to compete and campaign. Crucially, elections must be administered impartially, so that valid ballots are counted accurately and power is granted to those who win.

Liberty is the second leg of liberal democracy. A fully democratic system provides strong protections for freedom of speech, the press, association, and assembly. It ensures that these rights are equally protected for all social groups. And it promotes a culture of mutual tolerance and respect for the rights of political opponents.

The third leg—the rule of law—defends and strengthens the other two. It ensures that democratic procedures are impartially enforced by an independent judiciary and other regulatory bodies that check the abuse of power. In most advanced democracies, excluding the United States, these instruments of accountability include national bodies to administer elections and to monitor corruption.

Trump was the first U.S. president to demonstrate contempt for all three legs of the triad of liberal democracy. He attacked the media as “fake news” and “absolute scum” and called for his election opponent to be “locked up.” He invited his followers to commit acts of violence against protesting opponents. Upon his defeat, he insisted that the election results were fraudulent and had to be overturned. Throughout his presidency, he waged war on an independent judiciary, the Federal Bureau of Investigation, his own attorney general, the Office of Government Ethics, the civil service, and a host of other actors who refused to bend to his political will or sought to enforce the rule of law.

Many scholars of democracy perceived an unprecedented threat to U.S. democracy when Trump entered office in 2017 and feared grave assaults on the second and third legs of the democratic triad, in particular. This assessment was partially correct. Not since President Richard Nixon and rarely in U.S. history has there been such a determined effort to misuse and subvert administrative and rule-of-law institutions for nakedly political ends—but these attempts achieved only limited effect. The bulk of the press and the judiciary remained independent. The FBI avoided political capture. Outside the Republican Party and Trump’s own administration, freedom of speech thrived. From 2017 through 2020, liberty and the rule of law more or less held.

In three respects, however, most scholars misjudged the nature of the peril—and underestimated its gravity. First, many assumed that Trump himself constituted the biggest threat to U.S. democracy and that his defeat would lance the poisonous boil on the body politic. Second, with notable exceptions, including the Yale historian Timothy Snyder and the Carnegie Endowment scholar Rachel Kleinfeld, many underestimated the potential for violence on the part of Trump’s true-believing followers. And third, most underestimated the extent to which Trump would remake the Republican Party as an institution not only slavishly loyal to him but also hostile to democracy.

Fortunately, leading up to the 2020 election, democracy scholars and civic organizations correctly anticipated the threats to electoral integrity posed by zealous Trump partisans, as well as the staggering logistical challenges presented by the pandemic. As a result, they launched one of the most energetic civic campaigns in U.S. history to register an unprecedented number of voters, to give them safe and early access to the ballot, to ensure that local electoral administrations had the resources necessary to administer the vote, and to prepare to combat any potential efforts to overturn the legitimate results of the presidential election. The election was not a nightmare scenario, as some had feared. In fact, it proved to be one of the best-administered elections in U.S. history, leading election experts Nathaniel Persily and Charles Stewart III to call it a “miracle.”

IT COULD STILL HAPPEN HERE

Yet what followed was, in the words of Persily and Stewart, a “tragedy,” with “lies about vote fraud and the performance of the system [cementing] a perception among tens of millions of Americans that the election was ‘rigged.’” Such “manufactured distrust” has extended past the January 6 insurrection in Washington. Although President Joe Biden’s inauguration has deescalated imminent threats to civil liberties and the rule of law, the core element of electoral democracy—free and fair elections—is now under relentless partisan assault. Republican state legislatures are accelerating efforts to make it more difficult for African Americans, Latinos, and other Democratic-leaning constituencies to vote by passing laws that make it more difficult to vote by mail and to vote early, and that make it easier to purge voters from voting rolls. These changes are driven not by documented evidence of malfeasance associated with these practices but by deliberately false narratives about election fraud.

Now, the greatest threat to American democracy is posed by legislative initiatives seeking to subvert the independence of electoral administration, including the counting and certification of the vote. As the election law expert Richard Hasen has observed, “At stake is something I never expected to worry about in the United States: the integrity of the vote count.” A recent law passed in Georgia, for example, removes the secretary of state (currently Brad Raffensperger, who refused to manufacture the 11,780 votes Trump needed to win the state) as chair of the state Election Board and gives the state legislature—a highly partisan institution—the ability to name the new chair. Representatives in Michigan have politicized the Board of State Canvassers, which certifies election results, by replacing a Republican who voted to certify Biden’s election victory with a movement conservative. In Michigan and in Nevada, Trump loyalists are seeking to consolidate control over election supervision by running candidates for secretary of state—giving them the authority to preside over election administration and the tools to try to block Democratic votes. And at the federal level, Republicans could take back control of the House of Representatives (helped by their unilateral ability to redraw 187 congressional districts following the most recent census) and use their majority to manipulate the 2024 presidential results in their favor—especially if the 2024 election resembles 2020, when Democrats won a decisive popular vote victory but relied on narrow margins in a handful of states for an Electoral College majority.

Once a political system loses bipartisan consensus respecting the rules of the democratic game, it can be a short slide to autocracy. The world has watched this happen in Hungary, Turkey, and Venezuela. It is not inconceivable that it could happen in the United States.

THE GATHERING STORM

To warn of the failure of American democracy is not hyperbole or simply a slogan meant to motivate action. Political scientists may differ on the minimum conditions for democracy, but they agree on this: a country cannot be considered a democracy if it does not broadly ensure the neutral and fair administration of elections. If the outcome of a major national election in the United States were to be determined by fraudulent exclusion or the manipulation of votes, the country would cease to be a democracy, no matter how much freedom of expression might survive (for a time).

More than 100 prominent democracy scholars recently warned in a collective statement that Republican assaults on electoral integrity could bring about the demise of U.S. democracy. They appealed to Congress to pass the John Lewis Voting Rights Act and to adopt other measures to “ensure the sanctity and independence of election administration.” But with broad national legislation to ban partisan gerrymandering and strengthen voting standards unlikely in the near term, it will also be up to civil society to defend American democracy.

That defense is made more urgent by the gathering storm of democratic backsliding around the world. The United States’ outsize importance as a source of political diffusion, for good or ill, makes it an example that will influence struggling democracies and embattled autocracies alike. Both in backsliding democracies such as the Philippines and Poland and in deepening autocracies such as Turkey and Venezuela, Trump’s mantra of “fake news” emboldened strongman leaders in their assaults on the media. If the United States winds up disfiguring its democracy by politicizing electoral administration and suppressing minority votes, autocrats will gleefully seize upon the American precedent as justification for their methods of blocking democratic change. And in declining democracies, politically vulnerable incumbents will embrace similar methods of violating electoral integrity in order to hang on to power.

In short, what happens to democracy in the United States is likely to determine the fate of democracy around the world: whether this third wave of democratic reversals is turned back or gains horrific new momentum.

#### Global de-democratization prevents effective regulation of emerging tech.

Daron Acemoglu 21, Professor, Economics, MIT, "The Right Way to Worry," Project Syndicate, 05/14/2021, https://www.project-syndicate.org/onpoint/how-to-think-about-existential-and-immediate-risks-by-daron-acemoglu-2021-05.

But there are also anthropogenic – human-created – existential risks. As the University of Oxford philosopher Toby Ord argues in his thought-provoking new book, The Precipice: Existential Risk and the Future of Humanity, it is these risks that should most concern us now and in the coming century.

RISK AND REWARD

Ord recognizes that science and technology are humankind’s most potent tools for solving problems and achieving prosperity. But he reminds us that there are always dangers associated with such power, particularly when it is placed in the wrong hands or wielded without concern for long-term and unintended consequences.

More to the point, Ord argues that anthropogenic existential risk has reached an alarmingly high level, because we have developed tools capable of destroying humanity without the accompanying wisdom needed to recognize the danger we are in. He notes that the eminent twentieth-century astronomer Carl Sagan issued a similar warning in his 1994 book, Pale Blue Dot, writing:

“Many of the dangers we face indeed arise from science and technology – but, more fundamentally, because we have become powerful without becoming commensurately wise. The world-altering powers that technology has delivered into our heads now require a degree of consideration and foresight that has never before been asked of us.”

For Ord, this gap between power and wisdom could decide humanity’s future. On one hand, we could disappear entirely or suffer a collapse that wipes out most of the hallmarks of civilization (from vaccines and antibiotics to art and writing). But, on the other hand, Ord sees in humankind the potential for long-term flourishing on a cosmic scale: with both wisdom and technological ingenuity, humans could well outlive this planet and launch new civilizations across space.

This far-reaching vision of flourishing weighs heavily in Ord’s reckoning, because he recognizes that there may not be any other intelligent life forms in the universe. If we are indeed alone, a mass-extinction event that wiped out everyone on this planet would also eliminate all of the potential for intelligent, purposeful existence everywhere.

Based on this reasoning, Ord arrives at what mathematicians and economists would call a “lexicographic preference ordering.” In a situation where we care about multiple criteria, a lexicographic order assigns overwhelming importance to one criterion in order to provide clarity when two options are being compared. For example, in a lexicographic order between food and shelter, one would always prefer whichever option offers more food, regardless of how much more shelter the other option offers.

Ord’s philosophical stance is equivalent to a lexicographic order because it implies that we should minimize existential risk, whatever the costs. A future in which existential risk has been minimized trumps any future in which it has not been minimized, regardless of any other considerations. After establishing this basic hierarchy, Ord then proceeds with an expert overview of different types of anthropogenic existential risk, concluding that the greatest threat comes from an artificial superintelligence that has evolved beyond our control.

WHEN PROGRESS ISN’T PROGRESS

One can date science-driven existential risk at least to the controlled nuclear chain reactions that enabled atomic weapons. Ord is probably right that our (social) wisdom has not increased since this fateful development, with its earlier culmination in the bombings of Hiroshima and Nagasaki. Though we have established some institutions, regulatory tools, norms, and other internalization mechanisms to ensure that we do not misuse science, nobody would argue that these are sufficient.

Ord suggests that today’s inadequate institutional framework may be a temporary phenomenon that could be addressed in due time, so long as we survive the next century or so. “For we stand at a crucial moment in the history of our species,” he writes. “Fueled by technological progress, our power has grown so great that for the first time in humanity’s long history, we have the capacity to destroy ourselves…” And, in fact, in writing his book, Ord “aspires to start closing the gap between our wisdom and power, allowing humanity a clear view of what is at stake, so that we will make the choices necessary to safeguard our future.”

However, I see no evidence that this is really feasible. Nor is there any sign that our society and leaders have shown any wisdom when it comes to reining in the destructive power of technology.

To be sure, one could argue in favor of Ord’s optimism on the basis of what the German sociologist Norbert Elias famously called the “civilizing process.” According to Elias, the process of economic development and the emergence of state institutions for resolving conflicts and controlling violence since the Middle Ages have led to the adoption of manners and behaviors conducive to coexistence in mass societies. Elias’s nuanced case for why people in advanced economies have become less violent and more tolerant was popularized recently by the Harvard University cognitive psychologist and linguist Steven Pinker in his bestselling book The Better Angels of Our Nature: The Decline of Violence in History and Its Causes. Both authors offer arguments for why we should continue to expect a strengthening of the norms and institutions needed to control the misuses of science and technology.

But even if such a civilizing process is acting on individual behavioral norms and social intercourse more broadly, it doesn’t seem to have affected many political leaders or scientists and technologists. The civilizing process should have been in full swing by the first half of the twentieth century; and yet the Nobel Prize-winning chemist Fritz Haber enthusiastically used his scientific knowledge to invent and then peddle chemical weapons to the German Army in World War I.

Nor was the impact of the civilizing process much in evidence in the thinking of the American leaders who ordered the attacks on Hiroshima and Nagasaki, or in the attitudes of other political leaders who eagerly embraced nuclear weapons after World War II. Some may find hope in the fact that we haven’t had a repeat of WWI or WWII over the past 75 years. But this sanguine view ignores many near misses, not least the Cuban Missile Crisis in 1962 (the episode with which Ord opens his book).

One can identify many more examples contradicting the idea that we are becoming more “civilized,” let alone better at controlling anthropogenic risks or cultivating collective wisdom. If anything, controlling our bad behavior and adapting to the constant changes wrought by scientific discovery and technological innovation will remain a constant struggle.

This raises problems for the rest of Ord’s argument. Why should trying to eliminate future existential risks be given a superordinate priority over all other efforts to ameliorate the ills and suffering that our current choices are generating now and in the near term?

For the sake of argument, suppose we could significantly reduce the probability of our own extinction by enslaving the majority of humankind for the next several centuries. Under Ord’s lexicographic ordering, we would have to choose this option, because it minimizes existential risk while still preserving humanity’s potential to flourish fully at some point in the distant future.

Not everybody will be convinced by this argument. Count me among the unpersuaded.

THE AGE OF DEMONIC MACHINES?

To clarify the choice further, consider the main existential risk that Ord focuses on: the potential misuse of artificial intelligence. Ord estimates that there is a one in ten chance that humanity will fall prey to an evil superintelligence (which he calls, euphemistically, “unaligned AI”) in the next 100 years. By contrast, his estimated existential risk to humanity from climate change is one in 1,000, and one in a million in the case of collisions with asteroids or comets.

Even if many other experts would not assign quite so high a probability to the threat of superintelligence, Ord is not alone in worrying about the long-term implications of AI research. In fact, such concerns have become commonplace among many technology luminaries, from Stuart Russell of the University of California, Berkeley, to Microsoft founder Bill Gates and Tesla founder Elon Musk.

These figures all believe that, notwithstanding the existential risks, AI will bring many net benefits. But while Ord is well enough informed about these debates to know that even this last proposition is actually rather shaky, his lexicographic stance leads him to ignore most of the non-existential risks associated with AI.

But if one accepts that our scope of attention is finite, this weighing of priorities is problematic. My own assessment is that the likelihood of superintelligence emerging anytime soon is low, and that the risk of an evil superintelligence destroying our civilization is lower still. As such, I would prefer that the public debate focus much more on the problems that AI is already creating for humanity, rather than on intriguing but improbable tail risks.2

BACK TO NOW

As I have argued here and elsewhere, the current trajectory of AI design and deployment is leading us astray, causing a wide range of immediate (albeit prosaic) problems. Far from being inevitable or reflecting some inherent logic of the technology, these problems reflect choices being made (and imposed on us) by large tech companies – and specifically by a small group of executives, scientists, and technologists within these companies (or within their orbit).

One of the most visible problems that AI is causing is incessant automation, which is displacing workers, boosting inequality, and raising the specter of future joblessness for large swaths of the labor force. Worse, the obsession with automation has come at the expense of productivity growth, because it has led executives and scientists to overlook more fruitful, human-complementing uses of innovative technology.

AI is also being designed and used in other problematic ways, none of which inspire hope for humanity’s moral progress. Democratic politics has been defiled not just by an explosion of algorithmically amplified misinformation, but also by new AI technologies that have empowered governments and companies to monitor and manipulate the behaviors of billions of people.

This development represents a double whammy. Democratic politics is the primary means by which a society can rein in misbehavior by political and economic elites, yet it is precisely this process that is being undermined. If we cannot hold elites accountable for the damage they are causing because democracy itself has been impaired, how can we possibly escape our current predicament?

#### Unregulated tech causes extinction.

Nathan A. Sears 21, PhD, Political Science, University of Toronto, "Great Powers, Polarity, and Existential Threats to Humanity: An Analysis of the Distribution of the Forces of Total Destruction in International Security," International Studies Association, Annual Conference, Conference Paper, March/April 2021, ResearchGate.

Humanity also faces existential risks from the increasing power of technology and danger of losing control over a singularly powerful technology (Danzig 2018), such as biotechnology (Doudna & Sternberg 2017), nanotechnology (Drexler 2006), or artificial intelligence (Bostrom 2014). Artificial intelligence (AI) best captures the existential risk of humanity’s loss-of-control over technology. The term artificial intelligence is used broadly to refer to digital technologies “that are capable of performing tasks commonly thought to require intelligence” (Brundage et al. 2018, 9). The current wave of progress in AI is driven by a combination of gains in hardware (e.g., the exponential growth in computing power described by “Moore’s Law”), software (e.g., “machine learning” algorithms and techniques, such as “neural networks” and “deep learning”), and data (e.g., the abundance of digital information on the Internet). Today’s AI remain “narrow” systems, which means that they can only achieve or surpass human-level intelligence in specific domains, such as AlphaGo’s 2017 victory over the world champion, Ke Jie, at the strategy game “Go.” However, the idea that computers could one day possess human-level “general” intelligence (AGI) was first suggested by Alan Turing (1950), who described a test—the “Turing Test”—in which a computer would seek to persuade a human being that it too is human. Since then, AI experts have frequently expressed the concern that AI could one day far surpass human beings in general intelligence (Good 1966; Moravec 1988; Vinge 1993; Kurzweil 2006; Bostrom 2014; Shanahan 2015; Tegmark 2018; Russell 2019).

How could artificial “superintelligence” (ASI) threaten the survival of humanity? In one scenario (“takeover”), the ASI determines humanity to be a threat to achieving its fundamental goal, and “eliminates the human species and any automatic systems humans have created that could offer intelligent opposition to the execution of the AI’s plans” (Bostrom 2014, 95–7). In another scenario (“perverse instantiation”), the ASI pursues some seemingly benign goal—e.g., maximizing paperclip production—in a way that threatens human survival—e.g., by converting the biosphere into paperclips, humans included (Bostrom 2014, 119–24). While humans should not “anthropomorphize” the goals and motivations of artificial intelligence (Bostrom 2014; Pinker 2018; Russell 2019), this may be immaterial to whether or not it constitutes an existential threat to humanity. If AI is “goal-oriented,” then an ASI could pursue instrumental objectives— e.g., survival, self-improvement, and resource acquisition—that threaten human survival in pursuit of its fundamental goal (Omohundro 2008; Bostrom 2014). If humanity achieves superintelligence but fails to solve the “control problem”—i.e., the proper “alignment” or “compatibility” of AI with humanity—then human extinction could be the “default outcome” (Bostrom 2014, 115).

#### Particularly, it results in waves of disinformation and truth decay.

Mike Hynes 21, Lecturer, Political Science and Sociology, National University of Ireland, "Digital Democracy: The Winners and Losers," in The Social, Cultural and Environmental Costs of Hyper-Connectivity: Sleeping Through the Revolution, Chapter 9, 08/17/2021, Emerald.

The Digital Promise of Democracy

From the early days of the internet, the influence, power and reach of such hyperconnectivity was acclaimed as a potential vital instrument in democratising the world. There was a somewhat naïve assumption that once people were exposed to the virtues of democracy through the medium of the information superhighway, there would be inevitable civic transformation and popular uprisings towards such a political system driven by masses of well-informed citizens of former authoritarian and communist regimes. Many of these states and regions simply needed to know about the merits of democracy to completely and unequivocally embrace this political system, and they would get this understanding directly from the internet. And the basis of such optimism was promising. The outwardly unstoppable march of freedom that began in the 1980s and culminated, by the end of that decade, in the fall of communism in the Soviet Union brought with it a sense of victory: a sense of good winning over evil, of a world of common-sense politics prevailing over the perverse and malevolence. The triumph of the West, of the Western civilisation ideals, was evidence, it was claimed, of the total exhaustion of viable systematic alternatives to Western liberalism leading to what Francis Fukuyama termed ‘The End of History’.1 In other states like China that called themselves communist, political and economic reforms were also heading in the direction of a liberal order, he claimed. But lethargy set in over the following decades with regard to approaches to international democratisation that was to see the resurgence of authoritarianism and the emergence and rise of a new form of populism that has now engulfed countries right across the world: from Brazil to Hungry, to Turkey, Egypt and the United States. All the while digital information and communication technology (ICT) was held out as a sabre of freedom that merely needed a deeper appreciation and activation and a willing population to revolutionise, reform and direct towards democracy. People would march to freedom just waving their smartphones in the air. Whether it was the 2009 Iranian ‘Twitter’ revolution, the Arab Spring or the more recent efforts of Hong Kong residents to retain a semblance of their own autonomy, all these civic uprisings and conflicts would be powered by the freedom enabled by the internet and a host of smart devices and that would ultimately lead to good triumphing over evil.

It is a very seductive notion to think that information alone sets us free and that access to the internet with its vast stores of information will lead those oppressed by authoritarians into the light of democracy. But such technological determinist thinking ignores the underlying social and economic realities that pre-exist and that are the real grounds for civil disobediences and revolution within nations in favour of simplistic cyber-utopianisms unreal expectations as to the power of digital technology and a raft of technological quick fixes. Such high exceptions of what digital technology could achieve in democratising the world have today given way to the reality that these very same digital technologies have now been weaponised by other more sinister darker forces in the world and have ironically and skillfully been turned against the very pillars of democracy itself. While many within established Western democracies dithered and failed to truly understand and embrace the real power inherent in digital ICT, others were less hesitant and seized upon the opportunity to use the technology to undermine the institutions of democracy in some of the leading countries of the West. Western liberal democracies not only have failed to truly understand the power that has been unleashed, many of them have been complicit in allowing such a situation develop in the first instance. By failing to rein in the immense influence, power and reach of big tech authorities in the West has abducted their responsibilities to protect their democracies and, in turn, have left their citizens helplessly exposed to persistent misinformation, lies, fake news and manipulation on a vast scale. Yet while countries comforted themselves with blankets of cyber-delusionism, a few lone tech writers have been sounding the alarm bells for some time now.

Watching Freedom Fail

Evgeny Morozov is a writer, researcher, and intellectual from Belarus who studies the political and social implications of digital ICT. His 2011 book, The Net Delusion, challenges the myth of ‘internet freedom’ and argues that technology has failed to democratise the world as some had previously told us would happen.2 Behind many of the eloquent words spoken in high praise of digital ICT lies a combination of utopianism and ignorance that grossly misrepresents the internet’s political role and potential. Morozov argues that the West’s irresponsible promotion of technological tools as pro-democratic agents has provoked a backlash from authoritarian regimes to crack down on online activity, not just closing down or blocking websites but using online social platforms to infiltrate protest groups and track down protesters and dissenters. They are also sowing the seeds of their own agenda and propaganda online and generally out-resourcing and out-smarting their own beleaguered people and governments of the West. Two misapprehensions about digital information technology, in particular, concern Morozov: cyber-utopianism and internet-centrism. Cyber-utopianism is the belief that the culture of the internet is inherently emancipatory and the stubborn refusal to acknowledge its limitations and downsides. It stems from:

[T]he starry-eyed digital fervor of the 1990s, when former hippies, by this time ensconced in some of the most prestigious universities in the world, went on an argumentative spree to prove that the Internet could deliver what the 1960s couldn’t: boost democratic participation, trigger a renaissance of moribund communities, strengthen associational life, and serve as a bridge from bowling alone to blogging together. And if it works in Seattle, it must work in Shanghai.3

Internet-centrism is the conviction that every important issue and concern about modern society and politics can be framed in terms of the internet. It is not a set of beliefs rather; it is a philosophy of action that informs how decisions are made and long-term strategies are developed. Internet-centrists tend to react and response to every question about democratic change by first reframing it in terms of the internet rather than the exact context in which that change is to occur.

Morozov presents a good example of the unrealistic expectations and the broadly misrepresented impacts of digital ICT in social and civil unrest. In June 2009, mostly young Iranians took to the streets of Tehran and other cities to protest what they believed to be fraudulent and rigged elections, later to become known as ‘the green revolution’ or green movement. While these protests grew in number, a counter argument surfaced among many other Iranians that the elections were, in fact, fair and they set out to defend the incumbent president of the day Mahmoud Ahmadinejad. As the two sides – representing modernity and conservatism – squared up to each other, the country faced its worst social and political crisis since the 1979 revolution which saw the return of the Ayatollah Khamenei. Meanwhile, in the West, a very simple and alluring narrative of what was occurring began to emerge and of how the internet was beginning to usher in the building blocks of freedom and a new dawn of democracy in Iran. In a series of blogs for The Atlantic, Andrew Sullivan proclaimed ‘the revolution will be twittered’ in which he claimed that the microblogging site Twitter not only managed to avoid the shutdown of communications in the country but that it was becoming a critical tool for organising the resistance in Iran.4 He offered little evidence to back up such claims. But his claims did echo with many cyber-utopianists who had patiently waited for digital tech’s big break in beginning about global democratisation, and if the evidence was not as apparent yet it was only a matter of time before it emerged. Such optimism quickly went mainstream with prominent print media organisations such as The New York Times, Wall Street Journal, Los Angeles Times, Baltimore Sun and Financial Times, as well as other non-governmental organisation (NGO) and religious publications, eulogising the power of Twitter – and by default the internet – for its ability to organise and empower ordinary citizens in the face of authoritarianism and tyranny. Even when Twitter scheduled maintenance for the website, the US state department requested the company to postpone this work so that the service would not be interrupted as, they claimed, it was being used to rally people into the streets to protest against the election. Former deputy national security advisor in the George W. Bush administration, Mark Pfeifle, even launched a public campaign to nominate Twitter for the Noble Peace Prize arguing that ‘without Twitter, the people of Iran would not have felt empowered and confident to stand up for freedom and democracy’.5

But this story was not to have its fairy tale ending. The green movement quickly lost much of its momentum in the months following the election, and realisation slowly began to dawn on those people so fervent in their belief that the internet was destined to be a liberating tool against the oppressors across the world. Young people, merely armed with their smartphone and a Twitter account, were not, in fact, leading the charge to freedom and the spread of democracy. So, what had gone so wrong? It was later discovered that estimates of fewer than 1,000 active Twitter users were actually living in Iran at the time of the election and not all had joined the demonstrations.6 Many supporters of the green movement were from outside the country – the Iranian diaspora is highly active on social media – and got carried away by the enthusiasm of the protests, and numerous Twitter users across the world switched their location setting to Tehran in an attempt to confuse Mahmoud Ahmadinejad’s security forces and shield individuals within the movement. The Iranian government itself also worked to audaciously turn the technology against the protesters, and officials from within the regime started several fake opposition accounts on Twitter which began tweeting propaganda and misleading information. Iran had not undergone a Twitter revolution, and it was argued by Reese Erlich, an author and freelance journalist who had covered the election and had extensive knowledge of the circumstances of the protests, that the term simultaneously mischaracterises and trivialises the important mass movement that had developed at that time in Iran.7 Later, similar claims were made about the influence and power of Facebook, YouTube and other social media platforms as a catalyst for change in the context of the Arab Spring, and although there’s some evidence that these were used for mobilising and organising street protests and gatherings, the true vehicles for change during that period of time were the protests themselves and the underlying grievances against the ruling authorities. Digital ICT has the potential to play a significant role in unifying and rallying people around a cause, but in the face of subsequent leaderless disorganisation, what happens then? It is also easier for authorities to combat such protest by simply spreading misinformation through the same medium used to organise in the first instance and to single out individual deemed to be arranging such protests.

Weaponising the Internet against Liberty

One of the most significant developments with regard to the organisation and ultimate collapse of the green movement and protests in Iran was the way the authorities fought back, in particular, the use of sophisticated means of disinformation by governmental officials and agencies. The regime quickly understood the real potential in the use of social media and how it can be a willing and more than capable instrument in disrupting the messages of the protest. Propaganda and government misinformation are nothing new, but social media and other forms of online broadcast media available on the internet just makes it much easier and much more effective. The real state originator of this use of online misinformation was Russia, and they have been doing this for some time now. Long before the 2016, US presidential election and the Brexit referendum, Russia, China and Iran were just some of the states who had begun to use such cyber techniques to thwart dissidents and opponents of their own regimes. As the Soviet Union began to crumble in the early 1990s, the enormous cost and elaborate planning that went into surveillance of individuals began to be questioned. Such expense and time were also no indicator of success, and the human factor could easily ruin months of diligent surveillance work.

The shift in communications into the digital realm solved this problem. Not only was the storage of enormous amounts of data now possible, mining through such data was made much simpler. Identifying keep words or phrases in communications such as ‘democracy’, ‘freedom’, ‘free elections’ or ‘Putin must go’ was achievable by a simple search for such keywords or phrases, thus exposing the individuals involved in such communications. What’s more, much of this information is given up freely and widely available online, so authorities do not even have to hack communications for some evidence. Many people elect to freely give away much of their personal data on social networking platforms and are then surprised when authorities know so much about them. Morozov recounts the story of a young activist from his native Belarus who was called into his university to talk to the KGB, which still exists and remains very active in that country.8 The officers had detailed knowledge of Pavel Lyashkovich’s travel arrangements, his involvement with anti-government organisations and his associates in the dissident community, merely from checking his online social networking activity. While it is easy to say that Lyashkovich is to blame for his own predicament, the point is social networking platforms were initially set up as the means for us to stay connected to one another, but now even our most casual conversations broadcast online can be intercepted and misrepresented by authoritarian regimes and others to build a damaging case against us.

A persistent myth with regard to the internet and authoritarian governments has been that they are weak and ineffective regimes that do not truly understand the technology nor how to use it effectively. But anti-democratic forces have become very savvy and immensely sophisticated at manipulating the Web, and one of the main reasons for this is that they have surrounded themselves with the best and brightest talent and online visionaries from their country. The Kremlin, for example, have been particularly successfully in cultivating strong connections with Russia’s vibrant internet culture and have used such experience to their own ideological advantage. Morozov suggested that no one embodies this level of sophistication and linkage more than Kontntin Rykov, a key figure of the early Russian internet and now working as head of the internet department of the Russia’ Channel On, and creator of a range of political websites and staunch ally of Vladimir Putin. In addition to his own personal involvement in Russian politics, Rykov has also reportedly used his internet credentials and relationship with Kremlin officials to involve himself in various political campaigns and referendums in both Russia and other countries.9 Rykov developed tactics to help the Kremlin boost support for its image online and showed how to spread competing narratives on social media to deflect attention away from reporting that was critical of that regime’s activities. These kinds of disinformation techniques and campaigns were used to great effect when the Russian Federation annexed Crimea in 2014. In 2015, Rykov built a new website using the domain Trump2016.ru, which marked the beginning of active campaigning for Donald Trump and Russian interference in the 2016 US presidential elections. According to the special counsel investigation’s Mueller Report, the first indication of Russian interference was the use of the Internet Research Agency, a Kremlin-linked troll farm based in St Petersburg, to wage a social media campaign that favoured Donald Trump and disparaged Hillary Clinton in those elections.10 These campaigns also sought to provoke and amplify political and general social discord across the United States by spreading fabricated election articles and disinformation. For such transnational meddling to truly succeed, trust in traditional media and pillar institutions of state first needed to be undermined and then their creditably destroyed.

A Free and Open Press

One of the central tenets of a democratic system is a free and open press, which is also critical to sustaining the rule of law. In the past, traditional newspapers and broadcasters created the possibility of a single debate or conversation on an issue of national importance. This helped citizens to join together, not in a like-minded set of opinions but rather a singular conversation. Newspaper and broadcast journalism were required to conform to formal and informal ethical and moral codes of practice. But the rapid and wholesale shift in advertising revenue to digital internet giants has, within just a short decade or so, severely damaged the ability of both traditional print and broadcast media to investigate, collect and report on malpractice and convey essential information. This has led some media outlets discontinuing reporting news altogether, and yet others to assume an extreme partisan position in their reporting. The arrival and proliferation of digital online media means there are no longer common debates nor common narratives. Indeed, it is argued, that people have always had different opinions, but now they are presented with different facts.11 Anyone and everyone can now be anonymous and no one needs to take responsibility for what they report or say, or whether it is true or false. In what he termed ‘the cult of the amateur’, Andrew Keen argues that our most valued cultural institutions – our professional newspapers, magazines, music and movies – are being overtaken by an avalanche of amateur, user-generated free content.12 In this present self-broadcasting culture, where amateurism is celebrated, and anyone with an opinion, however ill-informed or ridiculous, can publish a blog or post a video on a sharing platform, the distinction between accomplished and experienced experts and uninformed amateurs has become dangerously blurred. When anonymous bloggers and videographers, unconstrained by professional standards or editorial norms, can alter the public debate and manipulate public opinion, then truth becomes a commodity to be bought, sold, packaged and reinvented. The ongoing erosion of trust, whether it be in the political realm or in the media, weakens the democratic system, and the ever-changing and developing digital ICT landscape and an evolution in the way people now consume news has brought about critical challenges in how we do politics and what we want our society to look like. Among them, fabricated and misleading news stories shared on social media sites and a tendency of readers to only consider news stories that adhere to their own political ideology are undermining press freedoms and leading to levels of mistrust that are severally damaging for a free, open and democratic society.

Lawrence Lessig suggests that unless we find a reason for democracy, there is no fight for democracy to be had.13 As a former self-confessed apologist for the internet, he is now a critic of how digital-enabled news has become fragmented and polarising and is damaging to the ideals and notion of democracy. In his analysis of the twentieth century and the development of the television, he notes that in 1977 almost 90% of people in the United States got their news from just three networks, and these were the sole sources of national and international news. Having this concentration of news through channels that was inherently understandable to everybody – the ordinary citizen as well as the elites alike – gave everyone an egalitarian exposure to politics because they were exposed to trusted sources that created a common understanding and common set of facts. Instead of just polarised extremes voting in elections, ordinary people were much more engaged and turned out to vote based on knowledge and understanding, and this shifted the political landscape of the country in true expressions of democracy. While he does not set out to eulogise this era and does point to issues of bias within this system at that time, he argues that the underlying architecture made it possible for a public to understand a common set of questions and issues. However, in the twenty-first century, there are no longer concentrated, universally trusted sources of news information, and many people now consume news through social media platforms, which are unrestrained by any form of editorship or formal and informal codes of conduct. Such fragmentation of news information means there are no longer any common stories, common facts, and the resulting radical polarisation is damaging for democracy. More worrying, he suggests that the business model now employed by many of these new digital media outlets is to increase this polarisation, thus increasing media brand loyalty.

