Bubbles and the Value of Innovation†

Valentin Haddad