But traditional media organisations themselves must foot some of the blame. The 24-hour news cycle, made possible by advancements in digital technologies over the recent past, can be distracting for many individuals and, it is suggested, is trivialising much of what we now understand as news.14 The Reuters Institute Digital News Report provides crucial new insights into key issues including people’s willingness to pay for news, the move to private messaging applications and groups and how people see news media from around the world performing their civic and public roles and responsibilities.15 The report is based on a survey of more than 75,000 people in 38 markets, along with additional qualitative research, which together make it the most comprehensive comparative study of news consumption in the world. The report pointed to a complex set of enduring challenges for the news industry specifically and for the media environment more broadly. This included the ongoing disruption of the inherited business model for news, persistent evolution in how individuals use digital media and the ways in which we are constantly reminded of how some of the information we come across is untrustworthy and sometimes spread with malicious intent and social upheaval associated with the rise of populism and general low trust in many state institutions. While the arrival of democracy usually results in enhanced levels of social trust, especially trust in government, this trend commonly reverses after several years or decades of citizens’ unmet expectations. It is argued that government is often the least trusted social actor, ranking below governing bodies, security institutions and the media.16 This finding is somewhat inconsistent across societies however. In some of the most populated countries, such as the United States, Russia and China, people have more trust in the government than the media, which is now the least trusted institution in both the United States and Russia:

From a normative perspective, these results should be viewed with some concern. In democratic societies, the media are entrusted with the responsibility of serving as a watchdog for the public interest and to scrutinize the movements of all three branches of government (executive, legislative, and judicative). A media system in which citizens do not place their trust in will be hardly able to watch over any authority or institution.17

Distrust, Disinformation and Discontent

Returning to the issue of interference in the 2016 US presidential elections, Russian’s use of social media platforms to spread propaganda and disinformation was expansive with the use of Facebook, Twitter and a host of other publicly accessible online outlets coming under the spotlight. Advertisements bought by Russian operatives for circulation on the Facebook social media website were estimated to have reached 10 million users, while many more users were also contacted by accounts created by Russian actors. In total, 470 Facebook accounts are known to have been created by Russians, of those accounts six generated content that was shared at least 340 million times according to research done by Columbia University’s Tow Center for Digital Journalism and New Knowledge, Canfield Research.18 The Mueller Report also found the Russian-financed Internet Research Agency spent some $100,000 for more than 3,500 Facebook advertisements from June 2015 to May 2017, mostly anti-Clinton and pro-Trump proclamations. Facebook initially denied that fake news on its platform had influenced the election and insisted it had been unaware of any Russian-financed advertisements. They later admitted that Russia-based operatives had, indeed, published about 80,000 posts on the social network platform over a two-year period in an effort to sway US political opinion, and that about 126 million Americans may have seen the posts during that time.19 While Facebook claim to have made significant changes to reduce the spread of misinformation and provide more transparency and control around political advertisements, in August 2019, a group of philanthropies working with the company to study the social network’s impact on democracy threatened to quit saying the company had failed to make data available to researchers as pledged.20

Further insight into the growing sophistication of user manipulation for unfettered purposes was made public during the Cambridge Analytica scandal. Cambridge Analytica, also discussed in an earlier chapter, was formed around 2013 initially with a focus on the US elections, with $15 million in backing from billionaire Republican donor Robert Mercer and the backing of former Trump White House adviser Steve Bannon and funding from several UK Conservative Party’s biggest donors. It was an offshoot of the wider SCL Group, which had worked on psychological targeting methods across the world. Having trialled their methodology in poorly developed countries with weak data protection laws, often on behalf of North Atlantic Treaty Organization (NATO) military objectives, they went on to commercialise their product for use in targeting voters during elections and referendum campaigns in some democratic states. Cambridge Analytica markets itself as providing consumer research, targeted advertising and other data-related services to both political and corporate clients and was staffed by mostly British workers. In an undercover investigation by Britain’s Channel 4 News, the company boasted it had developed psychological profiles of voters, which was the ‘secret sauce’ it used to sway such voters more effectively than traditional advertising ever could. They had been able to achieve this by harvesting the data of some 50–87 million Facebook users by a means that deceived both the users and Facebook itself. They were then able to specifically micro-target political advertising back at these Facebook users that would psychologically appeal to some of their base instincts as voters. In identifying people most susceptible to persuasion, they were able to induce them to vote in a particular way: to get voters to see the world as you wanted them to see it. They called this group of susceptible individuals ‘the persuaders’. The company also stands accused of voter suppression, particular in the context of 2010 ‘Do So’ campaign and election in Trinidad and Tobago.21 Ted Cruz had initially hired Cambridge Analytica to help with his presidential campaign, and Donald Trump’s 2016 presidential campaign further utilised the harvested data to customise messages and target specific voters in key swing states. Cambridge Analytica was also hired to assist Leave.eu and the UK Independence Party throughout 2016 and assist with efforts to convince voters in that country to support leaving the European Union in the Brexit referendum.22 In her testimony before a committee of UK parliamentarians, former Cambridge Analytica employee Brittany Kaiser, who has a deep understanding of the operations and techniques used by the firm, suggested that the psychographic micro-targeting used by the company should be classified as ‘weapons grade’ techniques and only used is conflict situations.23

The Retreat from Reality

In the 2016 BBC documentary HyperNormalisation, Adam Curtis maintained that over the past 40 years, politicians, financiers and technological utopians, rather than face up to the real complexities of the contemporary world, had retreated into a simpler version of the world in order to hang onto what they believed to be power.24 And as this unpretentious world grew more and more, people went along with it because the simplicity was reassuring to all. These were mostly the ‘starry-eyed’ former hippies who were hopeful that the digital age would deliver on the expectations and dreams of the 1960s that Morozov had referred to.25 But in this retreat from the reality, an entire generation was beginning to lose touch with politics and the realities of power and governance. Reflecting on the works of William Gibson, who coined the phrase ‘cyberspace’,26 Curtis suggested that by the middle of the 1980s, the banks and new corporations were beginning to link themselves together through computer systems creating a series of major networks of information that were invisible to ordinary citizens and politicians. Such networks gave these corporations and financiers remarkable new powers of control in a cyberspace where there were no laws or, indeed, politicians or governments to protect ordinary citizens. This was a vision of a future of raw brutal corporate power at work. Meanwhile, a group of technological utopian visionaries were emerging on the west coast of America, based around Silicon Valley, who began exploiting Gibson’s ideas of cyberspace and replacing this former frightening dark vision of a world oppressed and dominated by large and powerful corporations with a much safer cyberspace and world where radical dreams could come through. This, Curtis argued, was an ideal place for progressives and radicals to retreat to leaving behind the very harsh real world of Regan’s 1980s America and Thatcher’s Britain. What made this retreat from reality easier for these radicals were their roots in the counterculture of the 1960s and in particular the use of LSD.27

The activists of the 1960s counterculture believed that taking LSD offered much more than just a short escape from their ordinary lives, it opened people’s perception and the mind to an entirely new possible certainty normally hidden from them. It freed them from the normal day-to-day constraints of life, such things as basic political decision-making and the workings of mundane governmental power. The early period of the 1980s saw computer networks appear and offer a new alternative reality, a space to again retreat from the real world, only this time one that was not chemically induced. In this new cyberspace, corporeality people were freed from the realities of normal politics, decision-making and power, and individuals could begin to explore new ways of being and living. Indeed, one of the leading advocates of this new reality, John Perry Barlow, wrote A Declaration of the Independence of Cyberspace28 in response to the passing of the US Telecommunications Act of 1996 in which he sets out a rebuttal to government and against interference with the internet by any outside forces. It declared that the states and politicians did not have the consent of the governed to apply laws to the internet as it was outside any country’s borders. Instead, the internet was developing its own social contract to determine how to handle its own problems, based on language evocative of the US Declaration of Independence:

We must declare our virtual selves immune to your sovereignty, even as we continue to consent to your rule over our bodies. We will spread ourselves across the Planet so that no one can arrest our thoughts. We will create a civilization of the Mind in Cyberspace. May it be more humane and fair than the world your governments have made before.29

Barlow had laid out an alternative existence to the harsh existing world in which people could be freer without the unnecessary oversight of interfering politicians and governments and the old systems of power. This vision has come to dominate the internet as we know it today. Meanwhile, moves towards Perception Management in the United States and Britain were beginning to blur the lines between fact and fiction by telling dramatic stories that would capture the public’s imagination, and it did not matter if the stories were true or false.30 Many of these stories, Curtis claimed, were simply devised to distract people and other politicians from the realities of what was happening around them and the real complexities and challenges of the modern world. Then, with the collapse with the Soviet Union and the crumbling of the iron curtain, a new type of politics began to emerge in the West; one that was no longer focussed on trying to change things but simply managing things, trying to predict risk into the further and seek out ways of avoiding such risks.

The computer age affords us the ability to collect and analyse vast quantities of data rapidly. Computers were also beginning to hold a mirror to individuals, and they liked how that made them feel. They began to not only predict societal risk but also, at the micro level, what people liked and wanted based on their interactions with digital computing. Politics now became just part of the wider goal of managing the world in very simplistic ways. This was best epitomised by the approach adopted by George W. Bush and Tony Blair who viewed the removal of Saddam Hussein in Iraq as a simple fight between good and evil. Politics, democracy and movements for change began to become irrelevant in this new managed world, and a resentment to this began to grow and fester. Even when millions worldwide marched against the impending war in Iraq31 both Bush and Blair ignored this considerable public opinion in favour of their simple narrative and went to war in 2003 in the face of these protests. The effects of this are now widely felt. Not only did millions feel they were lied to when the true extent of the threat from the Saddam Hussein regime become known, but more importantly, they felt helpless in the world and believed that no matter what they did, it had no real positive effect on any eventual outcome. Protesting the war has been a waste of time, and they were demoralised and powerless about the world as it was around them.

Curtis went further to suggest that liberals, radicals and a whole generation had by now retreated into another world that was free of such hypocrisy and what they saw as the corruption of power and politics. They withdrew into cyberspace and here they found comfort in the company of like-minded individuals brought together by filter bubbles and algorithms which sought out and grouped people by means of their own personal data and preferences freely given to mega online corporations and platforms. Such online companionship created echo chambers which worked simply to reinforce beliefs and opinions rather than challenge them. But while such beliefs were being buttressed by compatible thought and opinion, such online environments did not allow for such beliefs to be challenged or developed by opposing viewpoints, a natural and healthy way for individuals to develop their own capacity for critical thinking. Moreover, it did not allow these radical or progressive thinkers to challenge opposing viewpoints from other individuals simply because the online platforms were keeping divergent groups apart. Through the worst of the financial crisis, which began in 2008 and lasted for several years, people retreated deeper and deeper into these online like-minded groups, shouting at the world but failing to understand or lay a transformative glove on power. There is an illusion of control and power online, but it was something completely different altogether; it is delusion.

Back to Reality: Enter the Showman

Then, in 2016, with the election of Donald Trump, the real fallacy of the power of cyberspace and the retreat from the real world became apparent. Here was a president who could regularly and pathologically lie to the camera and mismanage a pandemic in one of the most powerful nations in the world and yet remained unwavering in his views and unchallenged in his actions to any great extent. He is an extremely savvy media operator who can ‘suck the oxygen out of the room’32 and whose own online rhetoric is designed to make those tied to the ideals of liberal democracy shocked, insulted, angry and offended at every opportunity. Cyberspace is the preferred forum for many progressives and liberals to vent this anger, but this did not in any way affect nor change the Trump administration’s policy positions. Meanwhile, Facebook, Twitter and other such social media platforms were teeming with comments and suggestions from individuals and groups feeling insulted, mistreated and angry and suggesting all manner of ways of removing Trump from office. Ironically, these people themselves are more removed from power that ever before in the democracy age. Trump supporters and those on the right of politics have also mobilised on social media platforms, albeit in separate filter bubbles and echo chambers. And as his presidency comes to an ignominious end in 2021, it is now the turn of these individuals to feel cheated, marginalised, angry and resentful, notwithstanding this is based on the lie of a stolen election. People have become more and more dissatisfied and demoralised as time goes by and blame their unhappiness on everyone and everything but themselves. Few point the finger at digital media, which in some cases is having a directly negative impact on our collective well-being.33 Instead, retreating into cyberspace filled with echoes chambers and filter bubbles comforts individuals but also acts to widen the political divide between large sections of society, while fake news and the micro-targeting of voters with machine-generated messages designed to trigger individual prejudices reinforces the anger and hatred of the ‘other side’.

The playbook of many at the extremes of politics and society is to sow confusion and doubt about the legitimacy and authority of the institutions of democracy to represent all of the people. And whether its ‘fake news’ mainstream media or ‘draining the swamp’ of established politicians, digital ICT is today the extremist’s greatest weapon and means of spreading mistrust. There is an almost prefect symbiosis between conspiracy beliefs, such as the QAnon,34 and digital ICT, which acts to channel anger and negative energy towards irrationality and illogical thinking and which ultimately harms social cohesion and democracy. Groups and individuals at the extremes of societal thinking use the maxim: to change society, you must first break it, and thus such forces seek chaos and a total disruption to the existing fundamentals of democracy. All the while big tech looks on and does nothing and in many ways must be viewed as complicit in such chaos.

We now live in a world where most political debate happens on partisan public media outlets or bias online forums. In cyberspace, to challenge authority, protest against injustice or seek redress or social change, you no longer need to take to the streets and convince others who may not hold similar views. The new form of way of revolution looks nothing like past ones. You simply login to a website that articulates like-minded views or converse with others of comparable views on social media platforms: that never-ending cyber filter bubbles and echo chambers that comfort people into thinking they are doing something meaningful. Political action in the form of Facebook and Twitter ‘likes’ and ‘shares’. But such online activism is frequently perceived as nothing more than white noise which in many democratic countries is simply ignored but in more authoritarian states exposes the individual as a person of interest to repressive regimes and security forces. As frustration grows, leading to a change in the public mood, a shift in civic sentiment, and the collapse of political party allegiance, many in the West have turned to the strongman and embraced the concept of all-powerful authoritarian rule. Donald Trump is the poster boy president of the digital age, carried to victory in no small way on the shoulders of big tech and complicacy about the need to continually work to protect and strengthen democracy.

As authoritarianism grows and expands across the world, it is interesting to observe how such regimes develop and receive their support. Anne Applebaum, a leading historian of communism and contemporary politics, maintains that the ‘authoritarian predisposition’ is not one of closed-mindedness but rather simple-mindedness.35 People who are attracted to the notion and ideals of authoritarianism are bothered by complexity; they dislike diversity and prefer unity. They seek understanding and solutions in new political language that makes them feel safe and more secure. There is a revival of nostalgia, a disappointment with meritocracy, there is appeal in conspiracy theories, and a part of the answer may lie in the contentious, cantankerous nature of modern discourse itself; the ways in which we now read about, think about, hear and understand politics.36 And much of this authoritarian validation and unity is seeded and fomented online. Our new digital lives mean people have now become unaccustomed to the normal political and social public discourses that occur in functioning democracies and instead have become entrenched and obstinate in their opinions and mindset.

The Wizards Behind the Curtain

The digital ICT revolution promised much for democratic politics in the twenty-first century but so far has delivered little but disruption. The dawn of the internet age was to bring a decisive shift towards the citizen and information was to become free and limitless, and enlightenment and empowerment would follow. But while digital technologies provide us with the opportunity to accumulate quantities of information that one time may not have been possible, big tech and the state remains much better equipped than any private citizen to take full advantage of this opportunity. In many ways, digital technology has been weaponised against the very system it was purported to support and defend and the citizens it was meant to engage, protect and enlighten. Authoritarian regimes across the world have seized upon the opportunities provided by such technology to increase surveillance and control of their people while simultaneously spreading misinformation and confusion, undermining many of the established Western liberal democracies. It would be rather naïve to think that democratic governments are not also regularly using similar digital surveillance technique under various guises and security apparatuses. And all the while big tech is the real big winner. The pioneers of surveillance capitalism Google were emboldened and benefitted from historical events when a national security apparatus, galvanised by the attacks of 11 September 2001, saw the emergent capabilities and the promise of some certainty in how Google’s storage and use of huge stocks of personal data could be used to shadow and predict the behaviour of individuals.37 Zuboff believes that the concepts underpinning surveillance capitalism are facilitating the overthrow of the people’s sovereignty and is a prominent force in the perilous drift towards democratic deconsolidation that now threatens Western liberal democracies themselves.

And this is a common complaint in the twenty-first century; democracy itself has lost control of corporate power in the form of big tech companies, who use whatever means possible to hoard vast wealth and influence while fuelling inequality, damaging the planet and avoid paying their fair share of taxes.38 Today’s big tech behemoths exist in a political culture that has grown accustomed and accommodating to their every need, and Runciman argues, in the United States, this was further cemented by the Supreme Court decision in the Citizens United case of 2010 to grant corporations the same rights to free speech as individual citizens.39 The ideals and very notion of liberal democracy are now under constant pressure from many angles, and the traditional hierarchy of power is also under increasing danger. The power of modern corporate power, in the form of big tech, has grown exponentially over the past decade to the point where it now has the wherewithal to undermine how democracy itself operates and not be overly worried about the consequences. A major imperative now for every citizen and democratic nation must be to reassess the inequitable influence of big tech corporate power and the internet, particularly as it relates to our personal data, and to question: who owns and controls such power, and what right do they have to use and misuse our personal data to undermine our key democratic institutions? Democracy must be seen to represent the wishes of the people rather than viewed as a system of corporate tyranny.

#### AND resolving antidemocratic fissures deters adversarial hybrid attacks.

Mikael Wigell 21, Director, Global Security, Finnish Institute of International Affairs. Adjunct Professor, University of Tampere, "Democratic Deterrence: How to Dissuade Hybrid Interference," The Washington Quarterly, Vol. 44, Issue 1, 03/23/2021, T&F.

Western democracy is being attacked like never before, but not through overhyped “hybrid warfare.” It is in reality threatened more acutely by hybrid interference, attacks that are often subtle, manipulating for cover the very same liberal democratic values that the attack is designed to subvert. The cornerstones of Western democracy—state restraint, pluralism, free media, and economic openness—provide openings for authoritarian actors to interfere in democratic society through a host of covert, non-military means calibrated to undermine their internal cohesion and accelerate political polarization.

For instance, disinformation campaigns have become increasingly evident since the 2016 US elections and have stepped up in the midst of the COVID-19 crisis. Russia, and increasingly China, are deploying disinformation to aggravate the public health crisis in Western countries. Exaggerated and fabricated stories of how Western governments have been mismanaging the spread of the coronavirus have been used to play on the anxieties of Western populations.1 While not all such disinformation efforts succeed in persuading the public, the cumulative impact can be very effective in sowing distrust, rendering democratic states less capable of countering either the epidemic or the aggression itself.

Financial support is also being channeled to radical political parties and movements to accelerate centrifugal forces within and among Western democracies.2 Part of this toolbox involves exploiting the economic openness of Western democracies to capture strategic sectors of the economy—such as critical infrastructure, finance, and media—by which these authoritarian actors can attempt to destabilize Western democracies and purposefully corrupt them.3

Democracies urgently need to find means to defend against such hybrid interference without jeopardizing the values that they are meant to defend. Extending state control over civil society is not a viable liberal democratic strategy. Neither should Western democracies mirror the use of weaponized corruption, disinformation, election meddling, and other means of hybrid interference, as this would only further erode liberal democratic values around the globe.

For all their amassed military might, a particular advantage of Western democracies lies in their soft power and inclusive politics. Western democracy still commands widespread attraction and political legitimacy, and open societies are agile in responding to strategic challenges. Rather than the rigidness of state-based solutions, Western democracy harnesses market- and society-based approaches to dealing with risks and threats. These can readily be used to strengthen deterrence against hybrid interference. It is crucial to recognize the deterrent value of democracy itself, namely how it can provide means for deterrence by both denial and punishment. Wielded with confidence, democracy itself is a potent strategic weapon.

Open societies are agile in responding to strategic challenges

This article outlines the strategic logic of hybrid interference and how it puts Western democratic governability in jeopardy. It argues that deterrence policies need to be revamped in the face of this new challenge and suggests a new strategic concept—democratic deterrence—as a framework for dissuading hybrid interference. It asks what deterrent value democracy itself has and envisages a host of non-military means to adapt deterrence to the current non-military challenges. By evolving the concept of deterrence in this way, democratic deterrence shows how liberal democratic values need not be security vulnerabilities, but can be turned into strengths and tools to credibly deter hybrid aggressors, while making our Western democracies more robust and resilient.

The Strategic Practice of Hybrid Interference

Much of the debate on new “hybrid” threats has revolved around “little green men” and other grey zone military tactics, or hybrid warfare, essentially a military approach to conducting “indirect war” under special circumstances.4 The more pressing challenges from a Western perspective are the more subtle, non-military activities deployed by authoritarian regimes to penetrate democratic society. “Hybrid interference” is a concept developed to capture non-military practices for the mostly covert manipulation of other states’ strategic interests.5 As such, it bears resemblance to what was referred to as “active measures” during the Cold War and, more recently, in Russian strategic debate as gibridnaya voyna.

The idea of gibridnaya voyna is to avoid the traditional battlefield with the aim of destroying “the political cohesion of an adversary from the inside by employing a carefully crafted hybrid of non-military means and methods that amplify political, ideological, economic and other social polarisations within an adversary’s society, thus leading to its internal collapse.”6 While keeping diplomatic relations intact, and thus not breaking any official threshold of war, the aggressor mobilizes oppositionists and radicals within the target state through a host of means ranging from disinformation campaigns to corrupting political actors and financing subversive movements, carefully synchronized to compound the effect.

Hybrid interference avoids the use of overt kinetic means in order to maintain plausible deniability. Yet, it may include the use of targeted violence, through proxies, to inject fear and exploit emotional pressure points in the target society. In 2014, for example, pro-democracy protesters in Taiwan were attacked by some with links to the Chinese Communist Party to incite political tensions and undermine democratic governability.7 In Montenegro, Moscow instigated a coup attempt before the 2016 elections. The attempt involved harnessing its close relations to the Orthodox Church and the Serb minority population to foment distrust in the democratic process and, in the final phase, using Russian intelligence operatives disguising as police officers to create disruption by shooting protesters and blaming the Montenegrin government for killing innocents.8

Central to hybrid interference is subversion. Subversion refers to an aggressor state’s purposeful attempt to destabilize and undermine the authority of a target state by using local proxy actors.9 It specifically involves the use of disinformation and economic inducements to recruit and assist these actors inside the target country, detach their loyalties from the target government, and use them as interlocutors to transform the established social order and its structures of authority and norms. The aim is to weaken democratic governance and norms as a means of enhancing their own authoritarian standing. Not only are weakened democracies less able to directly confront these authoritarian aggressors, but they will also look less appealing as models of success and partners for others. By portraying Western democracies as corrupt and ungovernable, authoritarian regimes such as China, Iran, Russia, and Turkey are less at risk of being overthrown by their own populations.

As such, hybrid interference is designed as a flexible approach in which the tools and tactics can vary but will always be tailored to manipulate existing cleavages and sow internal dissension in target countries and alliances. Hybrid interference does not adopt a one-size-fits-all approach but exploits specific vulnerabilities depending on the context in the target country. The hybrid aggressor interferes in domestic politics by seeking to amplify divisions and hatred, undermining the “civic culture” that has been found to be so important for democratic governability by tempering the intensity of political conflicts and cleavages.10

The migrant crisis in Europe offered an excellent opportunity for authoritarian aggressors. By exposing rifts between “liberals” and “anti-liberals,” it allowed Russia and Turkey to leverage refugees as a disruptive force, fanning the already simmering political tensions in Europe. Following the outbreak of the Syrian civil war, Russia and Turkey began actively pushing migrants over the borders to Europe, while simultaneously engaging in disinformation campaigns by playing up rumored (or actual) misdeeds by immigrants and portraying European governments as unwilling or unable to manage the influx of people.11 To compound the polarizing effect, Russia also began channeling money to anti-immigrant and anti-EU political parties and movements such as France’s National Front.12 In this way, particularly Russia has contributed to Europe’s surge in anti-immigrant sentiments and populist support, as well as rising democratic dissatisfaction.13

Democratic Deterrence as a Novel Strategy

Western democracies urgently need to find counter measures against hybrid interference, recognizing that traditional military deterrence only works against predominantly military threats. By reducing clarity about who is doing what, or even whether somebody is actually doing anything, hybrid interference complicates traditional deterrence.14 Yet, the attribution problem is not insurmountable. The responsibility for interfering in US elections, for example, was traced and attributed to Russia, though only after a painstaking process involving much controversy and resources.15 But the question, then, even after the attribution problem is solved, becomes what to do about it? Any prudent deterrence posture needs to avoid unnecessary escalation, and a military response to political interference does not seem proportional.

The strategic concept of democratic deterrence suggests a novel way of thinking about deterrence to dissuade these hybrid interference activities by authoritarian states. Table 1 summarizes the differences between traditional military deterrence and democratic deterrence.

[Table omitted]

First, in contrast to traditional deterrence that is state-based, democratic deterrence rests on a whole-of-society approach, albeit one in which the state retains a coordinating role. It harnesses market- and society-based actors in an effort to pull together resources and take full advantage of democracy’s societal strengths and cultural capital. This difference is important because in this new era of subversive politics, where the classical Westphalian dichotomy between internal and external state affairs has been blurred, deterrence is harder to achieve by state action alone. Deterring hybrid interference requires a whole-of-society response whereby various societal actors build resilience capacities, support the state in maintaining preparedness, and ensure the continuity of vital societal functions and supply lines. For instance, private actors often own full or partial stakes in critical infrastructure such as energy pipelines, undersea cables, railways, banking and finance, health services, and food supply. Ensuring that they live up to their responsibilities with regard to safety measures must form an essential part of any modern defense. The whole-of-society approach is thus an inclusive model of cooperation and joint preparedness that aims to bring all relevant actors together into a comprehensive system of deterrence. It involves an effort to diversify and devolve responsibilities for security production to market- and societal-based actors, while maintaining a strong coordinating role for the state.

Second, while traditional deterrence relies on hard power, democratic deterrence uses the soft power base of Western democratic societies.16 Soft power rests on the ability to attract, and liberal democratic values and norms continue to exercise a strong international pull, not least among autocratic subjects.17 Democratic norms and values are thus strategic assets that can be used to deter authoritarian regimes. Western democracy promotion efforts have helped catalyze regime change in many parts of the world. During the Cold War, Radio Free Europe, a radio broadcasting company set up by the United States, contributed to the demise of communist regimes in Eastern Europe.18 By signaling strong and concerted preparedness to vigorously engage in democracy promotion, Western democracies can again help deter authoritarian leaders. Crucially, in an information age, when power is less hierarchical and social networks have become more important, projecting soft power is not only a matter for states. Nonstate actors such as NGOs, research institutes, and corporations are also important for generating soft power. The flexibility of non-governmental actors, such as the National Endowment for Democracy, in building relationships and networks across borders can provide a crucial gateway to strengthen normative legitimacy and mobilize the cause of advancing democracy.

Third, and related to the above, democratic deterrence crucially relies on non-military, democratic means. Democratic values and instruments such as transparency, the rule of law, and citizen activism provide tools for non-kinetic deterrence. Functioning under the threshold of war, they are well calibrated to avoid escalation, while helping deter grey zone activities such as hybrid interference. Hybrid agents thrive on being covert, so transparency is a key means of deterring hybrid interference. Similarly, a strong rule of law is essential to deny efforts to destabilize and accelerate polarization in democratic societies by means such as weaponized corruption. Citizen activism provides a force multiplier in efforts to both deny as well as punish hybrid interference by harnessing civil society’s capabilities and agility.

Fourth, while traditional deterrence often relies on “in kind” measures, namely a symmetrical response, democratic deterrence takes the response outside the domain in which the action occurs. In fact, asymmetry is a necessary feature of democratic deterrence. Responding in kind to hybrid interference—and thus mirroring the use of election meddling, corruption operations, disinformation campaigning, and other means of sharp power—will only contribute to the further erosion of liberal democratic values and undermine the normative legitimacy of Western democracy.19 Moreover, because outright attribution is a troublesome process with regard to hybrid interference, with the hybrid agent using proxies and artificial intelligence (AI) for obfuscation purposes, symmetry can seldom be the optimal response. Instead, by relying on a “democratic playbook” of response options that draw on Western democracy’s soft power base, outlined below, hybrid interference can be deterred without compromising normative legitimacy.

Democratic deterrence takes the response outside the domain in which the action occurs

Lastly, whereas traditional deterrence aims at wholly deterring any aggression, democratic deterrence accepts that some actions cannot be deterred. Indeed, absolute deterrence may even induce hostile actors to seek alternative and more dangerous ways to attack Western democracy. Unlike nuclear deterrence, deterrence against hybrid interference is more like crime prevention—not all crimes can be deterred and not all represent significant threats to national security. Conscious about the need to tolerate a certain set of hostile activities, democratic deterrence settles for a more restrictive aim whereby external interference is not wholly deterred, but modified to render it less effective and frequent.

Advocating for democratic deterrence does not mean that traditional deterrence has become obsolete. Military deterrence remains vital for dissuading armed aggression and various forms of sabotage. It may also contribute to deterring hybrid interference by instilling doubt about the level of response. Traditional military deterrence policies therefore need to be maintained and perhaps even strengthened. Yet, the argument here is that traditional deterrence measures fall short of effectively dealing with the challenge of hybrid interference and therefore need to be complemented by new measures, namely those proposed here.

A Two-Pronged Democratic Deterrence Strategy

Hybrid interference calls for new tools of non-military deterrence. Importantly, any new deterrence posture needs to maintain the openness of democracy and avoid sacrificing any of the Western democratic cornerstones in the name of security. Deterrence is based on increasing the perceived costs of hostile actions to the point of outweighing their potential benefits. In deterrence theory, measures to dissuade hostilities are often divided into two broad categories: denial and punishment.20 Both categories are also applicable to democratic deterrence. Indeed, much like traditional military deterrence, democratic deterrence can also be designed as a two-pronged strategy of deterrence by denial (i.e., resilience) and by punishment (i.e., compellence). Both are necessary to stop hybrid interference attempts.

Strengthening resilience is a necessary building block of any democratic deterrence posture, but it is insufficient as it is unlikely to deter hybrid interference. Russia’s meddling in Western democratic elections, for instance, has continued, despite being publicly exposed and despite measures to strengthen resilience against such external interference. Without any credible deterrence by punishment, these attacks are a relatively low-cost endeavor and can thus be expected to continue. Thus, democratic deterrence must consist of both measures to enhance denial through resilience and punishment, which has not been sought to date.

Strengthening resilience is a necessary building block of any democratic deterrence posture

Deterrence by Denial: Improving Democratic Resilience

Resilience refers to the ability to absorb, adapt, and recover from disruption and duress. High resilience will make it difficult for an aggressor to achieve its strategic aims, thereby making an attack not worth the costs and effort.21 Improving resilience helps modernize total defense doctrine by addressing vulnerabilities across state and society. The emphasis needs to be on continuity management of vital societal functions, supply lines, and critical infrastructure, including democratic infrastructure such as elections. As many critical functions are operated partly or even wholly by private sector actors, public-private cooperation is paramount for improving democratic resilience. Small Northern European countries accustomed to the idea of state-society collaboration and pooling resources to balance Russia may serve as an example: Finland’s comprehensive security model builds on enhancing preparedness through sustained cooperation between authorities, business operators, and civil society organizations in order to secure the vital functions of state and society.22 Similarly, democratic deterrence involves preventing or making hybrid interference difficult by harnessing and doubling down on liberal democracy’s strengths: activating autonomous civil society, increasing transparency of money flows, and broadening inclusive politics.

Activate civil society

While the open environment of Western democracy presents loopholes for covert interference, it simultaneously provides an enabling environment for citizen activism and market-based innovation. Citizen activism can play a major role in identifying interference and building institutional and societal resilience against it.23 The essential watchdog functions of the open media environment serve to enable citizen activism by shedding light on hybrid interference. Investigative journalism is a pertinent example, as evidenced by novel online sources like Bellingcat, whose investigations helped solve the Salisbury poisoning case, in which Sergei Skripal—a former Russian military officer and double agent for the British intelligence services—was poisoned together with his daughter by assassins connected to the Russian GRU military spy agency.24 Similarly, the Organized Crime and Corruption Reporting Project, a network of investigative reporters, helped uncover a Russian money laundering scheme through which funds were channeled to groups lobbying for closer relations between EU countries and Russia.25

Citizen activism can play a major role in identifying interference and building resilience against it

Western democracies should encourage investigative civil society groups and media to monitor and detect hybrid interference. Specific measures should include developing rapid alert systems and media literacy programs as well as training media professionals themselves in recognizing fake news. Finland’s Mediapooli, a joint organization set up by the country’s media companies, helps train journalists through capacity-building programs, anti-fake news education, and freely distributed guides on how to better protect sources and counter disinformation.26 In the United States, the Countering Foreign Influence Task Force of the Department of Homeland Security, in coordination with the FBI, began operations for countering disinformation before the 2018 US midterm elections. Its focus has been on raising public awareness about the dangers with foreign disinformation campaigns and working with social media companies and academia to better recognize, understand, and build resilience against foreign disinformation.

#### Pervasive disinformation AND hybrid interference trigger nuclear war.

Harold A. Trinkunas et al. 21, Deputy Director & Senior Research Scholar, Center for International Security and Cooperation, Freeman Spogli Institute for International Studies, Stanford University; Herbert S. Lin, Senior Research Scholar, Cyber Policy & Security, Center for International Security and Cooperation. Hank J. Holland Fellow, Cyber Policy & Security, Hoover Institution, Stanford University; Benjamin Loehrke, Program Officer, Nuclear Policy, Stanley Center for Peace and Security, "What Can Be Done to Minimize the Effects of the Global Information Ecosystem on the Risk of Nuclear War?" in Three Tweets to Midnight: Effects of the Global Information Ecosystem on the Risk of Nuclear Conflict, Chapter 10, Hoover Institute, 2021, pg. 193-201.

On August 11, 2017, President Trump tweeted: “Military solutions are now fully in place, locked and loaded, should North Korea act unwisely. Hopefully Kim Jong Un will find another path!”1 This message followed months of escalating rhetoric and military posturing between the United States and North Korea. The crisis became acute enough that, near the height of tensions in July 2018, polling showed that 60 to 75 percent of Americans were worried about the possibility of war between North Korea and the United States within the following six months.2 Tweets from President Trump often drove or narrated the crisis, adding fears that instantaneous, direct, 280-character threats could lead directly to nuclear war. As former acting undersecretary of defense for policy Brian McKeon testified at a Senate Foreign Relations Committee hearing on presidential nuclear authorities, “The statements the president makes through his Twitter account no doubt cause concern and confusion on the other side of the Pacific. . . . I’ll be very worried about a miscalculation based on continuing use of his Twitter account with regard to North Korea.”3

As this case illustrates, the new global information ecosystem may be having an important impact on the evolution of international crises. Widespread access to social media on a global scale has accelerated news cycles in traditional media and made it easier to spread mis information and disinformation. Intemperate, ill-considered, and impulsive outbursts have become an important part of crisis dynamics. In the decade since the founding of Facebook and Twitter, social media have added new arenas to conflicts in the Persian Gulf region among Iran, Saudi Arabia, and their respective allies; among Russia, Ukraine, and NATO; between nuclear-armed India and Pakistan; and, as we have just considered, among North Korea, Japan, South Korea, and the United States.4 If we were to include information operations meant to influence governments and publics, we could extend the list of cases to include Russian interference in elections in the United States, the United Kingdom, Germany, Spain, Italy, and France; operations by the Venezuelan government against its neighbors in South America; and operations between China and its neighbors in East Asia.5 Some of these crises involve nuclear-armed powers. Were one of these crises to spin out of control, the outbreak of nuclear war could have a catastrophic impact on humanity. Even a modest exchange involving one hundred relatively small warheads has the potential for producing a nuclear winter with dramatic effects on global climate and the prospects for human survival.6

While disinformation and misinformation have always been part of conflict, the chapters in this volume outline how the new global information ecosystem has created conditions for the spread of disinformation, misinformation, and other malign information in ways that threaten crisis stability, even nuclear crisis stability. Scholars of crisis stability have had well-established frameworks with which to analyze deterrence, decision making, and the role of public opinion in foreign policy. These approaches principally rest on rational actor models. While they acknowledge that misperception and miscalculation can have an impact on crisis stability, they tend to assume that leaders will make policy decisions rationally and analytically, based on the best available evidence and with the national interest foremost in mind.7

Social media and their disruptive effects are cause to reassess how existing analytical and theoretical frameworks for understanding crisis stability might be affected by the evolution of today’s information ecosystem. This volume fills a gap on whether, when, and how social media could contribute to international conflict—including deterrence failure and nuclear war. In particular, it makes four contributions.

First, it incorporates findings from cognitive psychology and decision analysis into analyses of how leaders and publics receive, process, and act on information, misinformation, and disinformation in the emerging global ecosystem. It highlights how social media have an impact on how much information individuals receive, how they receive it, and, in turn, how these factors affect and may increase the likelihood of engaging in heuristic thinking (i.e., intellectual shortcuts) to manage the overwhelming volume of information available.

Second, the authors in this volume examine how cyber-enabled influence operations may be deliberately conducted via the new tools made available in the present information environment to take advantage of human cognitive biases and affect the perceptions, preferences, and decisions of both publics and leaders in times of crisis.

Third, this volume examines how the intersection of human propensity to heuristic thinking and cognitive bias may have a dangerous impact on international crisis stability. Such mental shortcuts are common to decision making. The emerging global information ecosystem, combined with deliberate influence operations designed to affect leader and public perceptions, could further wear on leaders during crises—potentially even those involving major nuclear powers and the risk of war.

And fourth, this volume assesses the limits of what adversaries may actually be able to accomplish in the present information environment, including the risk that influence operations may cause blowback on the perpetrators. In addition, public preferences may actually be fairly resilient in the long run in the face of deliberate attempts to influence mass opinion, even if these may have an impact in the short run.

Human Cognition, Heuristic Thinking, and Implications for Crisis Stability

Digitization and global communication technologies make generating and sharing new information possible at an unprecedented speed and scale. Social media platforms provide vehicles (in many cases tailored to take advantage of human cognitive biases) via which to maximize the impact of targeted persuasion. Each year, more people around the world are part of this information ecosystem, as mobile phone penetration globally is estimated to reach five billion users in 2019 (and there is no reason to expect this trend to slow down).8 The transformation of the global information ecosystem is not just about speed, ease, or scale of communication. It has crucially democratized information production and information dissemination. Moreover, it is increasingly apparent that the new global communications ecosystem is producing new opportunities to influence humans by playing on traditional cognitive biases that we use to process information. Audiences could be more susceptible to such efforts when faced with time pressure, high volumes of information, and appealing post-truth narratives that are preferred by significant segments of the global public instead of evidence-based journalism and policies. Taken together, these trends call into question whether traditional models of crisis stability, which assume rational decisions made by elites based on the best available evidence, are an accurate way to understand the likely evolution of future international conflicts.

In chapter 2, Rose McDermott explores the psychology of the post-truth political environment. Applied to the political environment, post-truth denotes “circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief.” Absent special mental discipline, story and narrative are more important in shaping a person’s views than empirical fact or logically reasoned conclusions—and this applies both to ordinary citizens and to leaders. Importantly, McDermott argues that most people will regard a plausible story as true, whether or not it is in fact true. McDermott also points out two exacerbating factors. First, the decline of public trust in institutions and expertise has left individuals on their own to gather information and to make judgments about what to believe for themselves. Second, the rise of social media as primary information sources means that those who rely on such sources do not have the benefit of intermediaries who fact-check and place information in context. In this environment, people are far more likely to fall back on their own intuitive thinking, which places much higher value on factors such as simplicity, familiarity, consistency with prior belief, and how many other people appear to believe the same things. Analytical evidence-based thinking will struggle to keep pace.

Paul Slovic and Herb Lin consider the psychology of nuclear decision making, especially during crisis. Such decisions involve the highest possible stakes. The authors point to several psychological phenomena that affect nuclear decision making. Psychic numbing refers to a devaluation of life when large numbers of deaths are contemplated—the death of one innocent civilian is regarded as a tragedy, whereas the death of a million is merely a statistic. Indeed, in some cases, the death of millions is regarded as less tragic than the death of a few. Psychological devaluation of life likely underlies the ability of nuclear planners and decision makers to proceed in ways that they believe to be consistent with laws of war that are intended to minimize harm to innocent civilians. Tribalism reflects an “us versus them” mindset, enabling “us” to hate “them.” Tribalism enables the dehumanization of the enemy and treatment of the enemy in ways that do not seem to violate the laws of war. Decision makers often avoid making trade-offs between competing values, such as the value of protecting national security versus protecting noncombatant enemy civilians. Rather than finding a common currency to evaluate trade-offs, they will often prioritize different values and focus on achieving those of highest priority. Thus, a decision maker may well favor security objectives over lifesaving objectives because the former are more defensible. Combined with the affordances of social media (such as their use of short, simple messages and evocative visual and auditory content), the existence of such psychological processes means that social media messages are more likely to be processed with fast, intuitive thought rather than with reflective, deliberate thought. The same is true of leaders and decision makers who are active social media users, and they are just as likely to be pushed by their social media usage into fast, intuitive thought. Slovic and Lin conclude that where such leaders are concerned, exposure to social media may well increase the likelihood of taking rash action and of premature use of force.

Cyber-Enabled Influence Operations: The Impact of Disinformation on Leaders and Publics

The present revolution in the global information ecosystem has made propaganda cost effective again. Manipulating information with the intent to persuade is a tried-and-tested part of warfare, and skeptics are right to note that there is nothing new about propaganda per se.9 But the current information environment substantially reduces barriers to the conduct of information operations not just for great powers but also for small and middle powers as well as for nonstate actors. Unlike offensive cyberoperations—which require substantial investments in sophisticated cybercommands, recruitment of scarce hacking talent, and maintenance of up-to-date cyberweapons based on fresh exploits—information operations are much more affordable.10 As we learned from the investigation into Russian targeting of US elections, influence operations may cost millions of dollars, but they need not cost tens or hundreds of millions of dollars.11 Moreover, operations can be conducted on platforms made available largely for free by major social media platforms, designed to be used by the general public with the most minimal training. Lowering costs along all dimensions enables a wide array of states, great and small, and nonstate actors such as political parties and civil society organizations to conduct influence operations cheaply. In addition, states traditionally seem to treat influence operations as falling short of the threshold of armed conflict (more akin to subversion), which means that even great powers have avoided responding to such attacks by other state actors with military force. Since costs are low, both in financial terms and in terms of the likelihood of retaliation, we should expect the widespread use of influence operations intended to affect the behavior of leaders and publics, even against the great powers and even by weaker actors in the international system.

Misinformation and disinformation on social media have the potential to contaminate information flows, which could affect behavior during crises, as Mark Kumleben and Samuel Woolley show in chapter 4. People increasingly turn to social media for information during emergency situations, which creates an opening for nefarious actors to exploit that information ecosystem. For example, during a military crisis, an adversary could use a variety of computational propaganda techniques to interrupt and confuse information flows on social media in order to encourage publics and leaders to behave in a way that suits the adversary’s interests. Kumleben and Woolley explain some of the more important of those techniques and give an overview of how they have been used in political conflicts. The cases that the authors use illustrate the potential effects of misinformation and disinformation during military crises. The 2018 false missile alert in Hawaii is a useful hypothetical on how computational propaganda could provide an adversary with a cost-effective means to erode a target state’s civil defenses and interrupt its ability to mobilize resources. Political leaders might also be susceptible to digital information operations during crises. The microtargeting of Jeremy Corbyn by members of his own 2017 Labour Party campaign staff shows that disinformation on social media could affect political decision making. By showing how computational propaganda has the power to affect behavior, Kumleben and Woolley highlight the strategic importance of the information ecosystem during crises.

State and nonstate actors are already engaged in information operations designed to affect interstate relations, as Kate Starbird outlines in chapter 5 in this volume. Using the techniques analyzed by Kumleben and Woolley, these actors are conducting influence operations online to influence political discourse and generate false information, most likely with the intention of generating confusion and mistrust among their adversaries and competitors. Starbird’s work outlines how deliberate efforts by state actors, such as those aligned with Russia, can influence broader online conversations and activism among sympathetic audiences. In the case of NATO, both alt-right and fringe conservative voices and international far-left activists converged on a shared anti-alliance message that was influenced and driven in part by state-sponsored online actors working via social media. There is a pattern of state actors and state-backed trolls infiltrating authentic online and social media–based activist communities on both the right and the left to reshape their activities so that they unwittingly support state-sponsored messages and objectives, in this case Russia’s anti-NATO activities. The long-term impact of these activities remains to be seen, but they are already shaping conversations about and among major international actors, in this case NATO, possibly shaping the future strategic environment in ways that could undermine popular support for alliance activities to deter Russia.

The Risks to International Crisis Stability from the Global Information Ecosystem

During the Cold War, government leaders of the major nuclear powers received information from military and intelligence services that, while of course vulnerable to many errors, was nonetheless subject to a process designed to produce verifiable data on which leaders could base decisions. Publics received information via gatekeepers, whether in the form of official or private media, that also subjected information to a vetting process, admittedly not always designed to produce truth but at least to produce consistency and a consensus view of reality among audiences.

Publics and leaders are today exposed to masses of unverified information produced at high speed and distributed at high volume for next to no cost. It is much easier to produce polarization in target populations, to spur storms of public opinion to influence enemy leaders, to leak information deleterious to adversaries, and to conduct influence operations designed to target the psychology of enemy publics. Moreover, the same techniques, as Kristin Ven Bruusgaard and Jaclyn Kerr suggest, can target leaders, affecting perceptions of crises and of adversaries’ intentions. We already know that major government officials pay attention to social media, and they are also subject to the same effects from the global information ecosystem as the publics they lead. This raises the real possibility that influence operations may become an additional contributing factor to growing crisis instability in the world today.

In fact, the deployment of post-truth information during crises may contribute to escalation dynamics in dangerous and unpredictable ways. In the current information environment and given human propensity for heuristic thinking, deployment of convenient half-truths, rumors, or “extra-factual information,” as Kelly Greenhill argues, is attractive because it is a powerful mobilizer of public opinion and can magnify signals of resolve in international crisis. But precisely because it is so powerful and provocative, it can lead adversaries to escalate rather than back down. It can alarm public opinion among adversaries, putting opponents in the position of having to resort to their own escalation and provocation or else appear weak. In addition, as Jeffrey Lewis also documents in chapter 8, there is the possibility that both the general public and elites in the provoking country will come to believe extra-factual information, making it difficult to build off-ramps from international crises for fear of appearing weak or losing face. It may become difficult or impossible to “walk back” or discredit extra-factual information in a global information environment too prone to magnifying human heuristic thinking and spreading information that is appealing even if untrue.

#### A competitive tech sector reverses growing European populism.

Johannes Reck 21, MS, Swiss Federal Institute of Technology, "The Regulation of Tech Monopolies Will Decide the Fate of Western Democracies," Business Insider, 02/08/2021, https://www.businessinsider.com/regulation-of-big-tech-decide-fate-of-western-democracies-2021-2.

Over the past decade, the transformation of society's relationship to these companies is unprecedented. We use the products of the GAFA companies to manage almost our entire professional and private lives. And during the coronavirus crisis, the value of the tech giants have increased even more. The GAFA companies, taken together, currently have a market capitalization of 6 trillion euros, which is about five times as much as the entire German DAX index.

The dominance of the GAFA companies has become so great that individual governments have no choice but to submit themselves to them.

Google and Apple dictated the interface for the Corona app, not the other way around. When Australia wanted to introduce a new law on fair remuneration for its publishing houses, Google unashamedly threatened to switch off its search engine in the country.

It seems that many German and European politicians have given in to the superior power of the GAFA companies. Complicated regulatory issues are not very attractive election campaign topics, after all. This powerlessness is also reflected in the practically nonexistent media debate.

It is my steadfast conviction that how we handle the US tech monopolies will dictate Europe's future. Never before in the history of our continent have so few companies possessed so much power — never before were they able to exert such a profound influence on our lives.

And even if it seems almost unimaginable – the technological revolution is not yet over. Quite the opposite is true. We are still in the early phases. If we want, we can still make the rules that will shape our future. But the room for the maneuvering needed to do this is shrinking, and time is running out. If we want to know what to do next, then it is vital that we gain an in-depth understanding of the mechanisms by which the tech monopolies operate.

The tech monopolies survive on data, algorithms, and capital

In the 20th century, a company's wealth came from its factories, machines, and its qualified employees. In Germany and the EU, our entire education and economic systems are designed based on this formula. The problem is that the digital world functions completely differently.

In the digital world customer behavior is evaluated in real time. This allows digital services to be continuously improved and aligned precisely to customer needs and desires. The more data is collected, the better the algorithms work and the more relevant the offers presented to customers become.

Over the past decade, the GAFA companies have built a huge competitive advantage because they control the operating systems, the search engines, the browsers, and the cloud infrastructure. They also own the shopping marketplaces, the communication platforms, the networked household appliances, and the app stores.

To return to our picture of the 20th century, the tech monopolies not only have factories and machines, they also increasingly own the entire infrastructure of the value-creation chain, including all the businesses and all of the communication channels to the customer. With every new customer, this value-creation chain becomes more efficient and more profitable. Competition in the GAFA infrastructure is only allowed for as long as the competitor compensates the monopolist or helps the monopolist towards even more. expansive growth

The capital that the GAFA companies suck out of the system using this mechanism is fed straight back into undermining the competition or accessing new areas of business. The vendors on Amazon's marketplace are just as dependent on the goodwill of the platform as media companies are on Google or Facebook. This goodwill can only be acquired by consistently providing access to all content and data, thus continuing to feed the monopoly and making it more efficient.

The enormous profitability of these infrastructure services allows the GAFA companies to invest in new fields of business on a very long-term basis. To do this, the monopolists are willing to accept high losses for many years in order to weaken the competition and build up market shares. Due to the enormous market capitalization of the GAFA companies, it is necessary for them to continue to occupy large and lucrative markets to keep the expansive system going and to increase their stock market price.

Consumers are the human shields protecting the monopolies

The favorite argument put forward by the GAFA companies to divert attention away from their position of power is that the products are free for the consumers and that they greatly improve the lives of all of us.

That is a sneaky argument that might seem plausible at first glance. However, it is a deliberate trick. The products sold via the platforms must finance the high profits of the GAFA companies — which means that the users are paying indirectly.

What is even worse: they are forced to hand over their personal data. The tech giants have no reservations whatsoever when it comes to analyzing their customers down to the smallest detail. They know that we can no longer live without their products, and that is why antitrust fines or occasional political objections are a small price for them to pay on the way to increasing their market dominance.

Our dependency on the GAFA companies' infrastructure makes things extremely difficult for our government authorities because they want to help the citizens, not cause them problems.

A life without iPhones, Google Maps, WhatsApp, or Amazon is hard to imagine and not exactly something we are striving for. The bundling together of the different services such as the integration of maps into Google's search engine, or the linkup between the app stores and the smartphone operating systems makes it even more difficult to break up the monopolies.

Not that GAFA doesn't deserve our respect for having understood this connection years ago and for placing it at the heart of their strategy to defend what they do and exculpate them from any wrongdoing. None of us can imagine a world without their dominance anymore. And therein lies the problem, as well as the political and regulatory challenge.

The large tech corporations spend billions on image campaigns and employ an army of lobbyists in Berlin, Brussels and Washington, DC. There is hardly a single association, NGO, or start-up hub in the political sphere that they do not support in some way or other. Politicians and entrepreneurs who point out alternatives are – with only few exceptions – reeled in again by the lobbyists and opinion-makers, attacked in the media or have their businesses damaged.

The lack of regulation of the tech monopolies is the greatest danger to liberal democracies

There are two possible exit scenarios for the end of the tech monopolies. If we continue without strict regulation of GAFA, the polarization within society will grow. The economic opportunities of smaller businesses will shrink more and more because of the growing profits of the monopolies, and the GAFA companies will be able to take over whole new fields of business.

An ever-growing concentration of economic power will lead in the medium term to an erosion of the market economy. This will soon cause social unrest, distribution struggles, and an increasing destabilization of our liberal democracy, which will be unable to gain control of the economic inequality.

There have been clear signs of this development for many years, but our debate so far remains focused on the symptoms instead of asking about the causes. This is regrettable, because the facts are there for anyone to see. You only have to look at where record turnovers are being made in the midst of a great European recession, where profits are increasing permanently, and market shares are being gained.

And in my opinion, this dark scenario can only end in a state run by autocratic populists. We only have to look at protectionist China to see where this might lead. The GAFA companies have nothing to say there anymore, and the Chinese have constructed their own tech ecosystem. The Chinese population pays a very high price for this in the form of an illiberal dictatorship in which politicians decide who is allowed to have economic success.

#### The EU is uniquely positioned to resolve climate change---BUT requires internal coherence.

Sebastian Oberthür & Claire Dupont 21, Professor, Environment & Sustainable Development, Institute for European Studies; Assistant Professor, European & International Governance, University of Gent, "The European Union’s International Climate Leadership: Towards a Grand Climate Strategy?" Journal of European Public Policy, Vol. 28, Issue 7, pg. 1108-1110, 2021, T&F.

Our approach has enabled us to systematically assess the EU’s leadership record on climate change against a comprehensive set of criteria. It is based on the systematic distinction of different facets of exemplary leadership (credibility, policy learning and diffusion, market and regulatory power) and diplomatic leadership (coherence/unity, adaptation to international context), as well as the identification of key trends in international climate governance (polycentricity, multipolarity, geopolitics). This has allowed us to pinpoint key achievements of, and remaining challenges to, EU climate leadership.

This EU leadership has registered important achievements in (1) mobilizing its capacities and (2) adapting to external conditions over the past decades. The EU successfully plugged an early credibility gap by strongly developing its domestic climate policy from the 2000s, as most recently reinforced by the European Green Deal and the response to the Covid-19 crisis. Putting GHG emissions on a downward path, this domestic policy framework has also radiated beyond the EU that especially promoted key instruments such as emissions trading. Reformed arrangements for EU ‘actorness’ in the multilateral UN process in the 2000s have rendered the EU a more efficient and effective negotiator. Also, foreign services have been increasingly involved and coordinated across the EU for enhanced climate diplomacy. The EU has furthermore significantly adapted its international approach to the evolving turbulent context. This has included its new ‘leadiator’ strategy in response to the rise of emerging powers and resulting stronger multipolarity, including enhanced targeting of climate support. It has also entailed a stronger engagement with multiple international and transnational fora and initiatives beyond the UN process as well as in bilateral contexts in response to growing polycentricity. Notably, the EU has adapted its international approach in reaction to particular problems (inefficiencies in international negotiations; the EU’s Copenhagen failure).

A consequential first challenge for EU climate leadership going forward is maintaining and further expanding the aforementioned achievements. Coordination of climate diplomacy across foreign and climate policy, arrangements for the EU’s pursuit of multilateral climate negotiations, and advancing effective ‘polycentric’ engagement beyond it all require continued efforts and regular reflection. Effective ‘leadiatorship’ requires regular review of positions and opportunities/needs for coalition and bridge-building and well-considered deployment of support. Not least, the EU can – as foreseen in the European Green Deal – further develop its domestic climate policy to advance the EU’s regulatory and market potentials and to boost its own low-emission capabilities for maximum international influence.

Our approach leads us to another, interrelated major challenge towards the further development of the EU’s leadership capabilities, thereby advancing analysis of the strategic qualities of EU environmental foreign policy (Schunz, 2019). The geopolitical turn of climate policy impels the EU to embed its climate ‘leadiatorship’ into a broader ‘grand climate strategy’. Developing the EU’s more strategic approach that emerged in the 2010s (ibid.) into a full-fledged joint grand climate strategy would entail systematic, integrative and continuous consideration of climate strategy at highest political level, including the interplay between exemplary and diplomatic leadership across policy fields and fora. Given the particular character of the EU, such an integrated and comprehensive approach demands coordination between EU institutions and member states. The European Green Deal contains relevant elements (e.g., integration of climate objectives across policy fields), but as a Commission programme, it does not (yet) provide for the required coordination across the EU.

In this context, our analysis of past EU climate leadership highlights the strong potential of better integrating domestic and external climate policy development. EU climate diplomacy has helped diffuse domestic policies internationally and has significantly relied on the credibility of domestic action, as well as on EU regulatory and market power. In turn, EU leadership aspirations have been an important driver of domestic policy development. Hence, the distinction between EU exemplary and diplomatic leadership becomes blurred and the separation of EU domestic and international climate policy – reflected both in policymaking and related academic literature – elusive. Increasing the coherence and integration of internal and external climate policies remains a major challenge and source of untapped potential, which an EU grand climate strategy could aim to address.

#### Empirics prove---populism, regardless of affiliation, weakens climate action.

Aron Buzogány & Christoph Mohamad-Klotzbach 21, Institute of Forest, Environmental, and Natural Resource Policy, University of Natural Resources and Life Sciences Vienna; Institute of Political Science and Sociology, University of Würzburg, "Populism and Nature—The Nature of Populism: New Perspectives on the Relationship between Populism, Climate Change, and Nature Protection," Zeitschrift FüR Vergleichende Politikwissenschaft, Vol. 15, 08/24/2021, Springer.

Two existential challenges have emerged for contemporary democracy over the last decade: populism and climate change. Since the 2008–2009 Global Financial Crisis populist parties of different feathers have gained vote shares in many parts of the world, including Western and Eastern Europe, the United States, and Latin America (Manow 2018). During the same decade, climate change has become a salient political issue due to failures to find common solutions towards mitigation on the global level and increasingly frequent extreme weather events. These failures have led to the strengthening of political forces calling for more climate action and environmental protection and towards the end of the 2010s, the emergence of the “new climate justice movement” and massive climate protests by Fridays for Future, Extinction Rebellion, and the German anti-coal movement Ende Gelände has made climate and environment one of the most important political discourses in Western democracies (Evensen 2019; Fischer 2017, 2019; Marquardt 2020; Scherhaufer et al. 2021; Kocyba et al. 2019).

Both the rise of populism across Europe and the increasing salience of climate policy in political debates have been widely acknowledged in various fields of political science. However, the interlinkages between these two fields are rarely addressed comprehensively (but see Lockwood 2018; Radtke et al. 2019). Populist and/or radical right leaders, parties, and movements around the globe often mobilize against far-reaching climate policies and sustainability transitions. They either deny man-made climate change altogether (e.g. Trump in the United States) or prioritize economic issues before ecological ones (e.g. Bolsonaro in Brazil). While attacking the most important components of the liberal order, including democratic institutions, multilateralism, and the values associated with them, such as civil rights, pacifism, feminism, or environmentalism, populists have emerged as important forces defending the incumbent “imperial lifestyle” (Brand and Wissen 2017; Eversberg 2018) that characterizes Western political systems. At the same time, contemporary right-wing populism is often associated with science skepticism and stands in strong contrast to the markedly (and rather uncritically) pro-science attitude of the new climate movement (Evensen 2019; Zulianello and Ceccobelli 2020). The growing polarization on the epistemic dimension of climate change and environmental policies is reflected also in increasing partisan polarization and backlashes against sustainability transitions by popular movements, such as the Gilets Jaunes in France (Kinniburgh 2019). Taken together, these developments suggest that public support for far-reaching climate and environmental policies is not as uncontested as pro-climate mobilization discourse suggests and that social aspects and participation related to these policies must be considered (Lockwood 2018). The tensions between populism and climate-related policies are reflected in three main discussions in the literature, which we briefly introduce below (for a more detailed version of the next section, see Buzogány and Mohamad-Klotzbach 2021).

Varieties of climate and environmental populisms

The first discussion concerns the empirical variety of populism and its relation to nature and climate change. How different populist movements incorporate nature and integrate climate change-related issues into their discourses and policies is historically predetermined and has much to do with traditions of regional resource conflicts regarding land use, agriculture, or industrialization. Historically-informed work in this tradition shows how, in the case of early populist movements such as the Russian narodniks or US-American agrarian populists, opposition to ruling elites and populists’ classic Manichean “us vs. them” dichotomy has also influenced populists’ relation to nature. A particularly rich scholarship on far-right parties shows that these are often supportive of “green patriotism” and environmental conservation when this concerns the protection of “homelands” (Forchtner 2019b; François and Nonjon 2021). Nature protection-related discourses often emerged in parallel to nation-building processes of the 19th century, showing a strong association between the people and the landscapes they inhabit. These connections remain relevant today, as overlaps often exist between right-wing ecology and populist identity politics, including inheritances from agrarian populism and a rural voter base (Forchtner 22,23,a, b; Vihma et al. 2021).

While most of the current literature focuses on right-wing populists, the relationship between populism and environmentalism also has a left-wing heritage. This concerns not only branches of agrarian populism or the early environmental movement in the United States (Meyer 2008). More recently, activists and scholars such as Chantal Mouffe (2020) have advocated that the new climate movement should embrace pro-environmental populism both to counter anti-environmental populism and to broaden democracy (Beeson 2019; Bosworth 2020; Davies 2020). While disagreeing on the content, left and right-wing populist environmental discourses often overlap when it comes to anti-elitism, the critique of the establishment, and Manichean worldviews. At the same time, whether left-wing or right-wing discourses of “resource populism” (Lyall and Valdivia 2019) predominate is determined not only by national heritages but also by the embeddedness of different regions into the world economy.

Another aspect that contributes a further facet to the diverse landscape of populists’ environmental and climate-related discourses concerns the difference between the themes they focus on. While nature protection has an important historical heritage on the radical right, this often goes together with opposition to climate policies that are usually opposed because of their internationalist cross-boundary character or the threat they imply to traditional lifestyles (Schaller and Carius 2019; Tosun and Debus 2021; Vihma et al. 2021). Relating this perspective back to classic readings in the study of populism, one can see that climate policies affect populists’ “heartland” (Taggart 2000) in a more intrusive way than nature protection policies ever have. But even if “climate skepticism” or the denial of man-made climate change is undisputedly one of the most common positions of the “far right” (Lamb et al. 2020), there are also more ambivalent and even some affirmative positions due to populist radical right-wing parties’ need to rejuvenate and attract younger voters (Ruser and Machin 2019). This raises the question about the ideological underpinnings, or alternatively, the strategic considerations populist parties and their leaders make concerning their stances on climate change and environmental policy.

Drivers of climate and environmental populism

A second discussion thus concerns the drivers of populist parties’ environmental and climate change-related discourses. The scholarly debate hitherto differentiates between three main explanations. According to the interest-based account, economic challenges related to climate change and sustainability transitions lead to substantial and often painful structural changes in core industrial sectors and reinforce inequalities (Lockwood 2018). The Gilets Jaunes protests in France by the economically and geographically marginalized part of the population illustrate the importance of grievances related to social justice for those feeling “left-behind” on climate policy and sustainability transitions (Kinniburgh 2019).

An alternative explanation relates to (right-wing) populism’s anti-elite, socially conservative, and nationalist values and their aversion to globalism, liberalism, and the loss of national sovereignty (Lockwood 2018). This ideology-based account explains populist parties’ opposition to energy transitions as a conflict between the “detached urban elites” and the “ordinary rural citizens” (Eichenauer et al. 2018, p. 645; Fraune and Knodt 2018; Selk et al. 2019). Ruling elites are often thought to undermine sovereignty and propagate postmaterialist values that collide with national and local interests. Here populists offer protection, be it against foreign-minded neoliberal elites or immigrants.

Finally, a third perspective on the driving forces behind environmental populism points to the strategic nature of populists’ framing of issues. Climate policies, which are often delegated to experts, bureaucrats, or international bodies, can easily be exploited by populists to attract new voters or to react to the increasing success of competitor parties. At the same time, populist parties are particularly responsive to popular moods and might also easily adapt to climate change discourse if that seems a strategically convenient alternative, for instance when it comes to maximizing votes.

Effects of climate and environmental populism

The third discussion concerns the impact of populist parties on climate and environmental policies. A number of recent studies have analyzed both the output and the outcome dimensions, but this remains a less explored yet promising field of research (Huber et al. 2021; Petri and Biedenkopf 2021; von Homeyer et al. 2021). The few studies that focused on the output dimension of policies propagated by populist parties show that, against expectations, energy and climate policy efforts do not always weaken when populists are strong. Rather, and indirectly, populist electoral successes can contribute to policy improvements because the polarization of the political arena strengthens the government influence of smaller progressive parties (Ćetković and Hagemann 2020). Huber et al. (2021) find substantial differences in the climate change policy discourses, positions, and actions across and within different populist parties. The authors suggest that these differences are explained by left-right cleavages of host-ideologies, but not by populism per se. While left-wing governments tend to produce more climate policies on the aggregate level (Schulze 2021), Latin American case studies show that left-wing populist governments rarely break with extractivism and resource exploitation even if they campaigned against these policies initially (Andreucci 2018; Lyall and Valdivia 2019). When it comes to policy outcomes, research by Detlef Jahn (2021) argues that populists in government office produce more greenhouse gas emissions and are also linked to lower environmental performance in general (Böhmelt 2021).

#### Unchecked warming causes extinction.

Simon Beard 21, Centre for the Study of Existential Risk, University of Cambridge, Lauren Holt, Centre for the Study of Existential Risk, University of Cambridge, Asaf Tzachor, Centre for the Study of Existential Risk and Cambridge Global Food Security Interdisciplinary Research Centre, University of Cambridge, Luke Kemp, Australia National University, Shahar Avin, Centre for the Study of Existential Risk, University of Cambridge, Phil Torres, Institute of Philosophy, Leibniz University Hannover, Haydn Belfield, Leverhulme Centre for the Future of Intelligence, University of Cambridge, “Assessing Climate Change’s Contribution to Global Catastrophic Risk”, Futures, Volume 127, March 2021, ScienceDirect

As part of their discussion of BRIHN Baum and Handoh (2014) note that climate change is the planetary boundary for which the risk to humanity has received most meaningful consideration and they suggest that this attention is deserved. Yet little research attention has been paid to climate change’s extreme or catastrophic effects. Kareiva and Carranza (2018) argue that, despite currently falling outside of the area of high risk, climate change has the clear potential to push humanity across a threshold of irreversible loss by “changing major ocean circulation patterns, causing massive sea-level rise, and increasing the frequency and severity of extreme events… that displace people, and ruin economies.” Even if humanity was resilient to each of these individual impacts, a global catastrophe could occur if these impacts were to occur rapidly and simultaneously.

One scenario that has received comparatively more attention is that of the global climate crossing a tipping point that would trigger environmental feedback loops (such as declining albedo from melting ice or the release of methane from clathrates) and cascading effects (such a shifting rainfall patterns that trigger desertification and soil erosion). After this point, anthropogenic activity may cease to be the main driver of climate change, making it accelerate and become harder to stop (King et al., 2015).

Other scenarios can be discerned from the numerous historical cases in which the modest, usually regional, climatic changes experienced during the Holocene have been implicated in the collapse of previous societies, including the Anasazi, the Tiwanaku, the Akkadians, the Western Roman Empire, the lowland Maya, and dozens of others (Diamond, 2005, Fagan, 2008). These provide a precedent for how a changing climate can trigger or contribute to societal breakdown. At present, our understanding of this phenomena is limited, and the IPCC has labelled its findings as “low confidence” due to a lack of understanding of cause and effect and restrictions in historical data (Klein et al., 2014). Further study and cooperation between archaeologists, historians, climate scientists and global catastrophic risk scholars could overcome some of these limitations by identifying how the impacts of climate change translate into social transformation and collapse, and hence what the impacts of more rapid and extreme climatic changes might be. There is also the potential for larger studies into how global climate variations have coincided with collapse and violence at the regional level (Zhang, Chiyung, Chusheng, Yuanqing, & Fung, 2005; Zhang et al., 2006). However, these need to be interpreted and generalized with care given the differences between pre-industrial and modern societies.

Societies also have a long history of adapting to, and recovering from, climate change induced collapses (McAnany and Yoffee, 2009). However, there are two reasons to be sceptical that such resilience can be easily extrapolated into the future. First, the relatively stable context of the Holocene, with well-functioning, resilient ecosystems, has greatly assisted recovery, while anthropogenic climate change is more rapid, pervasive, global, and severe. Large-scale states did not emerge until the onset of the Holocene (Richerson, Boyd, & Bettinger, 2001), and societies have since remained in a surprisingly narrow climatic niche of roughly 15 mean annual average temperature (Xu, Kohler, Lenton, Svenning, & Scheffer, 2020). A return to agrarian or hunter-gatherer lifestyles could thus have more devastating and long-lasting effects in a world of rapid climate change and ecological disruption (Gowdy, 2020).7 Second, modern human societies may have developed hidden fragilities that amplify the shocks posed by climate change (Mannheim 2020) and the complex, tightly-coupled and interdependent nature of our socio-economic systems makes it more likely that the failure of a few key states or industries due to climate change could cascade into a global collapse (Kemp, 2019).

A third set of plausible scenarios stem from climate change’s broader environmental impacts. Apart from being a planetary boundary of its own, Steffen et al. (2015) point out that climate change is intimately connected with other planetary boundaries (see Table 1). Climate change is thus identified by the authors as one of two ‘core' boundaries with the potential “to drive the Earth system into a new state should they be substantially and persistently transgressed.” This transformative potential was elaborated on in subsequent work exploring how the world could be pushed towards a ‘Hothouse Earth’ state, even with anthropogenic temperature rises as low as 2 °C (Steffen et al., 2018).

The connection between climate change and biosphere integrity (the survival of complex adaptive ecosystems supporting diverse forms of life) is particularly strong. The IPCC is highly confident that climate change is adversely impacting terrestrial ecosystems, contributing to desertification and land degradation in many areas and changing the range, abundance and seasonality of many plant and animal species (Arneth et al., 2019). Similarly, the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) has reported that climate change is restricting the range of nearly half the world’s threatened mammal species and a quarter of threatened birds, with marine, coastal, and arctic ecosystems worst affected (Diaz et al., 2019). According to one estimate, climate change could cause 15–37 % of all species to become ‘committed to extinction’ by mid-century (Thomas et al., 2004).

Disruption to biosphere integrity can have profound economic and social repercussions, ranging from loss of ecosystem services and natural resources to the destruction of traditional knowledge and livelihoods. For instance, desertification, which threatens a quarter of Earth’s land area and a fifth of the population, is already estimated to cost developing nations 4–8 % of their GDP (United Nations, 2011). Many other rapid regime shifts involving loss of biosphere integrity have been observed, including shifts in arid vegetation, freshwater eutrophication, and the collapse of fish populations (Amano et al. 2020). There is a theoretical possibility of still more profound regime shifts at the global level (Rocha, Peterson, Bodin, & Levin, 2018). However, the contribution of loss of biosphere integrity to GCR is yet to be assessed. Kareiva and Carranza (2018) argue that it is unlikely to threaten human civilization, due both to a lack of plausible mechanisms for this threat and the fact that “local and regional biodiversity is often staying the same because species from elsewhere replace local losses.” However, in their classification of GCRs, Avin et al. (2018) suggest the potential for ecological collapse to threaten the safety boundaries of multiple critical systems with diverse spread mechanisms at a range of scales, from the biogeochemical and anatomical to the ecological and sociotechnological. Note that both these studies were conducted for largely conceptual purposes and should not be taken as rigorous analyses of this risk, this topic warrants further investigation.

3.2. Classifying climate change’s contributions to global catastrophic risk

Climate change's contribution to GCR goes well beyond its impact on the earth system. Taking Avin et al.’s list of critical systems, we note that previous studies have mostly focused on the effects of climate change on physical and biogeochemical systems (e.g. global temperature and sea-level rise) or the lower-level critical systems that are most directly related to human health and survival (e.g. Heath Stress). However, these represent a very limited assessment of risk as it only accounts for climate change as a direct hazard/threat and our "ontological" vulnerabilities to it. A more comprehensive risk assessment must consider the higher-order critical systems threatened by climate change passively (through a lack of alternatives) and actively (through intentional design).

The probability of a global catastrophe is higher when sociotechnological and environmental systems are tightly coupled, creating a potential for reinforcing feedback loops. If environmental change produces social changes that perpetuate further environmental change, then this could actively work against our efforts at adaptation. When this change has the potential to produce significant harm, via human vulnerabilities and exposure, we describe such loops as ‘global systems death spirals.’ These spirals could produce self-perpetuating catastrophes, whereby the energy and resources required to reverse or adapt to collapse are beyond the means of dwindling human societies. Feedback loops like this could thus create tipping points beyond which returning to anything like present conditions would become extremely difficult. Global systems would shift to very different states in which the prospects for humanity would likely be bleaker.

### 1AC---Solvency

#### Contention 3 is Solvency.

#### The plan inaugurates a revitalized presumption against the technology sector with regard to mergers.

Mark Glick et al. 21, Professor of Economics, University of Utah; Catherine Ruetschlin, Assistant Professor of Economics, University of Utah; Darren Bush, Leonard B. Rosenberg Professor of Law, University of Houston Law Center, "Big Tech's Buying Spree and the Failed Ideology of Competition Law," Hastings Law Review, Vol. 72, No. 465, 2021, Lexis.

Introduction

Big Tech dominates the technology sector in the American economy. Five technology firms - Google, Amazon, Apple, Facebook, and Microsoft - claim the top five spots on the NASDAQ by market capitalization. And Big Tech is hungry for more. All five companies are buying smaller companies at an unprecedented pace. Google has acquired 270 companies since 2001, including Android, YouTube, and Waze. Microsoft has made over 100 acquisitions in the last ten years, including acquisitions of Skype, Nokia Devices, LinkedIn and GitHub. Amazon has made a similar number of acquisitions, including its purchase of Whole Foods. Facebook has acquired ninety companies, mainly startups.

A growing chorus of commentators have argued that Big Tech's appetite for expanding through purchasing other companies provide a potential means for these dominant firms to solidify and protect their dominance. While we do not determine whether any particular merger was anticompetitive, this Article, relying exclusively on public information, joins that chorus but adds a new twist. It argues that existing law of mergers is ill-equipped to address the tech firms' acquisition of startups because of a rule called the "potential competition" doctrine. The potential competition doctrine addresses the effects of an acquisition where one firm is in the market and the other is "waiting in the wings" or on the periphery of the market.

The problem with the potential competition doctrine, we argue, is its extraordinarily high burden of proof. That burden can be traced back to Justice Powell's opinion in United States v. Marine Bancorporation. The Marine Bancorporation case imposed an extravagant evidentiary burden for a violation of § 7 of the Clayton Act based on elimination of potential competition. Decades later, that standard has gutted the proper role of competition law and rendered it effectively inapplicable to today's mergers in digital markets. A dramatic rethinking of the doctrine is needed to enable federal antitrust enforcement agencies to protect consumers.

In this Article, we explore how the proper use of potential competition doctrine might have halted the transactions that have led to massive Big Tech. We begin by examining the history of Facebook's acquisition strategy and how [\*468] it could have contributed to Facebook's rise to dominance and the maintenance of its dominance.

Facebook and other Big Tech companies maintain their market dominance by harnessing the network effects that reinforce user value in the consumer-facing market and advertiser benefits in digital advertising markets. Startup firms provide competitive pressure because they are able to siphon off or "cream skim" customers and collect valuable data. Big Tech acquisition of startup companies may benefit the incumbent by reducing competitive pressure of potential entrants on the periphery of the market or by preventing future entry and expansion by such firms that could undermine the incumbent's dominance.

Such acquisitions are typically analyzed under the potential competition doctrine. In the next Part, we discuss how the Court transformed a once workable standard into a completely unworkable, open-ended prediction of future conduct and performance that could not be practically discharged. We discuss how the Court split the doctrine in two, creating the actual potential competition doctrine and the perceived potential competition doctrine, each with different evidentiary requirements. It ultimately expressed disdain over one of the doctrines it created, suggesting that no plaintiff could meet such a standard.

We then discuss, using public information, the competition harm story of Facebook's acquisitions of Instagram and WhatsApp. In each Part, we detail why antitrust enforcement agencies failed to challenge mergers. We then describe why the potential competition doctrine as currently applied would lead to a false negative; namely, an acquisition that is competitively harmful yet not challenged by federal antitrust enforcement agencies. The high initial burden on the plaintiff to present a case concerning future conduct and competitive effects serves as a serious deterrent to potential competition mergers, even by dominant firms.

In the next Part, we seek to alter the potential competition doctrine. Using the 1968 Merger Guidelines and additions from the potential competition literature, we assert that with simple structural presumption, the Federal Trade Commission (FTC) could have elected to challenge these mergers and shifted the burden to Facebook to demonstrate why no harm to future competition could occur, and why, given Facebook's resources it could not internally innovate to achieve its competitive goals.

I. Facebook's History of Acquisitions of Small Potential Competitors

Big Tech firms operate in online platform markets where they provide critical facilitation services between buyers and sellers, users and content providers, and advertisers and consumers. Their services include search [\*469] engines, social networks, ecommerce, digital advertising, app stores, and operating systems, where platforms connect parties online to facilitate transactions. The increased functionality and speed of the internet has made platforms exceptionally efficient in connecting end users. The tremendous profits earned by these firms create strong incentives for others to enter these markets, yet two or fewer Big Tech firms have dominated many of these markets for years. Some observers contend that the Big Tech large-scale acquisition programs have diluted the natural process of competitive entry, with firms entering the market with the sole intent of being acquired, as there would be no other plausible endgame.

Online platforms typically operate in two-sided markets including a consumer-facing market for digital services and a market for online advertising. In order for a platform to maintain its position in both the digital services and the online advertising markets, it must maintain the most desirable platform for users and prevent users from switching to other platforms. In other words, user traffic is important to both markets because they each exhibit strong network effects. In social networking, for example, users value the social network with the most opportunities to reach others; advertisers benefit from [\*470] greater user numbers in terms of reach and consumer targeting. Such direct and indirect network effects have resulted in Facebook becoming a dominant provider. Once a dominant firm establishes itself in an online platform market, the network effects and data-driven efficiencies in digital markets tend to reinforce dominance even when new rivals improve or produce novel products.

While strong network effects can cause markets to tip and create a dominant firm, they can also allow small nascent competitors with a desirable alternative platform to scale quickly and challenge such dominance. Innovating startup firms provide competitive pressure in such markets when they exhibit rapid user growth and the potential to enter the dominant firm's core market. Prior to entry into the core market, these nascent firms demonstrate their potential by diverting users from the dominant platform or acquiring data that would be valuable on the advertising side of the market. This information provides a signal to the dominant firms, creating an incentive to absorb or eliminate the nascent rival. A nascent competitor can improve the economic performance of the market overall by preventing a dominant firm from reducing quality, raising prices, or curtailing innovation. The nascent startup that blossoms into a competitive rival can reinvigorate the competitive process within the dominant firm's core market. In this context, acquisitions of nascent competitors by dominant firms undermine both current and future competition, reinforcing the incumbent's dominance in the face of technological shifts.

[\*471] Facebook's record demonstrates how acquisitions can play a critical role in the rise to dominance and the maintenance of dominance by a Big Tech incumbent. At the time of Facebook's launch in 2004, the social media market was highly competitive, with multiple new social networks emerging each year. Facebook's famed beginnings in a Harvard dorm room filled a new niche in the social networking market. The site opened exclusively to the Harvard community - requiring a Harvard.edu email address to join - before extending services to Stanford, Columbia, and Yale. The interface was simple, providing a few core social networking functions, including profile pages where users could post a single photo and personalized information, as well as a "friend graph" or database of connections between individuals that could be searched via user names or other attributes to identify and request new connections.

The site was immediately popular and each new user added to its overall utility as more friends or potential friends joined the network. Despite its limited Ivy League user base, by December 2004 the site had grown to one million monthly active members. Its popularity drew the attention of funders. Funding drove expansion, first to more universities, then high schools, then workplaces, and finally in September 2006 to anyone in the world. By the time Facebook was opened to all people willing to register, the company had already received more than $ 40 million in angel and venture capital investments. This funding enabled the company to pursue an ambitious growth strategy, including early acquisitions, which made it possible for the company to take advantage of economies of scale and scope and network effects in the social networking market.

Social media use grew rapidly in the years of Facebook's early expansion. According to survey data from the Pew Research Center, just 7% of U.S. adults participated in social networking in 2005. Over the following decade, that number would rise to 65%, with the fastest growth occurring before 2011. Facebook positioned itself to take advantage of this market growth by expanding its user base, articulating a qualitative product differentiation between itself and [\*472] its competitors, and integrating new ways of engaging users into its suite of social networking functions by offering new features and functionalities.

Facebook operated in a rapidly changing competitive environment where the basic technological undergirding of the social network was evolving, including the increasing importance of mobile technology to connect users online. Beginning in 2007, the company initiated a series of acquisitions of both its potential rivals in the social media market and firms in adjacent markets that could divert user engagement away from the social network. This tactic arguably propelled Facebook's growth strategy as the company overtook its main competitors. Figure 1 shows the number of acquisitions Facebook completed each year from 2004 to 2018, as well as the number of monthly active users reported by the company each year.

Figure 1

Partially as a result of Facebook's acquisition strategy, when market user growth leveled off, competitors like MySpace, Windows Live Spaces, and [\*473] Google's Orkut suffered significantly, while the number of new users active on Facebook each year continued to measure in the hundreds of millions. Facebook first surpassed its main rival, MySpace, to become the most popular website in the United States in 2009, just five years after its founding. By 2011, when more than half of all adults and two-thirds of internet users were regular users of social networks, Facebook dominated the industry by a wide margin. Pew Research Center data from 2011 showed that while 92% of social network users regularly accessed Facebook, just 29% utilized the nearest competitor, MySpace, while 18% used LinkedIn and 13% used Twitter.

From 2007 to 2018 Facebook acquired or attempted to acquire more than 100 companies in competing and adjacent markets. The ninety acquisitions completed since the company's founding, and documented in the Appendix, range from small acquisitions like the $ 2.5 million purchase of location services network Nextstop to the $ 19 billion acquisition of popular instant messaging rival WhatsApp in 2014. They include deals that transferred key technology and expertise to the company in markets for app development platforms, instant messaging, photo sharing, location services, user information and surveillance, and advertising and analytics. Many of the acquisitions converted stand-alone apps, websites, and platforms that worked inter-operably across competing [\*474] networks into Facebook-exclusive features. Other products were simply shuttered in the days or months following their acquisition.

Today, Facebook is number three on the list of most-trafficked websites in the world. With Instagram, Messenger, Facebook, and WhatsApp, the company now owns four of the most popular mobile apps in the United States. Facebook is responsible for about ten percent of the mobile browser market, representing a substantial share of mobile users for whom Facebook is the main point of entry for online content. This remarkable influence over how individuals engage and consume online is the product of over a decade of strategic internal growth, as well as the acquisition of potential competitors and the integration of their user traffic and functionality within the Facebook structure.

Remarkably, Facebook's ascendancy in concert with its numerous acquisitions stimulated little interest by the antitrust agencies. A march to dominance, accompanied by numerous acquisitions of potential competitors, puts Facebook's strategy directly within the merger regulatory power of the government through its ability to enforce § 7 of the Clayton Act. Yet, few of [\*475] the acquisitions faced review from antitrust authorities in the United States. In 2012, the FTC conducted a nonpublic investigation of the $ 1 billion Facebook-Instagram merger and did not recommend any further action. In 2014, U.S. regulators cleared Facebook's $ 19 billion acquisition of the messaging application WhatsApp, though the FTC did send both companies a letter reminding them of their obligation to maintain privacy practices in accordance with the WhatsApp user agreement in place at the time that user data was collected.

Unlike many other companies acquired by Facebook, Instagram and WhatsApp remained separate from Facebook's social network in branding until 2019, and in some features of interoperability and data autonomy. They are also globally important market leaders in social networking, photo sharing, and instant messaging. The scale, innovation, and popularity of these products have made them frequent examples of potential competitors both at the times of the acquisitions and in the years since.

The question arises why the federal antitrust enforcement agencies demonstrated reluctance to seriously confront the competitive impact of these and similar mergers among high tech companies. We argue below that the potential competition doctrine, as developed during the years of the influence of the Chicago School of antitrust, has played an important role in insulating acquisitions of startups by the dominant tech companies from the levels of antitrust scrutiny necessary to protect consumers and the competitive process in technology markets.

II. The Potential Competition Doctrine

Facebook and other Big Tech companies maintain their market dominance by harnessing the network effects that reinforce user value in the consumer-facing market and advertiser benefits in digital advertising markets. Innovative startup firms provide competitive pressure in these markets despite the tendency toward tipping when small firms exist that have the potential to rapidly siphon off users to more desirable or innovative platforms, collect valuable data on end users, or both. In this context, the acquisition of startup companies may benefit the dominant firm by reducing the disciplining competitive pressure of potential entrants on the periphery of the market or by preventing future entry and expansion by such firms that could undermine the incumbent's dominance. Under the common law of antitrust, an acquisition of a potential entrant is [\*476] analyzed under the potential competition doctrine. Thus, to understand the ability and potential to regulate acquisitions by dominant tech firms it is important to understand how the law of potential competition mergers developed and why it has been so underutilized to date.

The history of the potential competition doctrine informs the analysis of tech industry acquisitions because it demonstrates how a shift in the standard of analysis beginning in the 1960s and culminating in the 1974 United States v. Marine Bancorporation decision undermined the applicability of the doctrine in a range of contexts including online platform markets. The potential competition doctrine emerged in the aftermath of the 1950 Amendment to § 7 of the Clayton Act. As described by the Supreme Court in Brown Shoe v. United States, the "dominant theme pervading congressional consideration of the 1950 amendments was a fear of what was considered to be a rising tide of economic concentration in the American economy." In 1963, the Supreme Court, in United States v. Philadelphia National Bank, explained that the "intense congressional concern" about increasing concentration "warrants dispensing, in certain cases, with elaborate proof of market structure, market behavior, or probable anticompetitive effects."

Under this standard, expectations of the market-disciplining effects of potential competition operated to preserve competition in cases where the doctrine applied. The Court explained that when there is a structural increase in concentration due to a merger, the merger "is so inherently likely to lessen competition substantially that it must be enjoined in the absence of evidence clearly showing that the merger is not likely to have such anticompetitive effects." Thus, the Court created a presumption of an anticompetitive effect from a structural increase in concentration, placing the burden on the merging parties to refute the presumption. The plaintiff would still be required to define the relevant markets involved and measure market shares and concentration, but a full-blown analysis of the impact of the merger was judged by the Court to be unrealistic and counter to the congressional intent to stem the rising levels of concentration in the United States.

The Court's approach is often referred to as a "structural approach," which is shorthand for the belief that mergers above a certain concentration threshold [\*477] have a reasonable probability of harming competition. The structural approach to merger analysis contrasts to the effects-based approach, which requires a prediction of the future competitive effects of the merger by use of detailed economic analysis. The Philadelphia Bank opinion implicitly rejected the effects-based approach because of its intractability. As the Eighth Circuit later commented, the structural approach is preferable in cases concerning potential competition since "proof of liability under either [potential competition] theory is certain to entail expensive, uncertain litigation, even if, as here, the acquiring firm is rich and powerful and the acquired firm's market highly concentrated." The practical requirements of proving the competitive effects of the threat of entry were deemed nearly insurmountable despite the importance of these effects.

In contrast to the Philadelphia Bank paradigm, later Supreme Court cases developed an unworkable legal standard for the potential competition doctrine. The Court imposed an initial stage open-ended proof requirement involving prediction of future conduct and performance that could not be practically discharged. In developing this standard, the Court divided potential competition into two separate legal doctrines - the actual potential competition doctrine and the perceived potential competition doctrine - with distinct evidentiary requirements. After separating actual and perceived potential competition, the Court twice expressed doubt regarding the viability of the actual potential competition doctrine. In these cases, the Court discussed the actual potential competition doctrine primarily in the context of acquisitions targeting a dominant firm, and not the context relevant to the current Big Tech mergers in which a dominant firm targets a startup.

The Supreme Court first addressed the issue of harm to potential competition from a merger one year after the Philadelphia Bank decision in United States v. El Paso Natural Gas Co. This case provides important insights [\*478] for the viability of the potential competition doctrine to Big Tech mergers since it is the chief example of the doctrine applied to a case where the potential entrant is the target firm. The case involved the merger between two natural gas pipeline companies and their impact on the California market. El Paso Natural Gas was the only supplier of natural gas to California when it attempted to acquire Pacific Northwest. The Court noted that Pacific Northwest had attempted to enter the California market by supplying Canadian natural gas to one of El Paso's customers in Southern California, Southern California Edison Co. The deal fell through only when El Paso agreed to a more favorable contract with its customer. The Court conceived of the potential harm from the merger as the elimination of influence of the potential entrant on El Paso, or the perceived potential competitive impact of Pacific Northwest. Pacific Northwest's threat of entry forced El Paso to act competitively, despite the company's monopoly in the California market. The evidence showed that El Paso did prevent Pacific Northwest's entry by matching and exceeding Pacific Northwest's offer to a California customer. If Pacific Northwest had captured the customer, it would have entered the market. Nevertheless, the Court chose to focus on the current impact of the entry attempt on El Paso's bid, rather than the more significant future impact Pacific Northwest might have had had it become a competitor in the California market. The Supreme Court would follow this emphasis on the impact of perceived potential competition in subsequent cases.

In the same year, the Supreme Court issued an opinion in another potential competition case. In United States v. Penn-Olin Chemical Co., the Court appeared to reject the structural approach of Philadelphia Bank, defaulting to a vague, open-ended analysis. Penn-Olin Chemical Co. involved a joint venture rather than a merger. All joint ventures raise potential competition issues because absent the joint venture one or both of the same companies might enter into the market alone.

In the Court's analysis, the joint venture eliminated a perceived potential entrant, removing the impact of an "aggressive, well equipped and well financed corporation engaged in the same or related lines of commerce waiting anxiously to enter an oligopolistic market" which disciplined the existing competitors. [\*479] Citing the El Paso Natural Gas case, the Court stated that potential competition "is not "susceptible of a ready and precise answer.'" It stated that analysis of the impact of a potential entrant depends on ""the nature or extent of that market and by the nearness of the absorbed company to it, that company's eagerness to enter that market, its resourcefulness, and so on.'"

In Philadelphia Bank, the Court had addressed the comparable complications of predicting the future effects of a horizontal merger by establishing structural judicial guidelines. Now, when addressing a parallel prediction of the impact of a potential competitor, the Court surprisingly defaulted to an ambiguous and open-ended narrative. The Court might be forgiven because it resolved the controversy by remanding the case back to the lower court to consider the perceived potential competition impact of the joint venture, but it did so without clear guidance on how such an analysis should proceed. In so doing, the case set a precedent in which the structural approach to potential competition was set aside in favor of a range of claims and presumptions about the intentions and perceptions of merging firms.

In 1967, the Supreme Court again confronted a potential competition problem in Federal Trade Commission v. Procter & Gamble Co., and moved the doctrine closer to the unworkable effects-based approach deduced from a subjective and imprecise evaluation of competitive conditions. Following the acquisition of Clorox Chemical by Procter & Gamble, the FTC blocked the merger, asserting, among other reasons, that Procter & Gamble was likely to enter the bleach market absent the acquisition. Procter & Gamble was a potential competitor in the market and had already launched an abrasive cleaner that was a differentiated substitute for liquid bleach. Procter & Gamble knew the liquid cleaning business, the customers of Clorox and Procter & Gamble largely overlapped, and the company advertised and merchandised in the same manner as Clorox. All of the factors led the FTC to conclude that the acquisition of Clorox by Procter & Gamble would eliminate a likely entrant into the liquid bleach market. Yet the court of appeals rejected the evidence of the closeness and proximity of the two markets and declared that there was insufficient evidence from the management of Proctor & Gamble that it intended to enter the liquid bleach market.

The Supreme Court reversed the court of appeals, but without offering a helpful analysis of the potential competition issues. The Court abstained from [\*480] analysis of actual potential competition and focused solely on the impact of Procter & Gamble as a restraining perceived potential competitor, even though the Court opined that it was "the most likely entrant" into the liquid bleach market. The Court also found, without explaining its basis, that Procter & Gamble did not face a barrier to entry and that "the number of potential entrants was not so large that the elimination of one would be insignificant." The focus of the court of appeals and the Supreme Court on aspects of competition such as the potential competitor's intention of entry, the likelihood of entry, and the number of potential entrants would support the inclusion of such difficult and even subjective or illusory criteria in the evidentiary standards for potential competition cases.

In 1968, the Department of Justice issued Merger Guidelines. While the Supreme Court was grappling with the early cases involving mergers that harm competition by preventing future entry, the Department of Justice developed a clear policy to protect new entry from mergers by dominant firms. According to the 1968 Merger Guidelines:

Since potential competition (i.e., the threat of entry, either through internal expansion or through acquisition and expansion of a small firm, by firms not already or only marginally in the market) may often be the most significant competitive limitation on the exercise of market power by leading firms, as well as the most likely source of additional actual competition, the Department will ordinarily challenge any merger between one of the most likely entrants in the market [and a firm with a large share of the relevant market.]

The acquiring or target firm must be one with the ability and incentive to enter and must be "one of the most likely potential entrants into the market." As discussed in a later Part of this Article, the 1968 Merger Guidelines faltered when addressing the evidentiary burden required to show that a target is one of the most likely potential entrants.

The 1968 Merger Guidelines' explanation of the required evidence to demonstrate potential entry is not a model of clarity. It requires that the Department of Justice marshal evidence demonstrating that entry by the firm [\*481] would be more profitable and less risky than other unidentified non-litigant third-party firms. In 1984, the Department of Justice would give more structure to this inquiry but would continue to require unworkable conduct and performance evidence that would make the potential competition analysis impractical and infrequent.

More clarity emerged from the Supreme Court's 1973 opinion in United States v. Falstaff Brewing Corp. The case involved the acquisition of Narragansett Brewing by Falstaff. Narragansett produced beer sold in the New England regional geographic market. Falstaff sold beer in thirty-two states and was the largest beer producer not in the New England market. The district court considered both the theory that Falstaff disciplined competition as a potential entrant and that Falstaff was a future actual entrant into New England. The district court held that evidence from Falstaff's management cast doubt on whether Falstaff was going to enter the New England market and that competition had not decreased since the consummated acquisition. Again, despite acknowledging the pertinence of the actual potential competition doctrine, in their decision the Supreme Court focused solely on the perceived potential competition aspect of the situation in which the merger "eliminates a potential competitor exercising present influence on the market." The district court erred by assuming that the subjective evidence from Falstaff's management meant that, as a matter of fact, Falstaff was not a potential entrant. Instead, the district court should have considered the objective evidence.

If the district court's approach had prevailed, it would have meant that plaintiffs asserting potential competition cases could be defeated by the uncontroverted testimony of the management of one of the merging entities. Instead, the Court thought that the proper inquiry was whether a rational incumbent firm would have perceived the acquirer as a likely entrant. It stated that "if it would appear to rational beer merchants in New England that Falstaff [\*482] might well build a new brewery to supply the northeastern market then its entry by merger becomes suspect under § 7." However, the Court does not inform us concerning what "economic facts about Falstaff and the New England market" should have been analyzed or what objective evidence should be consulted in order to ascertain the beliefs of a rational beer merchant. It appears that a complex, open-ended inquiry of this nature would lead to an unmanageable problem for a court. For an actual potential competition case, the Court offered even less, declining to even hold that a merger that prevents actual entry violates § 7 of the Clayton Act.

The Court's reluctance is puzzling. As described by Joseph Brodley, the Court has ample scope to apply and interpret the actual potential competition doctrine in both law and precedent. In early Supreme Court cases, the Sherman Act has been held to cover actual potential competition, and the Clayton Act "is an incipiency statute designed to prevent [mergers] that are beyond the scope of the Sherman Act."

The last and most influential Supreme Court case addressing the potential competition doctrine is United States v. Marine Bancorporation, Inc. The 1974 opinion, penned by Justice Powell, established the extraordinarily high requirements of proof that inoculate potentially anticompetitive mergers from scrutiny under the potential competition doctrine today. The case concerned the acquisition by Marine Bancorporation, a large Seattle-based bank, of the Washington Trust Bank, a smaller bank headquartered in Spokane, Washington. The government challenged the merger on both perceived and actual potential competition grounds. It argued that Marine Bancorporation's presence on the fringe of the Spokane market disciplined Spokane competitors, and that absent the merger, Marine Bancorporation would likely enter the Spokane market.

[\*483] The district court found against the government because Washington's state banking regulations prevented the kind of entry the government's theories predicted. The Supreme Court affirmed, but this time took the opportunity to develop a general methodology for analyzing actual potential competition mergers. According to the Court, "two essential preconditions must exist" before an actual potential competition theory "establishes a violation of § 7." First, that the potential competitor could enter the market at issue absent the merger. Second, that such entry would produce a likelihood of deconcentration or other significant procompetitive effects. Moreover, with respect to the first prong, the Court implied that "unequivocal proof" of actual future de novo entry is required. The standard of proof for the second prong is also exacting. The potential entry must accomplish more than simply increased competitive rivalry. It must deconcentrate the market or accomplish another "significant" but unspecified procompetitive transformation. Moreover, the Court expressed doubt that an actual potential competition case would be viable, even when these exacting standards are met. Because the government did not meet its burden regarding Marine Bancorporation, the Court would "express no view on the appropriate resolution of the question reserved in Falstaff."

Lower court interpretations of the binding precedent set forth in Marine Bancorporation demonstrate both the unworkable nature of the proof requirements and the difficulties attendant to requiring the judiciary to grapple with complicated conduct and performance predictions. For example, a few years after the Marine Bancorporation decision, the Fourth Circuit considered a potential competition claim by the FTC in 1977 in Federal Trade Commission v. Atlantic Richfield Co. The case involved the acquisition of Anaconda, a copper and aluminum mining and processing company, by ARCO, a large oil and petroleum company. The FTC claimed that ARCO was a likely entrant into the copper market. The Court interpreted Supreme Court precedent to require "clear proof" of entry (citing to the Marine Bancorporation standard of [\*484] "unequivocal proof"). The Court then relied on the testimony of ARCO's management. This is precisely the type of evidence eschewed by Falstaff. The Court found that "Arco would never seriously consider original entry or entry by toehold acquisition." Lack of proof of entry also doomed the government's cases in British Oxygen Co. International v. Federal Trade Commission, Tenneco, Inc. v. Federal Trade Commission, United States v. Siemens Corp., and Fraser v. Major League Soccer.

The Fifth Circuit, in Mercantile Texas Corp. v. Board of Governors of the Federal Reserve System, set forth a detailed analysis of its understanding of the proof requirements of an actual potential competition violation of the Clayton Act. According to that court, the required elements are: (1) a concentrated market; (2) no other potential entrants exist other than the target (or acquirer); (3) probability of procompetitive entry; and (4) procompetitive effects of independent entry. The court stated that when there are several potential entrants, the elimination of any one entrant would not be significant. It then added, following Richard Posner, that "economic theory suggests that, where oligopoly profits are available, a multitude of firms will eagerly seek to enter the market." Thus, the proponent of an actual potential competition case must show in the Fifth Circuit, contrary to the general case, that the specific facts at issue suggest that only the target (or acquiring) firm is a likely entrant. Thus, the court found that the plaintiff failed to demonstrate that the actual potential competition was "significant" because of the presence of other unanalyzed [\*485] potential entrants and that there was insufficient evidence that entry would have had a "significant" procompetitive effect.

The Department of Justice addressed the potential competition issue again in the 1982 Merger Guidelines drafted by appointees of Ronald Reagan, who were heavily influenced by the Chicago School of Economics. They were revised in 1984, and this was the last time potential competition mergers are addressed by the Merger Guidelines. The 1984 Merger Guidelines built upon but also significantly revised the Department of Justice's position developed in the 1968 Merger Guidelines. The 1984 Merger Guidelines treated perceived and actual potential competition together, thus implicitly rejecting the artificial division made by the Supreme Court. The Department of Justice considered four factors. First, the acquired firm's market must be concentrated, above 1800 HHI. Second, the acquiring firm must have specific entry advantages; otherwise, the elimination of the target still leaves many potential entrants. The number of firms likely to enter should be less than three. If there are more than three likely entrants then there must be direct evidence of likely entry. Third, the target must have a larger market share of twenty percent or more to make a challenge likely. Fourth, the 1984 Merger Guidelines required an analysis of the efficiencies of the proposed merger.

The 1984 Merger Guidelines were both a step forward and a step back from the 1968 Merger Guidelines. Unlike the 1968 Merger Guidelines, the 1984 version assumed that the acquiring firm is the potential entrant. The Department of Justice should have made clear that the potential competition doctrine can be applied in either direction; a merger can prevent entry by the acquiring firm or the acquired firm. The 1984 Merger Guidelines further provide that where entry is easy no merger challenge will be undertaken. This is a step backward from the 1968 Merger Guidelines. The 1984 Merger Guidelines never define ease of entry. At most, the 1984 Merger Guidelines declared that ease of entry is the likelihood and probable magnitude of entry in response to a small but significant and nontransitory increase in price. While the newer version of the Merger Guidelines added structure to the more opaque 1968 Merger Guidelines, it relied on another undefined concept, "entry advantage." As the antitrust scholar Joseph [\*486] Brodley points out, the most probable market entrant under the analytical structure of the 1984 Merger Guidelines is the firm that would achieve the greatest anticipated return from entry. According to Professor Brodley, "courts lack the expertise to resolve complex and speculative factual issues as to future costs and economic conditions. The cases are bound to be burdensome and expensive, especially when competing experts escalate the subtlety of the analysis."

Professor Brodley is correct. Analysis of entry under the Merger Guidelines requires a fairly sophisticated predictive financial analysis. To require a similar analysis for firms that are not parties to the analysis appears intractable.

Thus, the plaintiff asserting a violation of § 7 of the Clayton Act against a dominant firm in a digital market seeking to acquire a startup based on actual potential competition has a difficult uphill climb. First, many circuits do not recognize a reduction of actual potential competition as a viable theory under § 7 of the Clayton Act. Second, most courts, but not all, have considered the situation where the acquirer is the potential entrant rather than the incumbent, dominant firm. Third, the courts have demanded a high standard of proof for demonstrating that the startup would likely enter the market dominated by the acquirer. Fourth, even where entry is likely, the courts require that the target be uniquely situated to enter and not be one of many potential entrants. Fifth, the courts require proof that the startup's entry will significantly reduce the dominance of the dominant firm in its relevant market. These onerous requirements would deter even the most committed antitrust enforcer or plaintiff.

III. Application of the Potential Competition Doctrine to the Instagram and WhatsApp Mergers

In this Part of the Article, we describe the difficulty of applying the potential competition doctrine to Facebook's widely criticized acquisitions of Instagram and WhatsApp. Our intent is not to demonstrate that these acquisitions were anticompetitive but to show that the potential competition doctrine as presently formulated does not allow for a serious inquiry into tech mergers.

[\*487]

A. The Instagram Acquisition

When Facebook announced its $ 1 billion acquisition of Instagram on April 9, 2012, it was something of an anomaly. Although Facebook had made thirty-one acquisitions up to this point, none approached the price tag paid for Instagram. However, Instagram was different, and the opportunity arose at a critical crossroads for Facebook. On the eve of its May 2012 IPO, Facebook was under great pressure by investors to increase its revenue base. At the same time, the rise of mobile technology and its rapid adoption by consumers created hurdles for Facebook to satisfy these demands.

Two problems confronted the company as an increasing share of users accessed the internet from mobile devices. First, Facebook struggled to reorient its network from a desktop-based platform, and second, it had yet to monetize its mobile user base by incorporating advertising on the limited display area available on mobile screens. As other companies developed mobile-first applications that optimized web access using smartphones, Facebook elected to invest in an HTML5-based multi-platform strategy. On mobile devices, their HTML5 approach was slower and less stable than native iOS and Android applications. At the same time, mobile-native applications with social features such as Instagram and Foursquare were attracting growing user numbers and threatened to draw user engagement away from Facebook precisely when its revenue base was under scrutiny.

Photo sharing had been a key facet of Facebook's user engagement since its introduction on the network. By 2009, Facebook Photos was the largest photo sharing service in the world. In ensuing years as dramatic improvements in smartphone camera features made photo sharing an increasingly mobile-based activity, Facebook struggled to adapt to the shift to mobile technology. At this pivotal juncture, Stanford engineering graduates Kevin Systrom and Mike [\*488] Krieger launched the native iOS photo sharing social network Instagram. On Instagram, users could upload, edit, and share pictures from their iPhones and follow, comment, and like the images posted by others. The app also enabled users to post their Instagram images across social networks, including Facebook and Twitter. But the founders did not aim to be a mere content creator for other social networks. Rather, Systrom and Krieger envisioned their app as a rival to the incumbent social networking giants based on a community united under the premise that "the next network is people interested in sharing life visually." The company was poised to compete in the social networking market.

Within the first week of its October 6, 2010 launch on the Apple App Store, Instagram had garnered 100,000 user downloads. Ten weeks later it had accrued over 1 million registered users. The company quickly attracted the attention of venture capital that would allow it to scale. The firm's initial funding round brought former Facebook VP of Product Management Matt Cohler to Instagram's Board of Directors, who advised the company to pursue growth first without monetization in order to achieve the network effects that would drive advertising revenue later. One month before the company revealed its acquisition, just two and half years after its introduction on the App Store, Instagram founder Kevin Systrom announced that Instagram had reached 27 million registered users and "Facebook-level engagement." In the following weeks, Instagram branched out from iOS to launch on Android and brought in 1 million new users in the first twenty-four hours. When Facebook and [\*489] Instagram announced the acquisition six days after the Android launch, Instagram had over 30 million users and just thirteen employees.

According to Silicon Valley folklore, Zuckerberg invited Systrom to his home on a Saturday. By Monday the billion-dollar deal was done. Observers at the time registered their suspicions that the acquisition was an act of "squashing a potential rival" and pointed to the impending monetization of Instagram as a source of competition that could have driven down prices in online advertising markets. The merger triggered a Hart-Scott-Rodino filing, but ultimately the antitrust agencies took no action. The FTC investigation was nonpublic and enforcers did not disclose the basis for their decision at the time. One likely obstacle was the user price of zero set by Facebook and Instagram for their social networking services, which complicates estimates of markups above the competitive price or estimates of entry in response to a small price increase. In the social networking market, companies compete for user attention. The consumer-facing market generally has a price of zero, with services monetized in the advertising market by selling access to the user attention captured on the social network. Instagram operated in the social networking market and it was encouraging users to defect from Facebook to Instagram, but the competitive dimensions of this market are challenging to measure and interpret since users may participate on both networks and neither network charged for the services involved. Several economists have offered solutions to this problem, including measures of user engagement such as the [\*490] number of users or the amount of time spent on a website. By any reasonable measure, Instagram was already a competitor.

In contrast, advertising markets are not free. Digital advertising market analysts widely acknowledge the dominance of a duopoly in digital advertising composed of Google and Facebook, which jointly claim approximately 60% of total revenue in the market. For Facebook that dominance amounted to $ 16.6 billion in advertising income during the second quarter of 2019 and more than 98% of its total revenue. Facebook's advertising market power is even more significant when compared to similar advertising platforms. For example, during the 2007 investigation of the Google/DoubleClick merger, the FTC determined that search advertising (advertising delivered in response to a consumer search query) should be separated from display advertising (including image, video, rich media, etc., purchased on a webpage). According to the FTC, "the evidence shows that the sale of search advertising does not operate as a significant constraint on the prices or quality of other online advertising sold directly or indirectly by publishers or vice versa."

Today, Facebook leads the market in digital display advertising with a market share of over 40%. Arguably, an even smaller relevant market might exist for advertising on social networks. In 2011 and 2012, as Facebook struggled to monetize its mobile user base, Google and Facebook battled for the top spot, each controlling about 14% of the digital display advertising market in [\*491] 2011 and 15% in 2012. At the time of the merger, the majority of Facebook's revenue came from display advertising. Instagram did not sell advertising at the time of the acquisition, but it had been working directly with brands to support image-oriented ways of connecting companies with users. As the Instagram network grew, more businesses saw it as an important medium to reach consumers. When Instagram was ready for monetization, it would be unlikely to charge users for social networking services in a market where the going price was zero. Once Instagram introduced advertising it would likely compete with Facebook in the digital display advertising market as well as social networking. Instagram was an actual potential entrant in both of these markets. Thus, the Instagram merger presented a classic case of a potential competition merger under § 7 of the Clayton Act.

Although the FTC did not outline the considerations that guided its investigation, in August 2012 the United Kingdom's Office of Fair Trading (OFT) published an outline of its decision to refrain from referring the Instagram acquisition to the Competition Commission. OFT determined that Instagram was a current competitor in social networking services, and that Facebook's large share of the market achieved the threshold for investigation. OFT interpreted Instagram's rapid growth as an indication of low barriers to entry in social networking and photo sharing, concluding that Instagram did not evince a uniquely competitive product such that its acquisition would foreclose competition in either market. OFT considered Instagram as a potential competitor in digital advertising markets, but determined that Facebook's [\*492] competition from Google, Yahoo, and Microsoft dwarfed the potential competitive impact of entry by Instagram. It determined that there was "no realistic prospect that the merger may result in a substantial lessening of competition in the supply of display advertising."

Today, Facebook claims a dominant position in the social networking and online social photo services markets, and market power through the Facebook-Google duopoly over digital advertising. If the antitrust agencies faltered, it was likely because the potential competition doctrine created difficult obstacles for a merger challenge. Consider the following facts of the Instagram merger in light of the required proof under the 1984 Merger Guidelines to justify a Department of Justice challenge.

1. Market Concentration

The 1984 Merger Guidelines state that a challenge is unlikely if concentration in the acquired firm's market is below 1800 HHI. In the case of the Instagram merger, the relevant market to measure concentration would be the acquiring firm's market. Facebook operates in markets for social networking and digital advertising. By 2011, Facebook dominated the social networking industry by a wide margin in terms of user numbers and engagement, but HHI calculations lack defined measures for markets where the user price is zero. A workable measure of concentration is critical for markets like social networking in which the good or service is free. As zero-price markets proliferate, antitrust institutions must adopt new instruments for analysis or risk the amplification of consumer harms. Scholarship on the application of antitrust in these markets suggests that enforcement focus on attention and informational costs or metrics such as "time on site" to indicate the extent of competition for user engagement. Such a measure could have demonstrated [\*493] important implications of a Facebook-Instagram merger for competition in the market.

In the digital advertising market, the Facebook-Google duopoly already controlled 45.5% of revenue in 2011, although the majority of that share was attributable to Google. Narrowing the scope to the display advertising market, the top six firms in 2011 collected approximately 49% of the digital display advertising revenue and the HHI among those six firms amounted to just 546. In the years following the 2012 acquisition of Instagram, the Facebook-Google duopoly consolidated their market power in both the digital advertising and the display advertising markets. By 2018, both markets displayed HHIs of over 1800 and Facebook's share of display advertising revenue in the U.S. market rose to more than 20% - even higher if a more narrow market were defined. Thus, while it is likely that a measure of concentration for the social networking market would have satisfied the first prong of the merger guidelines analysis, the concentration levels measured for the display advertising market concentration levels would not have been sufficient.

2. Conditions of Entry Generally

The Department of Justice will not challenge a potential competition merger if entry into the market is easy. This protocol requires the Department of Justice to demonstrate some difficulty of entry or barriers to entry in the concentrated market. Through 2011, the markets for social networking and digital advertising had been dynamic as firms in these markets competed for dominance. The economies of scale and network effects that typify platform markets represent traditional barriers to entry that would reinforce the incumbency of dominant firms, but Instagram was showing the potential for a nascent competitor to siphon off users and gain market share. Entry into social networking or digital advertising markets was achievable for small and startup firms that operated in any of several adjacent markets if they exhibited the rapid growth in user engagement that would lead to increasing value on both sides of [\*494] the market and if they had access to the funding that would allow the company to scale up.

There is one significant barrier to entry in online platform markets that is unlike the traditional barriers considered in other markets: access to data. A dominant firm with access to broad user data has a significant advantage over new entrants. The data advantage allows a dominant firm to reinforce its market power in three ways. The firm can use data to review and improve user services in the core market and expand user engagement, generating more data. The firm can leverage its data advantage to reach new users through entry into adjacent markets and likewise expand its data access. Finally, the scope and magnitude of consumer data available to a dominant firm allows it to sell high-value, targeted advertising with revenues that may be invested in increasing user engagement and amassing more consumer data. These three advantages create a positive feedback loop for the dominant firm.

The drive to exploit user attention and access to data may translate to gains for consumers who enjoy higher quality services and seemingly individuated advertising. For startups with comparatively little data access, the competitive advantage of large firms' data scale and efficiencies poses a significant barrier to entry. As a result of these advantages, the dominant, consumer-facing platforms also dominate advertising markets - a tendency exemplified in the Facebook-Google duopoly.

Despite these structural barriers, demonstrating the difficulty of entry into the social networking or digital advertising markets presents a challenge. For one thing, the data barrier is specific to online platform markets. For another, competition for user attention forces the dominant firm to compete with platforms and applications operating across a variety of markets. There is no direct substitute for Facebook in the social networking market, but smaller firms offering complementary or adjacent features have the ability to capture user attention that draws engagement and profits away from the network, even if the smaller firm is not competing in social networking. This ability to capture user attention also makes these smaller, adjacent firms potential competitors in digital advertising. Extending consideration to potential competitors in adjacent markets where entry is relatively easy could undermine the government's ability to isolate any impact from the elimination of a single rival.

[\*495]

3. The Target Firm's Entry Advantage

If entry is not easy generally, then the Department of Justice has to show that Instagram had an entry advantage not possessed by three or more firms. For reasons discussed later, the potential for firms to enter social networking or digital advertising markets from a variety of adjacent or complementary markets makes it impossible to identify limits to potential entrants. Isolating the photo sharing market in the case of Instagram provides a good example of this difficulty.

Despite Facebook's dominance in photo sharing, several desktop-based and mobile applications existed at the time. Most of these platforms lacked the social features that distinguished the social networking elements available through Facebook and Instagram. Facebook even purchased several other photo-related services leading up to the Instagram acquisition, including the photo sharing and tagging website Divvyshot in April 2010, the file sharing, messaging, and commenting service Drop.io in October 2010, and video and image recording and editing app developer Digital Staircase in November 2011. In May 2012, after announcing the Instagram acquisition but before it was finalized, Facebook purchased Lightbox.com, a mobile social photo sharing application designed for Android, in the period before Instagram introduced its Android app. While Lightbox had amassed 1.5 million downloads in its first seven months of operation, Instagram's Android launch in April reached 1 million within a week. Facebook purchased and shuttered the Lightbox application, absorbing its employees and pulling the app from the market immediately. Facebook launched its own camera app, Facebook Camera, on May 24, 2012, weeks after announcing its intention to acquire Instagram.

The United Kingdom's OFT decision lists six competing apps in the photo sharing market, including Camera Awesome, Camera +, Flickr, Hipstamatic, Path, and Pixable. Of these services, only Camera+, Hipstamatic, and Camera Awesome included camera applications. Flickr is a photo storage and management tool and Pixable was an aggregator that scraped images from social networks including Facebook, Twitter, and Instagram. Path was a social network conceived as a competitor to Facebook that offered a more private experience, limiting social connections to invite more personal interactions. [\*496] Hipstamatic and Camera+ provided photo taking and editing tools but lacked the social features that distinguished Instagram. In addition, Hipstamatic and Camera Awesome had entered into a partnership with Instagram that streamlined posting photos taken with those apps to Instagram's social network. The OFT's list of competitors illustrates the difficulty of identifying potential entrants in the social networking or digital advertising markets. In online platform markets, new entrants often offer just a subset of the services offered by the dominant provider. Firms like Instagram that gain the popularity and funding to scale become rivals for user attention and potentially rivals for the market over time. Facebook would likely argue that Instagram is just one of many potential entrants into social networking, and that any of the other photo sharing apps could replace the potential competition lost through the Instagram acquisition. Moreover, when consumers multi-home by using several apps at once, entry by multiple firms becomes even more likely.

Facebook named Instagram as an important competitor, but it was not the only competitor. Instagram's entry advantages were the extraordinary user growth rate and venture capital investments that might allow the firm to overcome barriers of scale and data access in the social networking and digital advertising markets. These same advantages gained the attention of Facebook and its buyout proposal.

4. Deconcentration from Instagram Entry

The final criteria for a potential competition claim is for the government to show that Instagram's entry into the social networking or advertising markets would deconcentrate the market or have a significant procompetitive effect. Under the Merger Guidelines, this effect can be established by showing that Instagram had a market share of 5% or more. In 2012, the first year Instagram was included in the Pew Social Media Survey, 12% of adults - and a significantly higher share of young people - used Instagram despite the fact that it was a mobile-only application. There are no attentional measures such as [\*497] time on site available for the period before acquisition, but multi-homing and Instagram's own interoperability would suggest that the company claimed a small share of total social networking users' attention. The market draw for Instagram was its popularity with important demographic groups at a time when Facebook saw reaching young people and their preferred technologies as key to maintaining dominance in the market.

At the time of the Facebook acquisition, Instagram had not entered the digital advertising market and had no advertising revenue. It would be impossible to establish a procompetitive effect of Instagram's entry into the advertising market through the 5% threshold because competition from Instagram lay entirely in the future.

The potential competition challenge by the Department of Justice would have certainly failed under its own guidelines. But consider the post-acquisition information that retrospectively demonstrates how the guidelines produce a false negative result. Since the acquisition was finalized in 2012, Instagram has generated a significant share of user engagement and revenue for Facebook. With Facebook's resources and expertise guiding its evolution, Instagram reached 1 billion monthly active users in June 2018 even as Facebook's own user growth dwindled. According to the Pew Research Center, Instagram trails Facebook as the third-most popular social network in the United States with 37% of adults using the platform in 2019. It is the most-used social network for American teens. Although Facebook does not disclose Instagram's financial details, market analysts estimate that 15% of Facebook's revenues come from advertising on Instagram, a number expected to grow over [\*498] time. In 2019, Instagram launched a checkout feature allowing users to make purchases from within the app and delivering a new source of revenue to its parent company. It is impossible to know if Instagram would have developed into such a powerful position without Facebook's guidance, but it is clear that Facebook's ownership of Instagram allows it to reach a larger user base and achieve greater levels of user engagement and revenue generation than Facebook alone. The economies of scale and scope that characterize online platform markets are simultaneously a source of efficiency gains from the acquisition of Instagram and a barrier to entry reinforcing Facebook's dominance in the social networking market.

The Instagram case shows that the potential competition doctrine must be reformed. Common sense suggests that concentration must be measured either by an alternative metric in markets where goods are offered to the public without charge, such as user engagement, or possibly by the advertising dollars that flow to social networks. As we will discuss in the last Part of this Article, concentration should serve as a structural rebuttable presumption when a dominant firm purchases a potential entrant. Before turning to that issue, we briefly discuss Facebook's acquisition of WhatsApp.

B. The WhatsApp Acquisition

Facebook's $ 19 billion acquisition of WhatsApp was another landmark deal. In 2014, mobile messaging applications were the fastest growing app category in the mobile market as social media evolved to accommodate increasing smartphone usage. Users relied on these applications for far more than text messaging, with a variety of social activities taking place on the apps including voice calling, image and video sharing, and gaming. Five-year-old WhatsApp was already the largest and fastest growing of these applications worldwide. The app offered a reliable and affordable cross-platform technology for text, voice, image, and video sharing in one-to-one or group contexts that worked across national borders complete with end-to-end encryption. At the [\*499] time of the acquisition, WhatsApp had 450 million monthly active users and was gaining users at a record rate of one million per day. Importantly, WhatsApp users were unusually engaged; more than 70% of WhatsApp users accessed the app daily and its volume of messaging rivaled the global total of telecom SMS.

Two characteristics distinguished WhatsApp from its rival messaging services, and from Facebook's corporate model. First, WhatsApp's founders committed the service to almost complete data privacy. Second, WhatsApp was advertising-free. Instead of the intensive data collection, aggregation, and analysis driving advertising revenue on other apps and networks, the company elected a paid model with most users charged a $ 0.99 annual subscription fee after their first year of service. The app offered an alternative entry point into scaled-down social networking using only existing phone contacts to connect users; it was more personalized and lacked the privacy concerns and tracking characteristic of Facebook.

In February 2014 when Facebook and WhatsApp announced their merger, Facebook served over 1.2 billion monthly active users. Mobile devices had become an essential component of that usership. More than 75% of active users accessed the network through mobile technology and in the fourth quarter of [\*500] 2013 mobile Facebook users outnumbered those using personal computers for the first time in the company's history. Growth in user engagement was increasingly driven by mobile access to the social network and Facebook anticipated that future growth would similarly depend on mobile connections. In its 2013 Annual Report, Facebook identified mobile applications with competing social features including text messaging, voice, image, and video sharing as a key source of competition for the network.

Facebook's reorientation toward mobile-first engagement led the company to develop and release its own standalone messaging app, Facebook Messenger. As mobile users sought short, private, and real-time communication options, Facebook identified and acquired one of the best-received startups in the mobile messaging market, Beluga, and refashioned it into a Facebook product. Upon its release in August 2011, Messenger became the number one most-downloaded app on the Apple store overnight. Although Messenger quickly claimed the status of the most-utilized iPhone messaging application in the United States, Facebook struggled to make headway in markets like Europe where early movers had an established advantage and in emerging markets where consumers were more likely to access their networks through feature phones. In early 2014, when Facebook and WhatsApp agreed on their merger, Facebook Messenger had 200 million users compared to WhatsApp's 450 million. With [\*501] the purchase of WhatsApp, Facebook would claim ownership of the world's top two messaging companies in terms of market share by user numbers.

The $ 19 billion price tag made the WhatsApp acquisition one of the largest mergers in Silicon Valley history. Facebook's offer nearly doubled a prior bid from Google to buy the startup for $ 10 billion. Moreover, the $ 19 billion deal amounted to approximately one-tenth of Facebook's total market value, while the monetization opportunities associated with WhatsApp were as yet unproven. In 2013, WhatsApp operated at a $ 138 million loss. WhatsApp's commitment to maintain privacy precluded merging its users with Facebook's social graph and adding advertising or other monetization options would require a substantial change in WhatsApp's approach to the messaging market. For Facebook, the benefits of owning WhatsApp clearly involved future competitive advantages in messaging and social media. Firstly, the purchase thwarted rival Google's attempt to gain ground as a social network. Secondly, the transition from social sharing on broad networks to one-to-one and group messages promoting private, real-time interactions indicated a significant shift in the social networking services market. Facebook CEO Mark Zuckerberg increasingly alluded to this shift as an important guide for advancing social networking and other social media with his declaration that "the future is private."

True to form, the FTC cleared the merger without challenge in April of 2014, with a letter warning both companies about their responsibility to maintain the privacy agreements in place when WhatsApp users accepted the company's terms of service. The letter highlights the distinction between Facebook's data collection and advertising platform model and WhatsApp's promises that it will [\*502] not collect any personal or contact data from mobile phones or messages or send any marketing material without the user's consent.

The European Commission also conducted an investigation of the transaction and cleared the deal. The European Union (EU) primarily analyzed the merger within the confines of the relevant market for consumer communications services, not as a potential competition merger. Consumer communication services includes stand-alone apps such as WhatsApp, Viber, Line, WeChat, Facebook Messenger, Skype, and those integrated with smartphone hardware or operating systems like Apple's iMessage. In their analysis of consumer communications services, the Commission noted that low switching costs, the tendency for users to multi-home, and the overlap between consumers of the two platforms would undermine any barriers to entry derived from the network effects captured by the merged companies. On these grounds, they concluded that the merger would be unlikely to lead to increased concentration in consumer communications services.

The Commission ultimately found no competitive concerns in the online advertising services market, based on WhatsApp's abstention from advertising and data collection and the number of providers supplying online advertising at the time. The EU also analyzed the social networking market and again found no competitive concerns. According to the EU analysis, WhatsApp was not a participant in the social networking market. The Commission considered a social network to involve many functions in addition to communications, including contact lists, user profiles, relationship status, and other social features of online activity. Although the EU reported that several industry participants informed the Commission that they considered WhatsApp to be a social network already, and predicted that absent the merger WhatsApp would expand and scale in this market, the Commission dismissed these opinions. The EU placed considerable weight on statements from WhatsApp management, stating "no indication was found of WhatsApp's plans to become a social network [as defined by the EU] which would compete with Facebook absent the merger." In the Commission's view, identifying WhatsApp as a potential competitor in social networking would expand the scope of alternative sources of competition to include other prominent firms in the consumer communications market, [\*503] including LINE, WeChat, iMessage, Skype, Snapchat, Viber, and Hangouts. Such an expansion would only make it less likely that the elimination of a single rival would raise competitive concerns.

Next, the Commission evaluated the potential for Facebook to gain market power in social networking by integrating the two platforms. The addition of WhatsApp's consumer base to Facebook's social graph would reinforce the network effects that maintained Facebook's dominance in the market for social networking services. According to the Commission's report and later documents, Facebook testified that technical limitations would prevent any such integration without significant user involvement. The claims that technical issues prevented integration were proven false just two years later in 2016 when Facebook began to add WhatsApp user data to the Facebook social graph. The EU fined Facebook €110 million ($ 122 million) for misleading the Commission but did not reverse its authorization of the acquisition.

What the EU did not consider was the possibility that the social networking market could be disrupted by a mobile, reliable, private, no-frills competitor. While the Commission noted that innovation in communications services was driven by consumer demand for reliability, privacy, and security, and acknowledged that the social networking services and consumer communications services markets exhibited significant overlap, it did not identify the trends in consumer behavior pointing toward the increasing the importance of private, mobile social platforms. Facebook had honed in on the competitive threat that this shift in consumer preferences presented for social networking, especially as it manifested in demographic and geographic groups critical to user growth such as young mobile users and those in emerging markets.

WhatsApp may have posed important potential competition issues. The strength of its reliable private messaging capabilities, its social orientation connecting users through their address books, its access to unique user data, and its ability to scale untethered to a monetization strategy based on consumer [\*504] surveillance could have raised a threat to Facebook's social network strategy. WhatsApp also may have been able to partner with complementary service providers to generate revenue and develop innovative and competitive social communications products. We will never know.

The EU's analysis highlights the problems with the potential competition doctrine. First, the problems of evaluating concentration in the social networking and mobile messaging markets are identical to those pertaining to the acquisition of Instagram: enforcement agencies have yet to identify a workable measure of concentration or a credible data source. The European Commission's report notes the lack of appropriate measure, despite its own reliance on user numbers (provided by Facebook) as a proxy for market shares. Second, the perceived ease of entry and broad consideration of potential competitors ignores the data barrier that reinforces firm dominance in online platform markets and makes it difficult for the government to isolate the impact of eliminating individual rival companies. Finally, according to the U.S. Horizontal Merger Guidelines, a five percent market share would substantiate the potential for WhatsApp to have significant procompetitive effects in markets for social networking or digital advertising. The EU cites conflicting views on the distinct boundaries of social networking markets, but even if these boundaries were clear, proof of deconcentration still demands appropriate measures of market share and current participation in the market. Harm to potential future competition was alone inadequate to challenge the merger.

The high initial burden on the plaintiff to present a case concerning future conduct and competitive effects serves as a serious deterrent to potential competition mergers, even by dominant firms. Under a simply structural presumption the FTC could have elected to challenge the merger and shifted the burden to Facebook to demonstrate why no harm to future competition could occur, and why, given Facebook's resources it could not internally innovate to achieve its competitive goals. A structural standard of this type should be embraced by critics of agency intervention who believe that the government is poorly positioned to make a strong empirical case, since representatives of the private sector would be the first source of analysis.

[\*505]

IV. Reform of the Potential Competition Doctrine

The Instagram and WhatsApp examples demonstrate how the potential competition doctrine is designed to fail by placing an unrealistic burden on the government in a challenge to any of the hundreds of mergers by dominant technology firms. We do not think this case is merely the result of new technology that has rendered the law obsolete and unworkable. We argue that the law was made unworkable because of the ideological goals of the Chicago School of Economics.

A comparison of the law of horizontal mergers with potential competition mergers is instructive. The Philadelphia National Bank structural presumption remains intact today. The plaintiff, typically the government, bears the initial burden in a § 7 horizontal merger case of demonstrating that the challenged merger should be presumed to substantially harm competition. This is accomplished by showing that the transaction will lead to undue concentration. The burden then shifts to the defendant to rebut the presumption. If successful, the burden then shifts back to the government to present additional evidence of competitive harm. The structural presumption has survived despite erosion by the lower courts. For example, in United States v. Baker Hughes, Inc., Justice Thomas (then on the D.C. Circuit) sought to dilute the presumption stating:

The Supreme Court has adopted a totality-of-the-circumstances approach to the statute, weighing a variety of factors to determine the effects of particular transactions on competition. That the government can establish a prima facie case through evidence on only one factor, market concentration, does not negate the breadth of this analysis.

In contrast to the courts, when the Reagan Administration appointees to the Department of Justice revised the Merger Guidelines in 1982 they replaced the strong structural presumption in the 1968 Guidelines with a detailed multi-step effects approach that placed the full burden of demonstrating a merger will harm competition on the government itself. The shift was motivated by the Chicago School supposition that most mergers are efficiency producing, an assumption that was never backed by empirical evidence. The higher burden made it much less likely that the antitrust agencies would bring a merger challenge, and when [\*506] they did, defendants could point to any defects in the agency's proof induced by its own standards.

The shift away from the Philadelphia Bank structural presumption for mergers that impact potential competition came earlier. It was achieved in complete form in Justice Powell's opinion in United States v. Marine Bancorporation. This wrong turn in 1974 must be corrected in order for the potential competition doctrine to have any practical application in tech markets.

Thus, the starting point for our approach would be to resurrect the pre- Marine Bancorporation 1968 Merger Guidelines. Under the 1968 Merger Guidelines, a merger would be likely to be challenged when a firm with a large market share (above 25%) purchases a firm that is "one of the most likely entrants into the market." The determination of whether a firm is a likely entrant is based on the capacity of the firm to enter, an incentive to enter based on attractiveness or a special relationship of the market, and potential profitability of entry, or a manifested interest in entry. While a possible starting point, a further correction is required. The 1968 Guidelines' analysis of entry is open ended and not sufficiently amenable to a tractable structural presumption that could be used by the courts.

What is needed to address the intractability of proof in a potential competition merger is a reasonable proxy that can incorporate a structural presumption for the likely entry or entry advantage of the startup. Thus, the second component of our test is to adopt the proxy that Professor Joe Brodley referred to as a "legal surrogate to identify the entry advantage of the acquiring firm." Professor Brodley recommended the use of the concept of "proximate markets" to provide the structural presumption of ability to enter and entry advantage for a target firm. As Professor Brodley explained:

Market proximity is a concept of presumptive entry advantage. Two markets are proximate to the extent that a knowledgeable firm in one market [\*507] possesses the necessary production and marketing information and other capabilities to operate in the other. Market proximity provides a suitable surrogate for entry advantage because, other factors being equal, there is less risk and therefore less expense involved in entering a familiar market.

To establish proximity, Professor Brodley focused on the factors that would be critical to the entry analysis of a business: production, marketing, technology, and customer relations similarities. More pointed criteria can be defined given the accumulated knowledge concerning tech industry mergers. For example, proximity to the general search market in which Google is dominant would include factors such as specialized search features, search advertising abilities, and the overlap of users with Google properties. The criteria would capture a vertical shopping site that is supported by search advertising and would clearly be a proximate market to the general search market. There are many such vertical markets that are potential rivals to Google's general search advertising revenues. Proximate markets to the social networking market certainly would include markets that compete with the functions hosted by Facebook's social network for user engagement and/or compete for similar targeted advertising dollars. In addition, the ability to gather user data complementary to Facebook's may be indicia of proximity.

We pause to recognize that other scholars have proposed different tests. We argue here that these tests do not create a sufficient standard for potential competition cases, and would condemn the plaintiffs in such cases to unworkable standards.

To start, Professor John Kwoka proposed a test involving two components, one involving structure and one involving effect: "(1) satisfaction of one structural precondition for concern with mergers involving non-incumbent firms, and then (2) demonstration of certain features specific to the case of (a) a deconstraining merger or (b) an entry-negating merger."

The first step, demonstration of a structural precondition, requires that there be moderate concentration according to the 1992 Guidelines approach. Under recent guidelines, the standard for moderate concentration is substantially increased. Regardless, substantial concentration is a condition for bringing any merger challenge. Over-reliance on the guidelines (in any version) will effectively eliminate a potential competition claim and analysis we seek to bring.

Under Professor Kwoka's test, if the structural precondition holds, then the analysis hinges upon whether the merger is entry-negating or deconstraining. [\*508] If the merger is deconstraining, the transaction "would likely be challenged on the basis of convincing evidence that the firm represented an effective and significant constraint on competition among incumbents." Such "convincing evidence" would include "documents in the possession of incumbent firms indicating active monitoring of and reaction to the non-incumbent party to the merger" or "market data that demonstrate significant responsiveness by incumbents to actions of the allegedly constraining firm."

With respect to an entry-negating merger, Professor Kwoka would have the enforcement agencies challenge such transactions if the transaction meets a multi-factored analysis. These factors are all focused on intent and ability to enter.

One of the authors of this Article, along with Salvatore Massa, proposed a two-step approach for a party moving to show entry with an opportunity for the non-moving party to rebut the claim. In that article, the first step is to determine whether the firm intends to and has the ability to enter the market. Evidence that directly relates to the commitments and investments a firm has made for entry are the most direct and relevant. The difficulty with this test is that if the evidence is more equivocal, there is little guidance as to how to proceed - a point admitted to in the original article.

[\*509] The second step considered other factors that may influence the relevance of potential entry. The primary issue is whether the potential entrant firm has an ongoing influence on the market. To make this determination, the court may turn to external factors, such as general industry knowledge and the internal documents of competitors, to see if there is a perception that the firm is a potential entry threat. Econometric evidence that a potential competitor is constraining prices in the market is the strongest evidence. Where econometric evidence is ambiguous, courts could look to other evidence. Regardless, the party not asserting potential competition would have the ability to rebut the potential competition claim to demonstrate that the firm would not be able to discipline the market, have too remote an entry date, is unfit to enter the market, or is not unique in its ability to enter.

There are multiple problems with this approach. Most pressing apart from the test's complication, however, is that the ability to rebut will likely swallow the claim. In particular, uniqueness would likely be difficult to argue against.

Others have argued that the potential competition doctrine is "superfluous," and could be integrated into the recent Horizontal Merger Guidelines. The authors argue that the potential competition doctrine, whether actual or potential, is a meaningless distinction: "Whichever label is applied, the theory must involve a unilateral or coordinated horizontal effect, and its evaluation should be essentially the same. The new Horizontal Merger Guidelines are consistent with this approach."

[\*510] We consider this a weird flex. For one, it is not as if there have been a plethora of potential competition cases under any version of the Guidelines. To the extent that the Non-Horizontal Merger guidelines raised issues inconsistent with consumer welfare, those Guidelines have been disavowed. Moreover, even the Department of Justice has not consistently adopted a guidelines approach when seeking to prove potential competition, particularly outside of the area of § 7. Even within the realm of § 7, the Guidelines approach has proven problematic, and any rebranding of the Guidelines is unlikely to cure the issues we describe here. In short, neither Instagram nor WhatsApp would have been challenged successfully under any of these tests.

Under our approach, both the Instagram and the WhatsApp mergers might have been challenged. Instagram operated in a proximate market. In the months before the Instagram acquisition, Facebook identified photo sharing as a key component of social network functionality, particularly on the mobile platform, and a key facet of Facebook's own success. The social features common to Facebook and Instagram demonstrate considerable proximity between the two companies. The private messaging offered by WhatsApp was rapidly becoming a prevailing aspect of online communication for individuals and groups, with networks established via the user's own address book posing an alternative to the public platform approach. In both cases, users' increasing reliance on mobile technology for digital interactions forced a collision between Facebook and the proximate markets that provided the aspects of online interaction its users increasingly demanded. Under the structural approach, tech mergers like Facebook's acquisitions of Instagram and WhatsApp could be challenged and receive the scrutiny they deserve. Regardless of the particular cases engaged, the process would develop a new guide to judicial decision making in tech markets.

We advocate the informed development of a fully structural presumption for potential competition mergers in technology markets. We think that this is how the law of potential competition mergers should have developed after the Philadelphia Bank case but was derailed by United States v. Marine Bancorporation.

[\*511]

Conclusion

Big Tech has demonstrated that it has an insatiable appetite for acquisitions of small startups. The sheer number of acquisitions should raise red flags for the antitrust agencies. After many hundreds of such acquisitions, so few challenges or requests to fully investigate these acquisitions demands some explanation. We argue that one aspect of the problem is that the law of potential competition has developed in a manner that essentially ties the hands of the antitrust agencies because it demands levels of proof that are intractable, particularly for a court.

We have arrived at this point because of the widespread acceptance of the Chicago School's approach to mergers. The Chicago School asserted that only mergers to monopoly were a legitimate antitrust concern, and that mergers that do not result in monopoly are usually efficiency increasing and undertaken for that purpose. With these background presumptions, the Chicago School advocates jettisoned the structural approach to mergers and replaced it with an effects analysis that raised the burden to merger challenges and provided defense counsel with multiple avenues to attack a government challenge.

The efficacy of the potential competition doctrine fell to the same unsound premises beginning in 1974 in United States v. Bancorporation. The doctrine now embraces difficult tests of conduct and performance. In markets where tipping occurs, technology is rapidly changing, and startup firms can scale and challenge dominant incumbents, a viable potential competition law is critical to protect competition and consumers. What is needed is to untie the hands of government antitrust enforcers by articulating a clear structural test to identify acquisitions of potential competition. To achieve this standard, we contend that very little innovation in law or in economics is necessary. We need only reverse the damage brought by the Chicago School and its neoliberal revolution and return to the potential competition doctrine of the 1968 Merger Guidelines.

# 2AC

## Innovation ADV

## Democracy ADV

## OFF

### T Increase Prohibitions---2AC

#### ‘Expand’ means to increase the extent.

Merriam-Webster’s 21 Online Dictionary, ‘expand’, https://www.merriam-webster.com/dictionary/expand

transitive verb

1: to open up : UNFOLD

2: to increase the extent, number, volume, or scope of : ENLARGE

#### ‘Scope’ refers to activity at the present time, not the abstract potential application of law.

Frank G. Clement 16 Jr, Judge on the Tennessee Court of Appeals, “Hamer v. Southeast Res. Group, Inc.”, Court of Appeals of Tennessee, At Nashville, 2016 Tenn. App. LEXIS 176, 3/3/2016, Lexis

When interpreting a contract, ordinary words typically have their ordinary meanings unless there is evidence [\*13] that the parties intended for the words to have a special meaning. Madson v. Madson, 636 So. 2d 759, 761 (Fla. Dist. Ct. App. 1994). The ordinary meaning of a word is often described as its meaning in the dictionary. See Siegle v. Progressive Consumers Ins. Co., 788 So. 2d 355, 360 (Fla. Dist. Ct. App. 2001); Beans v. Chohonis, 740 So. 2d 65, 67 (Fla. Dist. Ct. App. 1999). The ordinary meaning of a word or phrase is also described as "a natural meaning or the meaning most commonly understood when considered in relation to the subject matter and circumstances." See J.N. Laliotis Eng'g Constr. v. Mastor, 558 So. 2d 67, 68 (Fla. Dist. Ct. App. 1990) (quoting Granados Quinones v. Swiss Bank Corp., 509 So. 2d 273, 275 (Fla. 1987)).

If parties wish to depart from the ordinary meaning of common words and assign uncommon meanings to them, they must do so explicitly. See Madson, 636 So. 2d at 761. "One who would ascribe an exotic meaning to a term in a contract which otherwise has perfectly ordinary connotations must take pains to define the term either expressly or by express reference." E. Ins. Co. v. Austin, 396 So. 2d 823, 825 (Fla. Dist. Ct. App. 1981); see Russ v. State, 832 So. 2d 901, 907 (Fla. Dist. Ct. App. 2002) ("[W]here a statute does not specifically define words of common usage, such words are construed in their plain and ordinary sense." (alteration in original)); Koplowitz v. Imperial Towers Condo., Inc., 478 So. 2d 504, 505 (Fla. Dist. Ct. App. 1985) ("Whether they appear in a statute or in a declaration of condominium, words of common usage should be construed in their plain and ordinary sense.").

Here, this dispute exists because the parties' agreement does not define "scope" or "scope and purpose." Furthermore, the agreement does not identify the point in time when the "scope" of [\*14] Action's business is to be determined. Southeast contends that "scope and purpose" is ambiguous because it is susceptible to multiple reasonable interpretations. According to Southeast, "scope and purpose" means "at a minimum any business opportunity to be marketed to credit union members, including the telemedicine opportunity." However, the entirety of the parties' agreement and the "inconvenience, hardship, or absurdity" that would result from Southeast's proposed interpretation demonstrate that the agreement is not ambiguous and that the parties intended for the words "scope and purpose" to have their ordinary meanings. See Branscombe, 76 So. 3d at 948.

"Scope" and "purpose" are commonly-used words with commonly-understood meanings. Therefore, if the parties intended to ascribe an uncommon meaning to "scope" or "scope and purpose," they should have explicitly defined those terms. See E. Ins. Co., 396 So. 2d at 825. Instead of explicitly stating that these words have an uncommon definition, the agreement provides that its terms, covenants, and provisions "shall be construed simply and according to [their] fair meaning[s] . . . ." Consequently, the failure to specify a unique meaning for "scope and purpose" and the inclusion of the above-quoted section [\*15] indicate that the parties intended for these words to have their ordinary meanings. See id.; see also Russ, 832 So. 2d at 907; Koplowitz, 478 So. 2d at 505.

Under Southeast's interpretation, Plaintiff agreed to disclose and make available every business opportunity "to be marketed to credit union members." Such a broad definition appears to encompass every product or service imaginable, whether they have anything to do with Action or not. Under this interpretation, Plaintiff would be required to disclose an opportunity to sell cars to credit union members even though Action's business is not related to cars at all. The inconvenience, hardship, or absurdity that would result are weighty evidence that the parties did not intend for "scope and purpose" to have this meaning, especially when interpreting the agreement based on the ordinary meaning of "scope" avoids these difficulties. See Branscombe, 76 So. 3d at 948 HN9 ("The inconvenience, hardship, or absurdity of one interpretation of a contract or its contradiction of the general purpose is weighty evidence that such meaning was not intended when the language is open to an interpretation which is neither absurd nor frivolous and is in agreement with the general purpose of the parties.").

HN10 The ordinary meaning of words is found in the dictionary and is the most commonly understood meaning in relation to the subject matter of the parties' agreement. See Siegle, 788 So.2d at 360; Beans, 740 So. 2d at 67; J.N. Laliotis, 558 So. 2d at 68. According to one dictionary, "scope" means "1. The range of one's perceptions, thoughts, or actions. 2. Breath or opportunity to function. 3. The area covered by a given activity or subject." The American Heritage College Dictionary 1222 (3d ed. 1997). The operating agreement is concerned with the relationship of Action's members to each other and to Action, and the subject matter of section 6.6 is the duty to make certain business opportunities available to Action in order to avoid competition between Action and its members. [\*18] Based on the dictionary and the subject matter of the parties' agreement, "scope" most naturally refers to the range or breadth of the business that Action is engaged in at the relevant time.

Southeast contends this interpretation renders "purpose" redundant because "by definition, scope would always be within the purpose." We respectfully disagree. Contrary to Southeast's contentions, "scope" and "purpose" refer to different concepts. "Purpose" is aspirational and refers to what Action is capable of doing in the future (i.e. all lawful business for limited liability companies). In contrast, "scope" refers to what Action actually is doing or has done at the relevant point in time. Thus, an opportunity might be within Action's scope but not its purpose if, for example, Action had been organized for a limited purpose (e.g. to acquire real estate in Florida) but was in fact also engaged in the business of selling disposable mobile phones to college students. In this example, a business opportunity to sell mobile phones to college students would be within Action's scope but not its purpose.

Therefore, under the ordinary meaning of "scope," a member is required to disclose a business opportunity [\*19] if that opportunity (1) is within Action's aspirational goal — its purpose; and (2) is within the area that Action's business has or is actually covering at the relevant point in time. As a result, interpreting "scope" according to its ordinary meaning does not render any part of the agreement redundant.

Having concluded that "scope" refers to the breadth of the business Action is or has engaged in, we must turn our attention to determining when Action's "scope" should be assessed. The agreement does not specify whether Action's scope is to be determined as of the date of the agreement, the date of the discovery of an opportunity, or some other date. After reviewing the agreement, we conclude that the parties intended for Action's scope to be determined at the time when a member seeks to pursue the business opportunity in question.

#### ‘Prohibitions’ include regulation---reject arbitrary distinctions.

John G. Koeltl 7, United States District Judge, “United States Baseball v. City of New York”, United States District Court for the Southern District of New York, 509 F. Supp. 2d 285, 297, 2007 U.S. Dist. LEXIS 63234, 8/27/2007, Lexis

The City responds that its home rule and police powers are broader pursuant to Article IX, Section 2(c) of the New York State Constitution, New York Home Rule Law § 10(1)(a)(12), and New York General City Law § 20(13) than the plaintiffs suggest. These provisions give the City the power to enact laws for the "safety, health, well-being, and welfare" of its residents. The City asserts [\*\*29] that the Bat Ordinance does not constitute a "prohibition" because it does not condemn all use of non-wood bats. It bars their use in competitive high school baseball games, but not for example in high school practices, junior high school games, "pick up" games, or youth league games that are not school-sponsored. Moreover, the City persuasively argues that the suggested distinction between "prohibitions" and other "regulations" is artificial and untenable, because all regulations prohibit some conduct that is incompatible with the regulatory standards and all "prohibitions" leave some conduct untouched. For example, a New York court upheld as a valid exercise of the police power a New York City law banning the possession in a public place of a knife with a blade of at least four inches in length in People v. Ortiz, 125 Misc. 2d 318, 479 N.Y.S.2d 613, 620 (Crim. Ct. 1984). The plaintiffs suggest the law at issue in Ortiz was a not a "prohibition," but it appears to be at least as complete a prohibition as the Bat Ordinance, which prohibits only certain uses of bats with certain defined characteristics.

#### ‘Anticompetitive practices’ are horizontal and vertical restraints on competition---mergers fall under the former.

OECD 90, Organisation for Economic Co-operation and Development, “GLOSSARY OF INDUSTRIAL ORGANISATION ECONOMICS AND COMPETITION LAW,” OECD, 1990, https://www.concurrences.com/IMG/pdf/oecd\_-\_glossary\_of\_industrial\_organisation\_economics\_and\_competition\_law.pdf?39924/e9f9a49f59fa42b7de2397532968788aa2855447

ANTICOMPETITIVE PRACTICES

INSTITUTION DEFINITION

Refers to a wide range of business practices in which a firm or group of firms may engage in order to restrict inter-firm competition to maintain or increase their relative market position and profits without necessarily providing goods and services at a lower cost or of higher quality. The essence of competition entails attempts by firm(s) to gain advantage over rivals. However, the boundary of acceptable business practices may be crossed if firms contrive to artificially limit competition by not building so much on their advantages but on exploiting their market position to the disadvantage or detriment of competitors, customers and suppliers such that higher prices, reduced output, less consumer choice, loss of economic efficiency and misallocation of resources (or combinations thereof) are likely to result.

Which types of business practices are likely to be construed as being anticompetitive and, if that, as violating competition law, will vary by jurisdiction and on a case by case basis. Certain practices may be viewed as per se illegal while others may be subject to rule of reason. Resale price maintenance, for example, is viewed in most jurisdictions as being per se illegal whereas exclusive dealing may be subject to rule of reason. The standards for determining whether or not a business practice is illegal may also differ. In the United States, price fixing agreements are per se illegal whereas in Canada the agreement must cover a substantial part of the market. With these caveats in mind, competition laws in a large number of countries examine and generally seek to prevent a wide range of business practices which restrict competition.

These practices are broadly classified into two groups: horizontal and vertical restraints on competition. The first group includes specific practices such as cartels, collusion, conspiracy, mergers, predatory pricing, price discrimination and price fixing agreements. The second group includes practices such as exclusive dealing, geographic market restrictions, refusal to deal/sell, resale price maintenance and tied selling. Generally speaking, horizontal restraints on competition primarily entail other competitors in the market whereas vertical restraints entail supplier-distributor relationships.

However, it should be noted that the distinction between horizontal and vertical restraints on competition is not always clear cut and practices of one type may impact on the other. For example, firms may adopt strategic behaviour to foreclose competition. They may attempt to do so by pre-empting facilities through acquisition of important sources of raw material supply or distribution channels, enter into long term contracts to purchase available inputs or capacity and engage in exclusive dealing and other practices. These practices may raise barriers to entry and entrench the market position of existing firms and/or facilitate anticompetitive arrangements.

### T Subsets---2AC

#### ‘Antitrust’ is broad and includes any instrument designed to make markets more competitive.

D. L. Rubinfeld 15, Professor of Law at New York University, International Encyclopedia of the Social & Behavioral Sciences, Second Edition, p. 553

Antitrust Policy

The term antitrust, which grew out of the US trustbusting policies of the late nineteenth century, developed over the twentieth century to connote a broad array of policies that affect competition. Whether applied through US, European, or other national competition laws, antitrust has come to represent an important competition policy instrument that underlies many countries' public policies toward business. As a set of instruments whose goal is to make markets operate more competitively, antitrust often comes into direct conflict with regulatory policies, including forms of price and output controls, antidumping laws, access limitations, and protectionist industrial policies.

### Capitalism K---2AC

#### Growth is sustainable---robust environmental progress and increasing resource reserves prove---BUT ALT can’t save the environment either.

Andrew McAfee 20, principal research scientist at MIT, codirector of the MIT Initiative on the Digital Economy at the MIT Sloan School of Management, Doctorate from Harvard Business School, two Master of Science and two Bachelor of Science degrees from MIT, "Don't Misunderstand Earth Day's Successes," Wired, 4-22-2020, https://www.wired.com/story/opinion-dont-misunderstand-earth-days-successes/

We should all be intensely grateful to the people who took to the streets exactly 50 years ago on the first Earth Day. The modern environmental movement that crystallized then has given us a cleaner, better planet. The pressure applied to governments and businesses on April 22, 1970, has not let up since, and it has yielded two huge victories.

The first is massive reductions in the amount of pollution we and our ecosystems have to endure. In the world’s richest countries, which are the ones where environmentalism has most taken hold, the air, land, and water are all much cleaner than they were 50 years ago. This is not because these countries have simply offshored degradation to poor nations. Germany, for example, has the world’s largest trade surplus, yet has seen steady reductions in air pollution in recent decades.

If globalization is not the reason rich countries are much cleaner now than they were half a century ago, then what is? Effective regulation. The United States established the EPA and greatly strengthened the Clean Air Act in 1970, added the Clean Water Act in 1972, and kept taking steps over the years to bring down all kinds of pollution.

Some of the most innovative and helpful of these steps are cap-and-trade systems that create markets for pollution. Companies can trade with each other for the right to pollute, but the overall total is set by the government and declines over time. Over the past 30 years cap-and-trade has proved to be both relatively cheap and highly effective; a triumph of smart environmentalism.

The other great triumph is the improved health of species and ecosystems that we had pushed to the brink. Throughout the 20th century, relentless hunting almost wiped out whales. A nearly global moratorium was finally passed 1982, thanks in part to the “Save the Whales” movement that started in the mid-1970s (no doubt helped by folk superstar Judy Collins’ 1970 hit “Farewell to Tarwathie,” which introduced many people to whales’ haunting songs).

Many other species, including wolves, bears, beavers, and deer, have also come back after being near extinction in America. They rebounded in large part because we limited when, where, and how they could be hunted, and we limited trade in wild animal products. It’s generally illegal, for example, to sell hunted meat in the US. For the past 50 years, the environmental movement has carried on the laudable traditions of conservationism, which got its start early in the 20th century as Americans reacted in shock and horror to the extinction of the passenger pigeon and near elimination of the bison and other iconic animals.

Paradoxically, the great victories over pollution and extinction highlight environmentalism’s greatest weakness: a continued hostility to economic growth. The “degrowth” movement, which started in the early 1970s, stressed that human populations and economies simply couldn’t continue to grow as they had in the decades leading up to Earth Day. As philosopher André Gorz put it in 1975, “Even at zero growth, the continued consumption of scarce resources will inevitably result in exhausting them completely. The point is not to refrain from consuming more and more, but to consume less and less—there is no other way of conserving the available reserves for future generations.”

This seemed like an obvious truth to many in the 1970s, especially when they saw that the use of many natural resources—fossil fuels, metals and minerals, fertilizer, and so on—had been increasing in lockstep with the size of the overall economy. Since these resources were finite, and since their consumption went hand-in-hand with growth, growth apparently had to stop.

Yet around the world, it didn’t. The pace has slowed down a bit since the inaugural Earth Day, but this is mainly because the years between 1945 and 1970 saw exceptionally fast growth as we rebuilt our societies after two world wars. Except for that 25-year stretch, economic growth since 1970 is the fastest the world has ever seen.

So how are natural resource stocks doing? Oil is a great indicator of the overall story (its recent pandemic-induced demand free fall notwithstanding). At present we have about 50 years of oil left, given projected consumption and known reserves. That sounds dire, until you realize that 40 years ago, we only had 30 years of oil left. How can this be? It’s certainly not because we’ve cut way back on oil demand; we consume almost 40 percent more oil now than we did in 1980.

It’s because we kept finding more supplies. The same is true for every other economically important natural resource. Proven reserves—the amount of the resource we know we can access—have increased as we keep developing better technologies for finding and accessing them. And because the supply-demand balance keeps getting more favorable, resource affordability increases. The world’s average worker can, with an hour of their labor, purchase a greater quantity of every important resource than was the case just a few decades ago.

We live on a finite planet, but an incredibly abundant one. It contains enough of everything we need for as long as we’ll be around. Especially since, in the decades and centuries to come, we clever humans will almost certainly figure out nuclear fusion or some other technology that gives us limitless clean energy and lets us ignore fossil fuels. In short, there’s no need to slam the brakes on our growth. This happy fact is deeply counterintuitive, and it trips a lot of people up. But the evidence is clear: Degrowth is unnecessary.

In fact, it’s a terrible idea. Recall that the countries that have cleaned up their environments the most since Earth Day are the richest ones. This is not a coincidence, as Indira Gandhi knew in 1972. In a speech given in Stockholm, she said “Are not poverty and need the greatest polluters?... The environment cannot be improved in conditions of poverty.” Prosperous people and societies can afford, in every sense of the word, to care about the state of the planet we all live on, and to improve it.

Economic growth does not irreversibly degrade and deplete the planet. Instead, economic growth yields more prosperous people, who demand to live in a better world—a world with less pollution and more healthy ecosystems. The 50 years since Earth Day have largely shown that they get what they want.

The Covid-19 recession has given us much cleaner air in cities around the world, but at a terrible cost. We don’t need to endure such hardship to reduce emissions from car traffic. If we just made pollution more expensive and energy and transport innovation cheaper (via subsidies or research funding), we’d get the same clean skies without any economic devastation at all.

We face no shortage of environmental challenges over the next 50 years. We continue to overhunt, overfish, and raze ecosystems in many parts of the world. More extinctions loom. And of course we have to reduce the greenhouse gas pollution that’s causing global warming. The good news is that, in the decades since Earth Day, we’ve put together an effective playbook for meeting these challenges. I hope the environmentalists of the coming half-century will study this playbook, and realize that it shuns degrowth rather than advocating it.

### Courts CP---2AC

#### The plan solves global dependency on US platforms by using the FTC.

Anita Gurumurthy 19, Executive Director of IfTC and Expert Advisor for the UN Secretary General, “PLATFORM PLANET DEVELOPMENT IN THE INTELLIGENCE ECONOMY,” https://itforchange.net/platformpolitics/wp-content/uploads/2019/06/Platform-Planet-Development-in-the-Intelligence-Economy\_ITfC\_2019.pdf

Platform governance: the way forward

Platform governance is an overarching development policy challenge of our times, not just a narrow technology policy issue. A planetwide restructuring of economic ecosystems by digital platforms has triggered new contestations over socio-structural relations and geopolitical power. This calls for a cohesive policy response that can adequately and appropriately reorient the platform mode of economic organization towards a more equitable distribution of the efficiencies of intelligence scale economies. Such a policy approach also needs to be multi-scalar (spanning interventions at global to national and local levels) as well as cross-sectoral (encompassing integrated actions in digital, economic and social policy domains). We summarize the challenges for policy development in this chapter, also discussing the key building blocks of a comprehensive policy framework.

4.1 Governance challenges in the platform economy

a) Old laws don’t work: Most countries in the Global South lack legislative frameworks that address the rights and development implications of platformization trends. For example, as we found, individuals engaged in platform-mediated service work across different sectors – domestic work in the Philippines, tourism in Indonesia, and transportation in South Africa – are not covered under pre-existing labor laws (Barrameda et al., 2019; Bentley & Maharika, 2019; Mare et al., 2019). Similarly, the interests of small and medium enterprises and consumers are not adequately protected against unfair trade practices of platform companies in emerging digital commerce markets such as Nigeria (Nuruddeen et al., 2018). Even developed countries with legal-institutional frameworks for human rights enforcement and corporate accountability – such as EU member states – face difficulties in coping with the ongoing digital disruption. In France and Belgium, robust pre-digital labor laws are proving inadequate in providing social protection to platform workers with atypical employment contracts. Similarly, the application of preexisting consumer protection frameworks to digital services in the EU has meant the use of blanket disclaimer clauses by platform firms, with no explanations about obligations arising in the online context (Delronge et al., 2019). When new legislation specific to the digital context, such as the GDPR, has been introduced, the penalties for violation may often not be deterrent enough (Hintz & Brand, 2019). It has been found that companies such as Google, which have been repeatedly fined by the European Commission for non-compliance with prevailing legislation, nonchalantly continue their illegal market practices by treating fines as the costs of doing business.

b) State responses are knee-jerk: Platform regulation often times tends to be ‘scandal-prompted’. For example, in China, it was public outrage over the rape and murder of two female passengers by DiDi Hitch drivers in 2018 that prompted the ministry of transport to set up a national supervision platform for systematic background verification of the drivers enrolled with ride-hailing companies (Chen et al., 2019). Similarly, in Uruguay, the central bank rushed in to hastily regulate the P2P lending sector without fully understanding its operational dynamics as a response to increasing negative national media coverage about the sector becoming a ‘financial Uber’ (Aguirre & GarciaRivadulla, 2019).

c) Platforms become boundary objects, interpreted differently by different state agencies: The conflicting imperatives to create an enabling environment for the growth of the domestic digital sector whilst guarding against the monopolistic and exclusionary tendencies of the platform economy seem to culminate in a Catch-22 scenario impeding effective policy development. For example, in Argentina, there was a bitter tug-of-war between the Ministry of Production and the Argentine revenue service (AFIP) about the application of tax laws to the regional e-commerce platform MercadoLibre. While the Ministry of Production called for exempting the platform from tax liability as part of its larger strategy of encouraging domestic digital industry, the AFIP was of the opinion that MercadoLibre ought to be treated as a commercial firm rather than as a technology company. The Ministry of Production had its way, but it is difficult to ascertain whether the decision to treat MercadoLibre as a technology company deserving of tax exemptions will fare better for the long term health of the Argentinian economy in comparison to the AFIP proposal (Artopoulos, 2019).

d) Big platforms are mythified as the necessary route to success: The myth-making that surrounds platforms also means that governments, especially in the Global South, adopt pro-platform policy approaches. The promise of innovation and opportunity has often led governments to valorize platforms as an enabling force in aiding national growth. There has existed in the tech industry, even before the platform era, an “alliance capitalism” between industries of innovation and policy (Higgins, 2015, as cited in Chen et al., 2019). Consider the 2018 bid by Amazon for its new headquarters, which had city and state governments in the US outdoing one another to offer sops, tax cuts, economic incentives and even political positions to the company, convinced by the potential for jobs and economic growth that Amazon could bring in for the economy (City Lab, 2018b). Or, as in China’s case, where the Internet Plus vision has catalyzed and championed the growth of private platforms in many ways (Chen et al., 2019).

e) Platform companies tend to usurp public policy spaces: By becoming a part of the multi-stakeholder processes that drive policy, platforms take on a direct role in norm and rule development. Such formal membership in governance spaces raises concerns about conflict of interest. In Argentina, when traditional banks raised concerns over MercadoLibre’s new offerings for fintech services, the company successfully negotiated with the government to set up a commission to liaison between the central bank and itself, also managing to get a seat on the commission (Artopoulos, 2019). In December 2018, Netflix’s director of regulation was appointed to Brazil’s film board, Conselho Superior de Cinema, a recognition that the platform is an increasingly important player in the country’s media regulation discussions (Valente & Luciano, 2019).

f) The lack of binding international law gives corporations runaway power: There is no binding global legal framework to check corporate abuse and violation of human rights. Transnational digital companies not only flout domestic legislation with impunity, but also exploit the lack of cross-jurisdictional rules. When faced with the risk of prosecution for unfair market practices in national courts, they evade responsibility by transferring liability to their parent company outside the jurisdiction (Mare et al., 2019; Van Eck & Nemusimbori, 2018). For example, in 2017, the South African Transport and Allied Workers Union brought a case to the national Commission for Conciliation, Mediation and Arbitration (CCMA) on how Uber’s arbitrary deactivation and termination of drivers enrolled on the platform constituted a violation of protections against unfair dismissal under the country’s existing labor laws. CCMA took up proceedings against Uber SA, the South African subsidiary of the global platform company, and ruled in favour of the plaintiffs. A year later, the company managed to get the ruling overturned in the Labor Court on the technicality that Uber SA was a mere recruitment and training agency for Uber BV based in the Netherlands, which provided the app and made payments to partner-drivers.

4.2. Curbing digital monopolies

The platform economy displays monopolistic tendencies that curtail economic innovation and deepen inequality; but by no means is this an inevitability (Mann & Iazzolino, 2019). Traditional legal approaches to managing the rights, relations and conduct of persons and businesses engaged in commerce demand a major overhaul in the digital context (See Figure 5). This pertains to both commercial laws and to new rules concerning techno-design.

4.2.1 Changes to commercial laws

a) Competition law: Current approaches in competition law tend to regard short term consumer pricing gains as an adequate indicator of vibrant market competition (Khan, 2019). Understandably, this signal becomes extremely misleading in emerging digital markets where dominant platform companies often pursue strategies of free/deep-discounted products and services with an eye on long term consolidation of the network-data advantage for market domination (Curbing Corporate Power Alliance, 2019). In this scenario, competition law must move away from a narrow, neoliberal consumer welfarist approach. Instead, it must adopt economic structuralism as a framework to address the undue advantage that digital platforms enjoy in their role as “unavoidable trading partners” in the multisided markets they control (Cremer et al., 2019). The unique vantage that platforms occupy enables them to engage in upstream and downstream price manipulation, which policy must be able to check. The opacity that surrounds such data-supported gaming by platform companies makes it difficult to identify and establish proof of willful anti-competitive conduct. The EU has attempted to address this through its February 2019 regulation for platform businesses. It has mandated a duty of transparency (to be effective by 2020) with regard to standard terms and conditions of service (including data practices and notice of changes in terms of services) on all platform intermediaries providing digital services. This covers search engines, e-commerce marketplaces, app stores, social media and even price comparison tools. In addition, it has provided user guarantees for a right to explanation pertaining to algorithmic ranking and prioritization of goods and services on platform marketplaces (European Commission Press Release, 2019).

### Infrastructure DA---2AC

#### Won’t pass---Manchin won’t be persuaded

Hans Nichols 9/16, Political Reporter at Axios, “Scoop: Biden bombs with Manchin”, Axios, 9/16/21, https://www.axios.com/scoop-biden-bombs-manchin-b2b4acbd-24d0-40a3-ba6f-c0509e0e0224.html

President Biden failed to persuade Sen. Joe Manchin (D-W.Va.) to agree to spending $3.5 trillion on the Democrats' budget reconciliation package during their Oval Office meeting on Wednesday, people familiar with the matter tell Axios.

Why it matters: Defying a president from his own party — face-to-face — is the strongest indication yet Manchin is serious about cutting specific programs and limiting the price tag of any potential bill to $1.5 trillion. His insistence could blow up the deal for progressives and others.

Axios was told Biden explained to Manchin his opposition could imperil the $1.2 trillion bipartisan infrastructure bill that's already passed the Senate. Biden's analysis did little to persuade Manchin to raise his top line.

Manchin held his position and appears willing to let the bipartisan bill hang in the balance, given his entrenched opposition to many of the specific proposals in the $3.5 trillion spending package, Axios was told.

#### COVID and Afghanistan are nuking PC

AFP 9/7 – i24 News, “Biden To Unveil New 'Six-Pronged' Plan On How To Stop Delta Variant”, 9/7/2021, https://www.i24news.tv/en/news/coronavirus/1631027956-biden-to-unveil-new-six-pronged-plan-on-how-to-stop-delta-variant

US President Joe Biden will make a speech outlining plans to combat the highly transmissible Delta variant of the virus that causes Covid-19 on Thursday, as he attempts to recover rapidly slipping political momentum.

A White House official said Tuesday that Biden will "speak to the American people about his robust plan to stop the spread of the Delta variant and boost vaccinations."

The "six-pronged strategy" will involve both the public and private sectors, the official, who spoke on condition of anonymity, said.

"As the president has said since day one, his administration will pull every lever to get the pandemic under control," the official said.

Biden, who took office in January, won praise for his administration's concerted effort to get the coronavirus pandemic under control. Mass vaccination campaigns quickly got off the ground, boosting the Democrat's image as a competent crisis manager.

However, the combination of the aggressive Delta variant and large, mostly Republican-dominated swaths of the country where vaccinations continue to lag, has fueled a stunning resurgence of the disease.

Despite the role played by Republican leaders in refusing to impose mask mandates in hard-hit areas, Biden is taking much of the blame.

Also damaged politically by the traumatic exit from Afghanistan, the 78-year-old Democrat has seen his political capital plummet in the last few weeks.

#### Federal spending decreases overall investment by trading-off with states and localities

D.J. Gribbin 19, Non-Resident Fellow at the Brookings Institution, Founder of Madrus, LLC, Former General Counsel of the U.S. Department of Transportation, “Three Reasons To Think Twice About An Infrastructure Bill”, Politico, 3/27/2019, https://www.politico.com/agenda/story/2019/03/27/infrastructure-funding-bill-000886/

In physics, Newton’s Third Law states that for every action there is an equal and opposite reaction. In policy, too, every action creates a reaction, albeit rarely equal or opposite. In fact, the challenge of policy is that reactions, while inevitable, are difficult to predict. When weighing federal expenditures on infrastructure, policymakers need to keep in mind that allocating more federal funds to infrastructure might backfire. Here are three ways that could happen:

The “coupon effect”

The prospect of federal funding can dampen state and local funding. While voters overwhelmingly support increased infrastructure spending, their strong preference is that someone else pay for it. This dynamic makes it difficult for state and local leaders (who own 90 percent of governmental infrastructure) to turn to their electorate and ask for a tax or fee increase if the federal government is offering “free” funding.

This dynamic can be called the “coupon effect.” Imagine if shoppers in the market for a new suit were told that there is a small likelihood they will receive a coupon for 80 percent off their next suit purchase. Consumers will rationally engage in what economists call strategic delay and postpone their purchase in the hope of receiving a coupon, even if the chance of getting the coupon is very small. Every time a consumer considers heading to the store and buying a suit, he will ask, “But what if a coupon arrives tomorrow?” As a result, many will continue to delay until their suits (or our infrastructure) become unacceptably shoddy and worn.

In my experience, the prospect of federal funding has this same impact on state and local leaders considering a tax or user fee increase to expand or improve the quality of their infrastructure. This dynamic was clearly apparent in Kentucky in 2014, for instance. That year, a candidate for the U.S. Senate encouraged the communities around the Brent Spence Bridge (connecting Cincinnati and Covington, Ky.) to oppose a toll increase, because if elected, she would get the federal government to pick up the $2.6 billion tab to replace the bridge. Her campaign successfully increased opposition to tolling. Yet five years later, the debate on how to fund the bridge is still unresolved, and the probability of full federal funding is still just about zero (notwithstanding the fact that the state is represented by the Senate majority leader, who is married to the Secretary of Transportation).

While further study needs to be done, the coupon effect could actually result in a net *decrease* in infrastructure funds, especially when coupled with the challenges of substitution; states and local governments receiving an influx of federal dollars frequently substitute the new federal dollars for funds previously allocated to infrastructure and transfer their dollars to other policy priorities. As a result, a dollar in new federal infrastructure spending does not necessarily result in an additional dollar available for infrastructure.

The current non-federal to federal ratio of infrastructure spending is 3:1. Thus, if a 30 percent increase in federal spending (along with celebrations that the coupon is in the mail) dampened by 11 percent non-federal spending increases, our nation would be left with a net national decrease in infrastructure funding.

#### There are still tons of gaps

David Smith 21, Vice President of Business Planning & Performance at National Grid, “The Grid in the Infrastructure Package – What’s In, What’s Out, What’s Next”, GridForward, 8/19/21, https://gridforward.org/the-grid-in-the-infrastructure-package-whats-in-whats-out-whats-next/

What’s Not In The Package

Demand-Side Flexibility

Demand response and wider demand side management capabilities are essentially not funded in the bi-partisan package. One section encourages utility demand side management considerations, but no real funding goes to bringing demand side resources on the grid. With the potential of FERC 2222 to bring aggregated demand side and distributed resources into markets, much more widely available and adopted controllable devices, and other market developments necessitating the type of resource coming on the grid, this is a bit striking.

Building Automation

Support to ensure that buildings have higher level controls and capabilities to respond to grid signals was also not in the package. See comments in demand side and DER integration above and below.

Distributed Resource Integration

It’s not a future state, but a current need, in which aggregated edge resources can provide significant value to the grid. Turning distributed assets (solar, storage, EVs, thermostats, generators, hot water heaters, and much more) into a resource requires new technology, evolved models, new partnerships and more. Support to help this transition is essential. When well established values can be equitably dispersed to owners and all grid customers (and for the benefit of the system itself), we will have reached a new milestone in the evolution of our energy system – the grid has not reached this place yet and investing to get there is critical.

Analytics & Digital Infrastructure

Real-time grid telemetry to better understand and optimize the dynamics of the system was essentially not in the package and is also not present in most parts of the grid. What’s the saying ‘you can’t manage what you don’t measure?’ Are there exciting things you can do with the roughly 70% of advanced meters that are now deployed? Absolutely! But additional investments are required to apply a suite of capabilities, largely powered by the cloud, to the grid and it’s time that we take them off the shelf and use them.

Renewable Energy

Remember that part of the grid that actually creates the energy we need to run our economy? There are a handful of minor areas of investment in targeted deployments and demonstrations here and there offering a few hundred million dollars. But this package does not help fund the build-out of clean energy resources, nor the grid capabilities to help facilitate it. Economics of resources like wind and solar in many jurisdictions are just so cost-effective that their additions have largely won out over recent years, but if we want a lower carbon society we have to dramatically expand renewable resources. And, importantly, we must build a grid that ensures affordable, reliable power gets to people and businesses when they need it. It seems that the reconciliation package may have central aspects to helping support the further build-out of clean energy resources, but if the IPCC report that came out this week didn’t wake you up to the needs I’m not sure what else may.

#### No backlash---it has support from the public, Congress, media, and interest groups.

Robert Manduca 19, Assistant Professor, Sociology, University of Michigan, "Antitrust Enforcement as Federal Policy to Reduce Regional Economic Disparities," The ANNALS of the American Academy Political and Social Science, Vol. 685, Issue 1, 09/10/2019, SAGE.

Among possible federal regional development policies, reinvigorated antitrust enforcement stands out in several ways that make its establishment as a policy more likely. First, it is salient and familiar to voters. Most voters have encountered monopolies in their daily lives, whether they be airlines, utilities, internet providers, or tech platforms. Almost everyone has had a negative experience with a company too large or omnipresent to avoid in the future. Breaking such companies up offers a response to angry customers who would otherwise not have any way to express their frustration.

Moreover, aggressive antitrust enforcement has a long history in the United States, and it was widely practiced within the lifetimes of many voters. It has been a stated principle of capitalist economics since Adam Smith (Smith 1827), albeit one that has often been honored in the breach. In the United States specifically, antitrust enforcement fits with a longstanding American skepticism toward “bigness” (Lemann 2016; Rosen 2016). Perhaps for these reasons, the current antitrust movement has managed to find support among both liberals and conservatives. A poll conducted in September 2018, for instance, found that 65 percent of Americans—and 54 percent of Trump voters—think the government “should do more to break up corporate monopolies” (Dayen 2018). And leading proponents of antitrust enforcement in Congress and the media are found on both sides of the aisle (Crane 2018).

Perhaps more important than its broad appeal among voters, antitrust enforcement has the potential to attract support, or at least avoid opposition, from a wide range of organized interest groups. Of particular note is the potential for corporate ambivalence on this issue. Unlike many progressive economic policies, many companies—including quite powerful ones—stand to benefit from a reinvigorated antitrust regime. Yelp, for instance, has been a major critic of Google’s abuse of its search monopoly for several years (Dougherty 2017). When AT&T attempted to acquire T-Mobile in 2010, some of the most vocal opposition came from competitor Sprint (Singel 2011), though that did not stop Sprint from initiating its own bid for T-Mobile recently. Even Walmart, the largest retailer in the country, recently joined with other brick and mortar retailers to call on the Federal Trade Commission (FTC) to examine “persistent oligopolies in other parts of the retail system,” specifically singling out the market power of Amazon and Google (Dodge 2019). Companies like these could potentially become strong supporters of specific antitrust enforcement actions or a new antitrust movement in general.

#### No capital---blame doesn’t stick

Liz Goodwin 20, staff writer at the Boston Globe, “‘Sleepy Joe?' Trump struggles to stick a label to ‘Teflon Biden’,” BostonGlobe, 7-11-2020, https://www.bostonglobe.com/2020/07/11/nation/sleepy-joe-trump-struggles-stick-label-teflon-biden/

But the 77-year-old Biden has been surprisingly hard to caricature, in part because he has largely stayed in his Delaware home due to the coronavirus outbreak while Trump has struggled to respond to the twin crises of the pandemic and racial justice protests.

Biden similarly survived blistering attacks on his record from his rivals during the Democratic primaries. Senator Kamala Harris memorably lambasted Biden for his decades-old stance against busing to integrate public schools, while liberals derided his stated willingness to compromise with Republican senators — even ones who defended segregation — and his assurances to donors that nothing would fundamentally change if he were elected.

Now, Trump has half-heartedly begun painting Biden as a secret radical, one who wants to “defund the police” and dramatically raise taxes, or at least who will be manipulated into doing so. The move fits into Trump’s larger strategy of warning his mostly white base that civil rights protesters seek to “erase” their history and transform the country, and that Biden will facilitate that.

“Joe is just — look, let’s face it, he’s been taken over by the radical left,” Trump said on Fox News on Thursday night. “I think they brainwashed him.”

In one of Trump’s campaign’s recent digital ads, Representatives Alexandria Ocasio Cortez, Ilhan Omar, and Senator Bernie Sanders silently leap out of the wooden cavity of a Trojan horse topped with the head of Biden, as ominous music plays in the background.

But Biden faced months of criticism from liberals for being too moderate in the Democratic race. Trump’s attacks face a credibility problem.

“They try to say he’s extreme. But of course Joe Biden has been ‘canceled’ every week for the last two years by people who think he’s too centrist,” said Sean McElwee, the founder of the liberal polling firm Data for Progress. “All the stuff that people really hated about Joe Biden in the primary, it’s ended up making it hard for Trump to attack him in the general.”

“It’s hard to say this man is this woke statue destroyer,” McElwee added, referring to Trump’s messaging around statues of Confederates and other historical figures that have been defaced or toppled in recent weeks.

Biden’s own relative blandness as a political figure hurts Trump’s attempts to define him negatively, as he does not inspire strong feelings in a significant portion of the electorate. Just 22 percent of Americans say they dislike Biden “a lot” compared to 40 percent who dislike Trump “a lot,” according to a July Economist/YouGov poll.

That lack of venom can be seen at recent Trump events, where relatively few fans sport anti-Biden gear, unlike in 2016, when Hillary Clinton was skewered on pins and T-shirts and other paraphernalia, often in sexist terms.

“While I don’t want to say anyone is Teflon, Biden in some ways is unique because of his generic nature,” said Ian Russell, a Democratic strategist who used to run the House Democrats’ campaign arm. “The truth is they don’t have a ‘lock him up’ chant, they don’t have a ‘Crooked Hillary’ equivalent.”

### Populism DA---2AC

#### It’s thumped.

MFEM 8-19, Masuda, Funai, Eifert & Mitchell, Ltd., "The Implications of President Biden's ‘Executive Order on Promoting Competition in the American Economy’," Mondaq, 08/19/2021, https://www.mondaq.com/unitedstates/antitrust-eu-competition-/1103288/the-implications-of-president-biden39s-executive-order-on-promoting-competition-in-the-american-economy.

On July 9, 2021, President Joe Biden signed a sweeping executive order titled the “Executive Order on Promoting Competition in the American Economy” (the “Order”), affirming the policy of the Biden administration to “enforce the antitrust laws to combat the excessive concentration of industry, the abuses of market power, and the harmful effects of monopoly and monopsony.” To achieve this, the Order, among other things, directs regulatory agencies to assert oversight over certain business practices and encourages regulatory agencies to develop and/or strengthen rules. The Order includes 72 initiatives by more than a dozen federal agencies.

The Order specifically cites the areas of “labor markets, agricultural markets, Internet platform industries, healthcare markets (including insurance, hospital, and prescription drug markets), repair markets, and United States markets directly affected by foreign cartel activity.” The scope of this order is broad. On the other hand, the Order itself does not create new regulations or laws, leaving the specific implications of it vague.

Although the implications of the Order are not limited to the area of antitrust, the Order reflects the Biden Administration's emphasis on it. For example, the Order encourages the DOJ and other agencies responsible for banking to update guidelines on banking mergers to provide heightened scrutiny of mergers. The Order also encourages the DOJ and the FTC to challenge prior “bad mergers,” meaning that mergers that went unchallenged under previous administrations may be challenged in the future. Another specific area that the Order focuses on is the right to repair; it encourages the FTC to limit equipment manufacturers from limiting consumer's rights to repair.

Other affected areas of law include, but are not limited to, labor and employment (e.g. non-compete agreements) and consumer protection (e.g. financial data portability). Corporations with any significant activity in the United States should assess the impact that the Order would have on their businesses and prepare for the materialization of the specific initiatives included in the Order.

# 1AR

## Courts CP

### AT: Ex-Post Review

#### Ex-post review AND the present doctrine they enforce are problematic AND fail.

Zachariah Foge 19, JD, Pepperdine University School of Law, "How the Enfeebling of Antitrust Law Corrodes the Republic," Journal of Business, Entrepreneurship and the Law, Vol. 12, No. 1, 2019, Lexis.

II. THE CORRECT APPLICATION OF ANTITRUST LAW SHOULD BE APPLIED TO THE TECH GIANTS, AND FULFILL THE GOALS OF ANTITRUST

A. The Narrow View Problem and How the Law Should be Applied

The current regime of antitrust law is narrowly focused and does not account for all forms of anticompetitive harm, especially in the context of the online market. The heart of the current approach is concentrating on market efficiency through the lens of consumer welfare. Not only is this view flawed when compared to the original goals of antitrust law, but this view also forgets other critically important consumer interests like customer service and product quality, variety, and innovation. Because the current antitrust framework fails to account for these concerns, the framework should be abandoned.

Abandoning the narrow market efficiency approach would faithfully apply antitrust law to the tech giants, who have escaped scrutiny due to their care to avoid implicating consumer welfare concerns. The tech giants have gained massive unprecedented power because the market efficiency approach has not been able to meet the unique internet marketplace. Certainly, the drafters of the Act and the later scholars of the Chicago School could not have possibly envisioned the online market place, but the original Act intended to stop the concentration of market power and the creation of monopolies and oligopolies. The current [\*133] framework betrays that intent and decades of jurisprudence. It ignores our interests as consumers, as workers, as citizens, and as competing entrepreneurs. It also falsely conflates the goal of antitrust law as an outcome-focused calculation, rather than addressing the concern over the distribution of power in the marketplace. Hence, the current framework needs to be abandoned: the focus on welfare betrays the legislative history of antitrust and fails in the face of the current market's unique challenges.

Moreover, the current approach fails because it does not prevent harm to competition. Since the current framework of antitrust law solely focuses on price and outcomes, enforcement cannot begin until after the harm has already occurred. Specifically, the current approach ignores how and when a company acquires market power and only acts after a company has become so dominate that the company already rendered the market noncompetitive. This approach is self-defeating and makes little sense: what is the point of protecting competition in the market once the market is no longer competitive? Logically, it makes much more sense to protect competition when the market is at risk of becoming noncompetitive and is conducive to the legislative purpose behind antitrust law.

The Chicago School's approach not only betrays legislative intent, but also fails on its own terms. The current approach misapplies the law because the approach conflates market power with market efficiency and assumes only price and outcome can indicate competition. In fact, growing evidence shows that the Chicago School approach has led to higher prices, but not to any efficiency gains.

Further, the current antitrust framework has allowed such a concentration of power in a few massive companies which bars small businesses and entrepreneurs from entering the marketplace and from competing. There has been a 50% drop in small business ownership in America since the Reagan administration castrated antitrust law by implementing the Chicago School philosophy. Every administration [\*134] since 1981 has implemented the current antitrust framework and has "all but suspended traditional enforcement of America's antimonopoly laws . . . ." As a result, "regulators have done almost nothing to stop the great waves of mergers and acquisitions, with the result that control over most major economic activities is now more consolidated than at any time since the Gilded Age."

Clearly, the current antitrust framework allows for a highly concentrated market structure. This harms consumer interests overall and betrays the goal of antitrust law. Robust competition promotes consumer interest in a way that market efficiency does not. An open market that is free from domination by industry giants, who hoard power, best promotes competition. This harms consumers and workers, and, as expressed by Senator Sherman, threatens democracy. Because of the current application of antitrust laws, a small minority can amass outsized wealth and can influence government. The legislators also wanted to avoid a state of affairs where "private discretion by a few in the economic sphere controls the welfare of all." With the unprecedented power of the tech giants and their enormous spending in Washington, D.C., it certainly seems that these companies are exactly to whom the original drafters of the Act wanted antitrust law to apply.

Unfortunately, the current doctrine of antitrust law does not see the tech giants as anticompetitive because they employ techniques such as predatory pricing, which the current framework does not recognize as harmful; in fact, the current framework presumes that a corporation's activities are benign unless they lead to higher prices. Moreover, courts and agencies mistakenly believe that practices like large corporate mergers actually promote efficiency, and these courts and agencies fear false positives. The current antitrust landscape has led to a paradox where "modern U.S. antitrust protects monopoly and oligopoly, suppresses innovative challenges, and stifles efficiency." Focusing on efficiency has been illogical because it has not taken into account any post-merger inefficiencies and has led to the creation of firms that are too big to fail. The best guardian of competition is a competitive process, and the unbridled power of the tech giants effectively acts as a barrier to entry. [\*135] The courts should abandon the Chicago School approach, and no longer ignore the tech giants.

## Capitalism K

### Sustainable---1AR

#### Green growth is coming and solves---their ev ignores recent shifts.

Andrew McAfee 20, principal research scientist at MIT, codirector of the MIT Initiative on the Digital Economy at the MIT Sloan School of Management, Doctorate from Harvard Business School, two Master of Science and two Bachelor of Science degrees from MIT, "How we can have 'Green Growth,' increasing human population and prosperity while taking better care of our planet," Newsweek, 2-10-2020, https://www.newsweek.com/how-we-can-have-green-growth-increasing-human-population-prosperity-while-taking-better-care-1486342

Fifty years ago it might have been reasonable to fear that because of our bottomless desire for growth, we humans were going to strip our planet bare and poison it with pollution. But not anymore. The past half-century has shown us that we can increase human population and prosperity while also taking better care of the planet we all live on.

We still face real challenges now and in the years ahead, of which global warming is the most pressing. The good news is that we now know the playbook for effectively meeting these challenges. The bad news is that we're not doing a great job of following that playbook at present. We have to do better. We have to get smarter about meeting the problems we face.

In 1970, people took to the streets on the first Earth Day because of how we were treating our world. It's easy to see why they were so concerned. The 20th century, and in particular the post-war decades, witnessed by far the fastest growth in human history. Around the world, populations grew more quickly than ever before, and economies grew even faster as people strove for a higher standard of living. Unfortunately, it seemed that along with this growth came three side effects, all of which were both inevitable and terrible.

First, we were using up the earth's natural resources at an ever-faster clip. In the U.S., for example, consumption of aluminum, fertilizer and other important materials was growing even more quickly than the overall economy was in the years leading up to Earth Day. On a finite planet, this was a scary trend. If it continued, disaster seemed unavoidable.

At MIT, a team led by biophysicist Donella Meadows built a computer simulation of the global economy and used it to run scenarios about how the future would unfold. Their conclusions, published in the 1972 bestseller The Limits to Growth, were stark: "We can thus say with some confidence that, under the assumption of no major change in the present system, population and industrial growth will certainly stop within the [twenty-first] century, at the latest. The system...collapses because of a resource crisis."

The second bad side effect of growth was pollution. Air, water and land were all getting steadily dirtier in the years leading up to Earth Day. Levels of atmospheric sulfur dioxide in the U.S. increased by more than 60 percent in the three decades after 1940, and in 1969, the Cuyahoga River caught fire in downtown Cleveland. There seemed no end in sight to the pollution. Life magazine reported in 1970 that "Scientists have solid experimental and theoretical evidence to support...the following predictions: In a decade, urban dwellers will have to wear gas masks to survive air pollution....By 1985, air pollution will have reduced the amount of sunlight reaching earth by one half."

The third negative consequence of constant growth was extinction of creatures that we share the planet with. The passenger pigeon showed that even huge numbers provided no guarantee of survival. It was an abundant bird early in the nineteenth century, yet gone by 1914. Animals from the North American bison to the sea otter to the snowy egret to the blue whale came close to extinction during the industrial era, and it seemed clear that many others would vanish. As U.S. Senator Gaylord Nelson wrote in 1970, "Dr. S. Dillon Ripley, secretary of the Smithsonian Institution, believes that in 25 years, somewhere between 75 and 80 percent of all the species of living animals will be extinct."

If we wanted to save species, reduce pollution and avoid running out of natural resources, it seemed that we had to do one thing above all else: stop growing. Perhaps the broadest idea coming out of Earth Day and the nascent environmental movement was degrowth: deliberate shrinkage—rather than expansion—of our populations and economies over time. Degrowth wouldn't be easy and it might not be popular with everyone, but it seemed like a necessity. Philosopher André Gorz spoke for many when he wrote in 1975, "The point is not to refrain from consuming more and more, but to consume less and less—there is no other way."

It's important to be very clear on the following: people and societies around the world have not embraced degrowth since Earth Day. Global economic and population expansion did decelerate a bit after 1970, but this is largely because the 25 years after the end of World War II were a time of extraordinarily fast growth as countries rebuilt themselves. Except for that brief period, growth in the world's economies and populations has never in human history been as fast as in the years since 1970. Degrowth is nowhere to be found.

So what has happened with the three nasty side effects of growth: resource depletion, pollution and species loss? They must all have increased, just as populations and economies have, right?

Not at all. In the years since Earth Day, something weird and wonderful has happened: we ingenious humans figured out how to tread more lightly on our planet, even as we become more numerous and prosperous over time. This happy phenomenon is most advanced in the richest countries, but it's spreading around the world. Almost nobody anticipated that it would happen, and even today very few people are aware that the apparently ironclad trade-off between human prosperity and the state of nature has been eased. But it has. To see this, let's take another look at the three big problems.

First, resource depletion. The surest sign that something is becoming more scarce is that it's becoming less affordable. But without exception, important resources like fuels minerals, and foods have been getting more affordable, not less, for the world's average worker (in other words, not just for people in rich countries). Researchers Marian Tupy and Gale Pooley have calculated this hypothetical worker's ability to buy each of 50 resources over time—everything from crude oil to coffee to cotton. They find that the same market basket of all 50 that could be bought with one hour of labor in 1980 could be bought with only a bit more than 20 minutes of work in 2018. Not a single resource became more "time expensive" to the world's average worker over this period.

How can this be? One of the most important reasons is that many if not most resources are not nearly as scarce as we used to think. 1972's Limits to Growth provides a fascinating demonstration of this because it included a list of the proven reserves of several natural resources, along with predictions about how long these resources would last under various scenarios. If exponential economic growth continued, one of the team's main computer models showed that the planet would run out of gold within twenty-nine years of 1972; silver within forty-two years; copper and petroleum within fifty; and aluminum within fifty-five.

These weren't accurate predictions. We still have gold and silver, and we still have large reserves of them. In fact, the reserves of both are actually much bigger than in 1972, despite almost half a century of additional consumption. Known global reserves of gold are almost 400 percent larger today than in 1972, and silver reserves are more than 200 percent larger. And it's probably not too early to say that we're not going to run out of copper, aluminum and petroleum as quickly as estimated in Limits to Growth. Known reserves of all are much larger than they were when the book was published.

One other thing to keep in mind about natural resources is that in much of the rich world we're now using less of them year after year. And not just less per person, but less in total. In the U.S., which accounts for about 25 percent of global GDP, annual consumption of resources as diverse as copper, paper, water for agriculture, timber, nitrogen (a critical fertilizer component) and cropland is now trending downward. In addition, total American energy use has been essentially flat since 2007, even as the economy has grown by almost 20 percent. Developing countries, including fast-growing ones such as India and China, are not yet de-materializing. But I predict that in the not-too-distant future they'll start decreasing their consumption of some resources, just as high-income countries have.

As I explain in my book More from Less, two powerful forces are combining to drive this de-materialization of the economy. The first is tech progress, especially progress with all things digital (think of how much better and lighter today's LCD computer screens are than the cathode ray tube [CRT] monitors that preceded them). The second is capitalism, or intense competition among profit-seeking companies (think how much pressure CRT makers faced as LCDs took over their markets). This competition provides strong incentives for companies to save money on resources, and tech progress provides plenty of opportunities to do exactly that. So internal combustion engines are simultaneously lighter, more powerful and more fuel-efficient; smartphones replace entire shelves full of devices; and the economy de-materializes in countless other ways.

It's true that we live on a finite planet. But when it comes to thinking about resource consumption and availability, this fact is essentially irrelevant. Our experience since Earth Day has demonstrated that our planet is easily vast enough to supply us with all the materials we'll need, for as long as we'll need them. The real danger is not that our growth will deplete the planet, but instead that it will befoul it. So let's look at pollution next.

As every Economics 101 student learns, pollution is the classic negative externality, or bad outcome from a transaction that affects people who are not part of the transaction. If a factory pollutes a nearby river with its waste, for example, people living downstream suffer even if they don't buy any of the factory's products. Competitive markets do a lot of things well, but they don't deal with externalities. Instead, they often create them. So governments need to step in by forbidding the pollution (as we've done with the chlorofluorocarbons that were responsible for the hole in the ozone layer), placing an upper limit on it, or placing a price on it.

The logic of the latter approach is simple: if pollution is expensive, companies will work to reduce how much they spend on it, just like they work to reduce their spending on other materials. The cap-and-trade program adopted in the U.S. and other rich countries in recent decades to reduce atmospheric pollution is an attempt to reduce pollution by making it costly.

Cap and trade has been a huge success. As Smithsonian magazine summarized, it "continues to let polluters figure out the least expensive way to reduce their...emissions. As a result, the law costs utilities just $3 billion annually, not $25 billion [as they originally estimated]....It also generates an estimated $122 billion a year in benefits from avoided death and illness, healthier lakes and forests and improved visibility on the Eastern Seaboard."

The bans, limits, pricing programs and other pollution-control efforts established in high-income countries since Earth Day have been extraordinarily successful. They've caused pollution levels to go down in the rich world, even as economies and populations have continued to grow. The U.S economy is more than two-and-a-half times as big as it was in 1970, yet atmospheric sulfur dioxide levels have declined by more than 90 percent, and other kinds of air, water and land pollution have also declined dramatically.

A half-century ago, the conventional wisdom was that pollution was an unpleasant but unavoidable consequence of economic progress; as an American mayor said during debates in 1970 about strengthening the Clean Air Act, "if you want this town to grow, it has got to stink." But we now know that this is not true at all. To depollute, we don't have to embrace degrowth. We just have to put smart anti-pollution measures in place, then enforce them.

In recent decades, rich countries have done both. And what about low-income countries? Here the news is not as good. As researchers Hannah Ritchie and Max Roser summarize, "We see that the death rates [from air pollution] tend to be highest across Sub-Saharan Africa and South Asia...Outdoor air pollution tends to increase as countries industrialize and shift from low-to-middle incomes."

This is not a surprising finding. There's a hypothesis, based on the work of the economist Simon Kuznets, that low-income countries will pollute as their economies grew, but only up to a point. As people escape poverty and have more of their basic needs met, they will start to demand a cleaner environment. The government will respond to these demands, and overall pollution will start to go down, even as economic growth continues.

This pattern of rising-then-falling pollution is known as the environmental Kuznets curve (EKC), and in recent years we've seen it with air pollution in China. In March of 2014, Premier Li Keqiang announced to the National People's Congress, "We will resolutely declare war against pollution as we declared war against poverty." The government mandated that coal plants reduce their emissions, shelved plans to build new ones in highly polluted regions and even removed coal furnaces from many homes and small businesses (without, in some cases, providing anything to replace them).

These efforts worked. Economist Michael Greenstone found reductions in fine-particulate pollution of more than 30 percent throughout the country by 2018. He estimated that these reductions, if they were maintained, would add 2.4 years to the life of the average Chinese citizen. As Greenstone wrote, "It took about a dozen years [after passage of the 1970 Clean Air Act] and the 1981–1982 recession for the United States to achieve the 32 percent reduction China has achieved in just four years."

The EKC tells us something fundamental: that economic growth is at first the cause of pollution, then the cure for it. So to reduce pollution, we don't have to pursue degrowth; instead, we should encourage growth around the world. China's example gives us confidence that this approach works, and that growing countries will turn the corner and start polluting less in the years ahead.

We all need this to happen, because some kinds of pollution are global, not local. For example, the huge amount of plastic trash in the world's oceans that doesn't come from ships comes mainly from rivers that flow through low-income countries in Asia and Africa. The surest way to stop this flow of garbage is to make people in these countries prosperous enough that they can afford to care about the environment. We also need to ensure that rich countries don't start backsliding on their environmental successes. The Trump administration's moves to roll back wetlands protections, methane pollution standards and other safeguards are moves in the wrong direction, and should be reversed.

Of course, the greenhouse gases like carbon dioxide that cause global warming are the pollution most damaging to the long-term health of our planet and ourselves. Reducing future greenhouse-gas emissions is a major challenge, requiring political will and technological innovation. But it won't require any other radical departures from our current trajectory. Instead, the same approach that has worked for reducing other kinds of atmospheric pollution—namely, making it expensive—would also be effective

<<MARKED>>

at lowering greenhouse gas emissions.

A carbon dividend is an ingenious way to both make greenhouse gases expensive and to help people afford the resulting price increases. This dividend is a tax on carbon with an important twist: instead of keeping the money collected from companies, the government sends it right back out to the people as a "dividend" to each household. William Nordhaus was one of the winners of the 2018 Nobel Prize in economics in large part for his work on the carbon dividend. Clearly, it's an idea whose time has come.

Species loss is one of the most heartbreaking predicted consequences of global warming. Our greenhouse gas emissions could make some habitats uninhabitable, adding more animals to the sad parade of those already eliminated by our actions. But the threat of extinctions doesn't imply a need for degrowth. Instead, it makes more urgent the opposite: a world of nations and people prosperous enough to be good stewards of our planet and the life on it. Over the past 50 years we've seen remarkable increases in this kind of stewardship. In 1982, for example, most nations agreed to a complete moratorium on the hunting of whales, and populations are rebounding.

In addition to protecting species, we're also protecting territory. In 1970, less than 2.4 percent of the earth's land area was designated as parkland or otherwise conserved, and only 0.04 percent of the world's waters. By 2018 these figures had increased to 13.4 percent and 7.3 percent, respectively. In China, which has long been the world's largest market for endangered animal products, another important EKC has taken shape. As the country got wealthier, it eventually applied less pressure, not more, to some important animals. Strict bans are in place on buying, selling and possessing rhino and tiger products, and trade in ivory has been prohibited since 2017.

Are efforts like these helping in the fight against extinctions? They are. Documented extinctions appear to have slowed down in recent decades; for example, no marine creatures have been recorded as extinct in the past fifty years. It is far too early to declare victory over the forces of annihilation, but not too early to say that we know what works. Less pollution and more protection and prosperity are core elements of a winning strategy for protecting life on Earth.

Degrowth is not. The past half-century has exposed it as an unreasonable idea (given human nature), and an unnecessary one. This is not the same as saying that environmentalism is unnecessary. We should all be deeply grateful to the modern environmental movement born around Earth Day. On that day and countless others, concerned people took to the streets; put pressure on businesses, policymakers and elected officials; and otherwise advocated that we take better care of the planet we all live on. It worked.

But gratitude toward environmentalism does not mean continuing to support all of its original ideas. We now know that the core idea of degrowth—that there is no other way to conserve the earth for future generations—is simply wrong. With a few smart moves, including limiting pollution and protecting vulnerable species, we can have both greater human prosperity and a healthy, endlessly abundant planet. So let's get to work on building one.

#### They systematically underestimate market ingenuity – the same reason every other historically similar argument was wrong

Rune **Westergård 18**. Entrepreneur, Engineer and Author, founder of the technical consulting company CITEC. 2018. “Real and Imagined Threats.” One Planet Is Enough, Springer International Publishing, pp. 71–80. CrossRef, doi:10.1007/978-3-319-60913-3\_7.

Threatening reports about our ability to create disasters and even exterminate ourselves are not a new idea. A standard example is the British national economist Thomas Malthus in the early 19th century, who predicted that population growth would come to a halt because of starvation. Malthus calculated that the available food in the world couldn’t feed more than one billion people. He extrapolated the development from a still picture of his own time and couldn’t fathom that food production would increase tremendously thanks to new knowledge and technology. Our present food production is sufficient for seven times as many. Malthus didn’t pay attention to the fact that we live in a continuously changing civilisation, and the same kind of miscalculations are still made today. There are people who have even achieved the status of media superstars by presenting various dystopias and catastrophe scenarios. As early as 1968, Professor Paul Erlichs at Stanford University published the bestseller The Population Bomb, where he predicted that an imminent population explosion would result in hundreds of millions of deaths by starvation in the 1970s and 80s. Basically, he made the same mistake as Malthus, i.e. he treated knowledge and technology as if they were static phenomena. The most widely read environment report in the world, State of the World, was a loud whistle-blower when it was first published in the early 1980s. The Swedish version, Tillståndet i världen, was published yearly from 1984 and some years into the 2000s by the Worldwatch Institute Norden; I still have some of the early issues left. This report contains many valuable observations and suggestions, but also several basic analytical mistakes. In other words, it acts as an eye-opener, but it suffers from being tainted by political ideology. Its main weakness is that it doesn’t take the intrinsic driving forces of progress into account. State of the World was translated into most major languages and is, as already mentioned, the world’s most widely read environmental report. It has affected us all, directly or indirectly, through school and media. Even if the Swedish version I refer to was written some years ago, it is still worthy of discussion, firstly because it maintains an appearance of scientific validity, and secondly because it has served as a trendsetter for the general ideology which has been adopted by many later books and reports on the subject at hand. It still lives on as an engraved pattern in our conception of the world. In the report we can, for instance, read the following: A world where human desires and needs are fulfilled without the destruction of natural systems demands an entirely new economic order, founded on the insight that a high consumption level, population growth, and poverty are the powers behind the devastation of the environment. The rich have to reduce their consumption of resources so that the poor can increase their standard of living. The global economy simply works against the attempts to reduce poverty and protect the environment. We stubbornly insist to regard economic growth as synonymous with development, even though it makes the poor even poorer. Even if we up to this point have mainly described the environment revolution in economic terms, it is, in its most fundamental meaning, a social revolution: to change our values. Massive threat scenarios are still presented, for instance in the British scientist Tim Jackson’s book Prosperity Without Growth from 2009, which is one of the most widely read and frequently quoted works in this area. Tim Jackson, who is an economist and professor in sustainable development, explains how we humans are indulging in a ruthless pursuit of new-fangled gadgets in a consumption society running at full speed towards its doom. He also claims that material things in themselves cannot help us to flourish; on the contrary, they may even restrain our welfare. In other words, we cannot build our hopes that the economy, technology or science can help us to escape from the trap of Anthropocene, which has brought us to the brink of an ecological disaster. There are hundreds on books on this theme, and they all agree that the general state of the world is pure misery; everything is getting worse, the resources are being depleted, and that man will soon have destroyed the entire planet. The apparent reason for this, of course, is due to the consumption culture and the present financial system—which exposes man as a greedy, ruthless and ultimately weak creature. This attitude may serve a purpose as an eye-opener. But it is not very credible, and it may even be counterproductive. Of course, we can see a lot of problems ahead of us; but to solve them, we need the correct diagnostics instead of dubious doomsday prophesies. Focus: The Problem Since the focus of attention is so profoundly fixated on the problems in the climate and environmental debate, the progress already made—and the opportunities at hand—are often overshadowed. The example below will help to illustrate this point: In the year 2014, the Nobel Prize in physics was awarded to three scientists who had invented blue light emitting diodes—a technology that has made high-bright and energy-efficient LED lighting possible. As lighting accounts for 20% of the world’s total electrical consumption, this invention has the potential to radically reduce energy consumption and greenhouse gas emissions. In an interview made by the major Swedish daily newspaper Dagens Nyheter, one of the prize winners, Hiroshi Amano, says the following about energy-efficient, inexpensive and high-bright LED lights: “They are now being used all over the world. Even children in the developing countries can use this lighting to read books and study in the evenings. This makes me very very happy”. Shortly after this announcement, the news headlines declared that LED lighting was a threat to the environment. This statement was based on a report showing that LED lighting could be hazardous to flies and moths, which in turn might disturb the eco system. This is a typical example of how progress pessimists and, not least the media, think and act. In this case, they focused on a potential problem associated with LED lighting, and ignored the tremendous possibilities that the new technology offered to dramatically reduce greenhouse gases and thus spare the eco system (not to mention all the other advantages). Books and reports of the kind mentioned above tell us repeatedly about disasters, threats, problems, collapses and famines. On the other hand, they are notoriously silent about the great improvements actually made—the reduction of extreme poverty (not only as a percentage but also in absolute numbers), longer lifespans, dramatic global progress in education and healthcare, etc.

<<MARKED>>

The lack of positive media coverage on the environment means that many people believe that too little is being done, which is quite understandable considering the one-sided nature of the information they are presented with. Alarmist reporting almost always reminds me of pirates: they are unreliable and half their vision is blocked by their eye patches. It is vital that the media not only one-sidedly focus on the misery without presenting the progress made and suggesting constructive courses of action. The quality of our decisions in all respects depends on our knowledge, insight and attitude. Real and Imagined Threats Many people are convinced that the climate and environmental problems are growing. It is certainly true that our planet has its limitations, but many of the predictions from alarmist literature have been proven false. In the 1980s, the forest dieback was a frequently discussed subject. To quote the well-known German news magazine Der Spiegel, an “ecological Hiroshima” was imminent. Most experts at the time claimed that a wide-spread forest death seemed unavoidable. Additionally, the general mood of impending doom was augmented by the threat of a nuclear disaster during the cold war. I remember the pessimistic discussions among friends and how frequently the gloomy reports appeared in Swedish and Finnish television. The future of humankind appeared to be depressingly bleak. But the forest dieback never happened. On the contrary, the forest area has been constantly expanding in Europe, even during the entire period when the forest was believed to be dying. Today, only two thirds of the yearly accretion in Europe are cut down, according to the Natural Resource Institute in Finland. There are different opinions as to why the large-scale forest dieback didn’t occur. One theory is that the researchers’ evidence and conclusions had been incomplete and too hasty; the forest was actually never in danger. Others suggest that the emission limitations implemented prevented the disaster. My point is that the environmental catastrophe did not happen. Some other environmental problems, exaggerated or not, that have concerned us during the last decades have also disappeared from the immediate agenda: overpopulation, DDT, the ozone hole, heavy metals, lead poisoning, soot particles, the waste mountain, and the acidification of our lakes. Unfortunately, some environmental problems, like soot particles and waste, still remain in some areas, especially in poorer countries, where there are other, even worse problems that have yet to be resolved. The conclusion is, however, that we and our society in most cases have handled threatening situations quite well. When alarming symptoms are noted, scientists and other experts are summoned, and we act according to their diagnoses. It is no big deal that the diagnoses are sometimes wrong, as long as the side effects are not too severe. The main thing is that we do our best to avoid disasters, and on the whole, humankind has succeeded rather well this far. As individuals, we react very differently to various kinds of threats. The closer and more tangible the threat is, the more violent are the reactions—while distant and invisible symptoms, like the depletion of the ozone layer, concern us less. In the latter cases, we have to trust the scientists’ and later the politicians’ reactions. Does this mean that disasters are avoided thanks to war headlines, threats, and anxiety? I don’t think that this is the most important explanation; rather, it is factual and science-based information that produces effective results. But if exaggerated threat scenarios and reports of misery are needed to inspire the necessary political opinion, acquire research funding and create behavioural changes, we will have to live with that. The most important thing to remember in this context is that the actions shouldn’t cause more harm than the original problem itself. The risk with exaggerated threat and misery reporting is that it may inspire an over-reaction based on misleading diagnoses, or the opposite—a paralysing feeling of helplessness. It is necessary to take threats against the climate and the environment seriously, but not to a degree where our ability to reason and act is blocked by fear or anxiety. Many environmental debaters claim that the fall of the Inca and Roman empires were caused by the same causes that are now threatening our present civilisation—a short-sighted over-exploitation and rape of nature. Easter Island is another popular example. However, in my opinion it is both worthless and irresponsible to judge the world situation of today by copying the outcome of earlier cultural endeavours in history. The inhabitants of the Inca empire and Easter Island didn’t have anything even remotely comparable with the organisations, technology, medicine or general knowledge of today. It would be like comparing a case of appendicitis in the past to a case today. In pre-modern times, it was a fatal condition. In this day and age, it is cured by a simple routine operation. Today, humankind is conscious of the climate changes and other ecological challenges. And we also have the knowledge and resources needed to act. Facts, Propaganda and Hidden Messages During all the years I have followed the development of technology and society, I have repeatedly observed how a mishmash of serious research, political propaganda, and the hidden agendas of individuals have been distributed more or less randomly by the media. There are of course many different kinds of alarmism— everything from well-founded research reports to exaggerated prophesies of doom. It is far from simple to separate the wheat from the chaff. The actions taken against ozone depletion, lead emissions and the toxic chemical, dioxin, are all examples of how research has shown the way to successful results. Today, greenhouse gas emissions top the list of issues deserving our gravest attention, as it is a global phenomenon—just as the depletion of the ozone layer once was. There are also a considerable number of local environmental problems, such as drought, air pollution, forest depletion and overfishing. All of these are real threats that have to be acted upon, even though they are not global. However, I am always disturbed when a single global environmental issue is bundled with an assortment of several local issues, rather like a simplified trademark advertisement for the negative consequences of civilisation. This makes the information abstract and inaccurate, ignoring the fact that different locales require different solutions. Fear and alarmism are natural reactions that once protected us when we were living at the mercy of nature—they are evolutionary relics from our life in the savanna. Today, the same properties can be significant drawbacks. The transition from a primitive, animal-like state to the society we have today must, on the whole, be counted as a great success. But many people regard the same world as over-exploited, depleted, unjust, war-ridden and balancing on the brink of destruction. How can people living in the same epoch have so entirely different views of the world? In the sustainability debate, there is one faction dealing with the natural resources and ecosystems, and another focusing on the redistribution of wealth. There is even a third faction discussing a minimalistic lifestyle; for example, downshifting, with less work and less material welfare. When all these ingredients are mixed without discretion, the result is an anxiety soup that many have choked on. In a situation like that, we cannot expect any constructive initiatives to materialise. Instead, it would be far better to explore, research and discuss each dimension separately. What Is the Real State of the Planet? It is easy to generalise and say that we over-exploit the planet’s resources and pollute the world with our waste. But how many care to examine these statements in detail and ask exactly which resources are over-exploited? • Are fish becoming extinct? It is true that overfishing occurs in many places, which is, of course, unsustainable. However, this is not an unavoidable threat to the world’s total food resources. Fortunately, there are several examples of fish stocks that have either recovered or started to replenish once the fishing effort has been eased. • Is the air being poisoned? Many are convinced that the air we breathe is becoming dirtier all the time. But that isn’t true, at least not in the Western world. From the year 1990, emissions of sulphur dioxide have been reduced by 80%, nitrogen oxides by 44%, volatile organic substances by 55%, and carbon monoxide by 62%. Despite these dramatic improvements, 64% of Europeans believe that pollution is increasing. • Are the forests dying? It is a general belief that the forests in the developed countries are dwindling. But that isn’t true; on the contrary, the wooded areas are expanding. However, the forests are decreasing in the poor countries, where forestry and farming are still major sources of income, as they once were in the industrialised countries. • Are we drowning in waste? There are many who believe that we are surrounded by constantly growing mountains of waste. In the developed countries, the truth is that increasing amounts of waste are being recycled and the landfills are decreasing. • Will there be enough phosphorus? Phosphorus is an important nutrient in farming, extracted from phosphate ore. Many scientists fear that the finite natural resource of phosphate ore will become depleted in the future, which may jeopardise the world’s food supply. But there are already working solutions for this problem, such as by reclaiming phosphorus through digestion residues and sewage sludge. There are also technological solutions for the chemical extraction of phosphorus from polluted water—the remediation of lakes and rainwater by removing phosphorus is already a common procedure. Here we achieve a win-win situation—phosphorus is collected while preventing the eutrophication of lakes. • Will there be enough energy to go around? A common statement is that the earth’s population is too large, and that we consume too much energy with respect to the climate. This is one of those issues where we have to think in terms of symptoms, diagnoses, and medication. The symptoms are there for all to see: climate change. On the other hand, the diagnosis that we consume too much energy is wrong. The correct diagnosis is that we are not using the right technology; i.e. energy efficient power production without harmful emissions. Consequently, the correct statement would be that we consume energy that is produced by technologies that are harmful to the climate. The difference in wording is important. As the first diagnosis is “too high energy consumption”, the remedy will be to use a different medication than a diagnosis based on “the wrong technology”. Alarmist reporting can inspire bad decisions if the statements aren’t systematically reviewed and evaluated. It can also be misguiding to express environmental threats in general terms. Actions must be based on precise specific symptoms with corresponding diagnoses. If the doctor discovers that the patient is lame and suffers from a high fever, it doesn’t help to predict imminent death. Maybe the lameness and the fever have different causes altogether! A successful cure would probably include two different diagnoses with separate medications. Several recent surveys of the general conception of the world have been made— one is Project Ignorance by Gapminder and Novus in Sweden. One of the questions asked was whether CO2 emissions per capita and year had increased or decreased in the world during the last 40 years. The surveyed group was large and representative in order to give a fairly accurate picture of the common opinion. No less than 90% believed that CO2 emissions had increased. The truth is that they haven’t increased at all. It is important that decision makers on all levels learn how to see the wood from the trees. Decisions based on false preconditions can halt technological development, and thus also the development of the economy, welfare, and a healthier environment. The flow of innovations in the climate and environmental areas is accelerating rapidly. This can be seen in the number of improvements that have occurred in recent years, which can be counted in the thousands. Such improvements have to be weighted on the same scale as the problems in this area. That is not to say the problems should be ignored—they need to be acted upon. But they should not be allowed to occupy our brains to the extent that our power to act is paralysed. Is the Notion of Sustainable Technology-Driven Growth Over-Optimistic? The development of a technological society has always been questioned. In the 19th century, critics claimed that the technological revolution would create poverty. In the 1970s, it was generally believed that the forest dieback would cause a disaster. In the 1980s, the acidification of lakes and throwaway mentality of society were regarded as manifestations of the devastating properties of growth and industrialisation. Today, many fear the environmental effects of air travel and the production of electronic devices. There are people who seriously wish to halt economic growth and wind back the clock to the society of the 1960s. They recall this time period as small-scaled and down-to-earth, stress-free and idyllic. But they tend to forget that the refrigerators of that time required 90% more electricity than today, and that our teeth were repaired with mercury fillings instead of plastic. There were no X-ray CT scanners and no medicines against ulcers. In addition, there were many more people living without electricity. There was also more widespread malnutrition, a higher infant mortality, and, in fact, more wars. Cars were fuelled by leaded petrol, and sulphur emissions were 90% higher than today. The acidification of lakes, as well as polluted streams and fields, were serious concerns. Since then, technological innovations have reduced sulphur emissions and removed the lead from car fuel. At any given point in history, there have been critics claiming that this was the time when we had reached the optimal point in the development of the modern society. But we hadn’t, not then and not now. And the more our countries are modernised, the greater our possibilities to care for animals and nature become. In the mid-1800s, the killing of large animals like sperm whales didn’t concern people to any significant degree, despite the cruel hunting methods using harpoons. The benefits of the whale fat, mainly used for lamp oil to facilitate reading in the evenings, overshadowed any empathic impulses. In the 1850s more than 70,000 people were employed by the American whaling industry. There were 900 ships in the world hunting whales, and during one of the most active years, 8000 whales were butchered, which provided more than 300,000 barrels of oil. The oil extracted from the head of the sperm whale, the so-called spermaceti oil, was especially sought-after. It was of very high quality and sold for 1.50 US dollars per litre in today’s monetary value. As a consequence, the number of sperm whales in the world rapidly dwindled. However, when oil drilling started in Pennsylvania in the year 1859, the price of whale oil began to fall. The fast transition to petroleum products for lighting and other applications is considered to have saved the last of the sperm whales. Thus, new technology can both contribute to the protection of threatened animal species and provide the wealth to make it affordable for us to even save predators. Imagine what would happen if we were able to bring back someone from the 19th century and tell them that today we move wolves though the air by helicopter in order to save the species and expand its habitat; our ancestor would probably rather go back to sleep than listen to such apparent stupidity. Pessimism Does not Support a Sustainable Development There is a lot of progress going on in the world today, but not without negative side effects. When improving the world and dealing with the side effects, an optimistic attitude provides us with a much better chance of success than a pessimistic view. The optimist carries a positive inner beacon to follow, while the pessimist is always looking for potential traps and drawbacks. As visions and conceptions of ideas often become self-fulfilling, it isn’t difficult to realise what’s most constructive. All decisions—big or small, conscious or not—are affected and guided by our inner beacon. When solving a problem, such as developing a new product for example, it is necessary to have a conception of a working solution in mind. As a product developer, it is of course necessary to review every minute step in the process and question the choices made. You have to ask yourself if there may be a better material or a smarter design. Strange as it seems, this continuous struggle in the mind of the developer may appear to be a kind of pessimism, as it is all about looking for weaknesses in the imagined solution. It is not dissimilar from the process a doctor follows when selecting a diagnosis and a remedy. You start with certain hypotheses, examine, exclude, test, question and verify until you are satisfied that you have made the correct diagnosis. Then the choice of medication becomes much simpler. It would be fatal if the doctor was pessimistic from the start and worked in the belief that it would be impossible to find a reason for the illness, or a working remedy. This could then be the conclusion that such a doctor would unconsciously try to verify. Would you like to have a doctor like that? The same is true for climate and environmental problems—we need optimists armed with critical thinking to solve them. There are also so-called climate change deniers, who believe that man hasn’t really affected the planet and its ecosystems to any significant degree. Some of them claim that the influence of the sun and other natural phenomena are so enormous that human activities have no bearing on global warming. Perhaps these deniers are so deeply pessimistic that they cannot imagine any possible solutions. For ages, man has harboured a certain distrust of his own species. Throughout history, various religions have emphasised human shortcomings and presented assorted consequential threats. During the last 30 years, such prophesies have increasingly often been introduced by environmental activists and some political groups, whose messages have been significantly supported by the media. The underlying conception of humanity isn’t flattering. The human race is considered to be fundamentally ruthless, greedy, short-sighted and evil. Threats against the climate and much other misery on earth are caused by human failure. However, if we take the time to study the progress that has been made by the human race throughout the ages, we actually get the opposite picture. Can it really be evil, greedy, and short-sighted beings who put their own lives at stake to treat people infected by Ebola or HIV in poor countries? Who are the ones that are continuously reducing the number of starving people on earth? Who are the ones that invent vaccines for the children of the world? Who are the ones that have developed a civilisation where an increasing number of people get educated, and who struggle to reduce the casualties of war? Why blame an entire species for atrocities that are actually committed by a mere fraction? Establishing a firm belief in humankind should be the first step on the road to sustainable development.

#### Yes absolute decoupling – prefer consilience of data and experts – their indictments don’t assume urbanization and peak car

Peter **Newman 17**. Professor of Sustainability at Curtin University, and is a Lead Author for Transport on the IPCC (Intergovernmental Panel on Climate Change). 06-06-17. “Decoupling Economic Growth from Fossil Fuels.” Modern Economy, 8, pp. 791-805. DOI: 10.4236/me.2017.86055

Abstract The decoupling of fossil fuels from economic growth has not been imaginable for most of the industrial era but is now underway. The data for this are presented for the world and for typical nations. The mechanisms behind this are outlined and suggest that climate change goals to end poverty and to achieve the phasing out of fossil fuels are achievable if the trends are mainstreamed. Keywords Cities, Decoupling, Decarbonisation, Disruption, Forecasting 1. Introduction In 2015/16 the world’s governments committed to two core sets of goals for the future: the Sustainable Development Goals (SDG’s) and the Paris Agreement on Climate Change action which essentially aims to phase out fossil fuels by 2100 and 80% by 2050. Inherent in the achievement of the SDG’s is the need for economic growth, especially in the dominant SDG1 “End poverty in all its forms everywhere”. Yet in the past this has led to increases in fossil fuels and hence greenhouse emissions. Are the SDG’s therefore inherently in conflict with the need for climate change action? Or can the world achieve decoupling of economic growth from fossil fuels and greenhouse gases? For some these two sets of goals are incompatible as GDP and fossil fuels, hence greenhouse gases or GHG, have been totally coupled for over a hundred years so climate change action can only be achieved by degrowth [1]. For others, especially the UNEP and IPCC, who are more positive about sustainable development achieving economic and environmental goals together, the agenda of decoupling GHG and GDP is critical [2] [3]. 2. Methods The data for wealth and economic growth are from the World Bank [4] and GHG are from the International Energy Agency [5]. GDP and GNI are very sim ilar measurements of economic growth and GNI was chosen due to greater availability. They are standardized to a date back in the 1990’s where it is possible to show how growth rates begin to separate out and indicate the level of decoupling. 3. Evidence In 2016 the International Energy Agency announced that the world had changed. For the first time in hundreds of years the world was producing less greenhouse emissions than the year before without this being caused by an economic crisis [5]. In 2015 the amount of GHG emitted to the world’s atmosphere decreased by around 0.5% whilst economic growth continued at more than 3%. A few scientists had predicted this but mostly the fossil fuel lobby had been in complete denial over its possibility [6]. Figure 1 shows that for the first time the industrial world was producing wealth without this meaning more fossil fuels and more emissions. Despite its huge implications for a world that has faced down the global climate issue for decades without much good news, the world’s media were virtually silent. Perhaps this was because the EIA (from the US Government Energy Information Administration) were more sanguine predicting a continued growth in GHG of 1%, though their data were only up to 2012 and in reality, they did not consider the possibility of major changes often picked up by groups such as Carbon Tracker [7] [8]. Indeed, Carbon Tracker has shown that the new trends in GHG are following the kind of projections made by the IPCC’s carbon emissions targets much more closely than any other conservative projections. The decoupling of greenhouse gas emissions (GHG) from wealth (usually measured as GDP or GNI, gross national income) has been a UN agenda for several decades [2] [3]. The first signs of decoupling began in the 1990’s as Figure 1 indicates and their trajectories have been separating out quite rapidly for most of the 21st century. For many commentators and scientists, such as the IPCC, this relative decoupling was not significant enough for a world needing less total GHG until the actual decline in global emissions began. Now we appear to have reached a point where this can be seen in a peak in global greenhouse emissions. We now have absolute decoupling for the first time. 3.1. Understanding the Causes The simplest explanation for decoupling is to see the mathematical relationship between GHG and GDP as: GHG GDP GHG Energy Energy GDP = × GHG/Energy is called the carbon intensity and Energy/GDP is called the energy intensity. GHG/GDP will lower as the national carbon intensity decreases from greater use of renewables and natural gas; energy intensity will decrease as energy efficiency increases. Thus the increased use of renewables and the increased energy efficiency result in a relative decoupling of GHG from wealth. This relative decoupling was set as the basis for many national goals rather than seeking actual absolute reductions in GHG, especially if the countries were growing fast economically (like China and India) and were expected to have low but rapidly expanding GHG. This was seen by coal companies as the reason why coal consumption would continue to rise [9] [10]. This paper will show some of the rationale for why this is happening faster than expected but in particular that it is being driven by the world’s cities. Because cities are rapidly growing we can expect decoupling will continue to rapidly phase out the world’s dependence on fossil fuels and continue its goals in ending poverty. Figure 1 shows that Gross National Income (a way of measuring global wealth) has decoupled from greenhouse gas emissions. The notion of decoupling is based on the same idea as the economic theory of Environmental Kuznets Curves (EKC) which shows that as wealth grows people begin to choose environmental quality increasingly once their basic needs have been met. Decoupling is however happening earlier than expected in many parts of the world and indicates that there is more to this process than just simply getting wealthier. Perhaps one very key finding by a 2012 ADB report [11] is that the EKC was related to how quickly urbanization is happening as it is in cities that the phenomenon is most easily observed. In Figure 2 we have begun to breakdown the GHG into its main components by looking at consumption of coal and oil which are the largest causes of GHG. The spurt in global coal growth in recent times was mostly due to China but as can be seen in Figure 3 that era is over and China declined in its coal consumption in 2015 for the first time. Oil has also begun to plateau in China. The driving forces behind this will be outlined later but it is important to see that the growth in China’s economy is no longer dependent on growth in fossil fuels; they are phasing out and yet China is becoming wealthier. Indeed it is important to see that the reduction in GHG has occurred after an extraordinary 15 years of economic growth that took China out of its extreme poverty. The global Millenium Goals that were agreed to by the world’s nations in 2000 for the period up to 2015 focused on how development could halve extreme poverty—defined as being less than $1.25 per day. Across the developing world people living in extreme poverty declined from 47% to 22% between 1990 and 2010—achieving the Millenium Goal 5 years ahead of time [12]. This was mostly due to rapid declines in extreme poverty in China and India with some improvement in Africa. This process has taken around 1 billion people out of extreme poverty in a 15 year period. It is also important to see that the rapid decline in poverty has mostly occurred within cities and that the extremely rapid urbanization in China (and now in India and Africa) is what fuels this dramatic change. It also clear that there was not as much concern in this rapid urbanization in the emerging world’s cities for achieving fossil fuel and GHG reductions; these cities were and are focusing on ending poverty. But the new trend in China which decoupling is now underway provides great hope that the process will now spread to the whole emerging world. China is probably going to show the way to decouple rapidly as they invested $90 billion in renewables in 2015 (more than 60% of their investment in energy) so much of their continuing growth will be based around solar and wind rather than the fossil fuel-based economic growth of the past 15 years. The data on developed nations below show that decoupling can be anticipated in all elements of economic growth as GHG decoupling has set in. The big question for the world is whether this process of absolute decoupling is likely to continue in places like China and India and Africa as they develop. In Figures 3-6 we have presented the data on decoupling for Australia, China, Denmark and the United States to show that the extent of absolute decoupling is now considerable and this indicates that the global decoupling process is likely to continue. The largest decoupling is in Denmark, where the absolute decoupling began back in 1994 and has continued since, with coal and oil in significant decline without impacting on their overall growth in wealth. The US and Australia are following this pattern though not as spectacularly as Denmark. Most European nations have similar trajectories to Denmark. The EIA data in Figure 7 on China and the US, which stops in 2012 and so is not as clear in recent times about absolute decoupling, is certainly showing the sources of the decoupling. The reduction in the ratio of GHG over GDP in cludes both energy efficiency and growth in renewables leading to a decoupling of GHG and GDP of around 60% over 22 years in China and nearly 50% in the US. Similar data is available on India, the fourth largest emitter of GHG, and they have made very significant commitments to solar and to electric rail as well as energy efficiency that could lead to their absolute decoupling well before EKC theory would have predicted [13]. Although many would not want to say that the battle is over, there is no doubt that the era of fossil fuel dependence is ending. It obviously needs to keep going and gain exponential decline momentum. Our book is written to help with that momentum through the powerful forces of change that are potentially available in our cities. 3.2. Causes of Decoupling The three primary sources of change are government, business and the community [14]. Each of these play a part in any change and will be briefly outlined here before showing how the power of cities to bring these together is what is really driving the world to decouple from fossil fuels. 3.2.1. Government The Paris agreement in December 2015 (COP21) was an historic accord signed by 196 nations. It was significantly easier for the nations of the world to sign than in Copenhagen in 2009 (COP15) as the world could now see that the changes being required on them were already underway without causing a decline in wealth generation, as outlined above in Figure 1 and in national data. Nevertheless it was a significant achievement and a lot of lobbying and government activity was required to demonstrate that it was now possible to commit more heavily to the journey of removing fossil fuel-based economies 100% by 2100. The follow-up commitments began in New York in April 2016 with a signing by 175 nations to ensure that climate change is “well below” the 2 degrees’ limit accepted by the world’s scientists. Each country must now deliver the ratification from their parliaments and begin the adoption of annual goals that are meant to be ratcheted up whenever the country feels able to do so. The role of government in providing regulations and infrastructure to enable higher energy efficiency and low carbon electricity, fuels, appliances, buildings and vehicles has been a driver of change over the past 30 years. This process is usually one of being a small step ahead and pushing the system to produce a better low carbon outcome without too much cost or change. However it is possible, that disruptive innovations will begin to take over markets much quicker than governments have allowed for. In our view the emergence of solar and battery storage is likely to rapidly displace coal and the reduction of car dependence and electrification of transport is likely to rapidly displace oil. This process will depend on whether business and the community can see the changes emerging and accept that new fossil fuel-free goals are achievable in their cities and towns and how they can use the opportunity to ride this new economic wave rather than try to prevent the change. 3.2.2. Business There is a lot of evidence that the next area of growth for business is the green economy and that there are large groups of businesses partially or totally funding green innovation, products and services as the basis of their future [15]. The most significant driver in the 21st century leading to the removal of coal (and gas) for electricity production has been investment in renewables. As shown in Figure 8 below there was a point around 2004 when investment in renewables by the world’s bankers and financiers outstripped fossil fuel-based power investment. In the most recent data from Bloomberg New Energy Finance the investment ratio is now 2:1 in favour of renewables. Business has often been seen to have very short term goals of a year or so in terms of their strategic plans for market gains. But investors from the financial community look to see how they will make profits right through the lending period which is usually 20 years. When governments are debating the world’s scientists about how quickly they can remove fossil fuels from the market place, then it is easy to see why they would not want to invest in potential stranded assets like coal-fired power stations when other options that governments and scientists want to see, are available. In the US in the past 5 to 8 years the phasing out of coal was made easier by the availability of natural gas. At the same time dramatic growth in solar and wind power was underway and now it is possible to see how the natural gas limits are being reached as renewables becomes the fuel of choice [17]. The combination of these two forces led to the collapse of coal consumption in the US and decline and fall of the largest coal company Peabody with many stranded assets. In Australia a similar process has been underway with gas as the preferred option over coal but in the past 5 years the dramatic growth in rooftop solar has created a significant market that was not considered likely for decades. 1.5 million homes purchased roof top solar in 5 years without any real subsidies like feed-in tariffs and in Perth this reached 25% of households. The 550 MW produced is the largest power station in Western Australia and has led to the Minis ter Energy saying growth will inevitably reach 70% of households by 2020 and the utilities will never again need to purchase a coal or gas-fired power station [18]. The implications for the cities and towns in Western Australia for how they manage a potentially carbon-free power future in the next decade are a journey they were not prepared for. Fossil fuel-based power stations are now likely to be stranded assets and the key questions are about how battery storage can make this transition seamless for business and households and how it can enable the electricity grid to be maintained as a way of equitably distributing solar electricity. This is likely to be different in different parts of the city as well as in different types of settlements in rural and remote locations, as will be discussed later. 3.2.3. Community The reasons why Perth grew so rapidly in adoption of solar include factors like easy access to Chinese mass produced PV cells (and now batteries) and the high price of electricity, but mostly it was driven by the community. In the 8 years from 2008 there was a significant economic boom in Western Australia when the rest of the world was frantically adapting to the GFC. For a period Perth became the wealthiest city in the world and the new money was flooding into many household bank accounts. Many chose to put it into a rooftop solar system because of their interest in long term sustainability (over 80% of the community want to see action on climate change) as well as having an investment that would pay for itself in 5 years. Community values can easily be underestimated when facing the future and the need to address complex matters like climate change. The majority of this rooftop solar has gone into the outer suburbs which in Australia are generally less wealthy though they often have bigger homes with bigger roofs and bigger power bills [19]. The inner suburbs are much higher socio-economically and have higher green intentions but in reality it was the outer suburbs who have made the majority of the investment in solar. It indicates that the mechanisms for decoupling economic growth and fossil fuels will vary between cities and within cities [20]. 3.3. What about Oil? The reduction in consumption of oil is another example of how cities have begun to drive the change to remove fossil fuels at a faster rate than many anticipated. It also is a combination of government, business and the community. The collapse in oil price as well as the collapse in coal price appears to be due to demand issues as well as supply issues. Demand for coal will continue to drop as renewables takes over; especially as battery storage becomes cheap. Demand for oil is expected to decline as vehicle fleets are electrified; however it has already gone down despite there being only a few percent of electric vehicles. The reason why this has happened appears to be a city issue: for the past 150 years cities have been moving out as they have followed first trains and trams then automobiles, but now they are coming back in [21]. Our work on cities has shown that there is an exponential relationship between urban density and car use/fuel use as in Figure 9. If cities have begun to increase in density rather than decrease then they will drop down the steep curve quite quickly creating exponential decline in car use. This is what we are now seeing in all the world’s developed cities and the phenomenon is being called ‘peak car’ as car use per capita has peaked and is in decline. Indeed it is very similar to the decoupling phenomenon described above. Many cities are now seeing that their economic growth is dependent on them reducing their car use. Hence cities like Washington DC and Portland, Oregon have demonstrated that their wealth is decoupling from car use (Figure 10). It should be no surprise that this strong decoupling is associated with cities that have invested in good urban rail systems in recent decades. The decoupling is closely associated with this as our data also shows that urban traffic is slowing in all the world’s cities and urban rail is now able to outcompete cars as they can go around, over or under the traffic [21]. Governments play a big part in generating this transition as they are needed to help plan, though not necessarily finance, such large scale infrastructure. However business and community are also critical as they are responding to a new economic process associated with the knowledge economy. Many businesses that are part of the new innovation economy are locating in city centres where they can have creative face-to-face contact with people from various backgrounds and professions [22] [23]. These new jobs are generally for the young Millenials and also older wealthier professionals. As a result they are the backbone of the social movement that has rediscovered inner and central city living. The market demand for dense urbanism that is not car dependent has therefore grown dramatically in the 21st century city. In more recent work the same phenomenon of peak car has been found in Shanghai and Beijing which are now carrying 8 and 9 million passengers a day on their new Metro systems enabling a significant reduction in the relative use of cars [24]. The rapid change into decoupling growth in wealth from growth in car use is happening in emerging cities because they have the urban fabric that is not suited to large scale growth in car use. The building of 81 electric urban rail systems in Chinese cities and 52 in Indian cities is indicative of how the switch away from oil is happening before many expected. 3.4. The Future A range of fossil fuel demand forecasts and the forecast just for renewables are set out in Figure 11. The nine organizations they used include the main fossil fuel companies and the main government-based forecasting bodies of the IEA and the EIA. The average growth from 2015 to 2050 is 580 EJ to 850 EJ with the oil companies suggesting well above these forecasts. The average for renewables goes from 80 EJ to 200 EJ with the fossil fuel companies suggesting much less than this. What this suggests is that based on historical trends global energy demand is forecast simply to grow in line with global population growth and growth in GDP. Owing to this, fossil fuel companies predict that due to the continued forecasted rise of energy consumption, their products will meet the majority of this demand with a small though growing proportion of renewables. This is business as usual. However, this paper has shown that decoupling of fossil fuel energy from GDP has become firmly established and is being driven by cities from across the world. What if the trends are setting in to be significantly faster than most of these forecasts dare to suggest? What could happen if we took seriously the kind of disruptive innovations in our cities that are likely to lead to rapid decline in fossil fuels and rapid growth in renewables? In order to separate out the components of GHG growth we have used the Kaye simplification (based on the old Paul Ehrlich simplification of Impact being a combination population, resource consumption per capita and technology efficiency per unit of resource consumption): GHG = carbon intensity per unit of energy (GHG/Energy) x energy intensity per unit of wealth (Energy/GDP) x wealth per capita (GDP/population) x population The carbon intensity of energy is going down globally and will increase rapidly as renewables increases its proportion. The decoupling of energy from wealth has been increasing with a 35% decoupling between 1990 and 2015 and from the above diagrams this will grow even more rapidly as the two separate out. Wealth per capita is likely to continue its gradual growth and population growth is predicted to continue but slow. The interactions between all these factors are complex for example urban population growth will help propel the economic processes that enable solar adoption and urban regeneration as well as being a factor in creating more GHG until these processes enable the phase out. The scenario considered most feasible will be something like the IPCC goal of 80% less GHG by 2050 as in Figure 12. 4. Conclusion The end of fossil fuel dependence is hard to imagine but it is getting easier because the trends show it is underway. This paper has shown the kind of exponential growth in renewables and decline in fossil fuels

<<MARKED>>

that we have just started to see could in fact decouple economic growth from fossil fuels much more rapidly than most have forecasted. The continuation of these trends will require a combination of different forms of solar-based power, different forms of electric transport and different forms of urban fabric. Most of this will be in cities. Government, business and the community will drive it in different ways in different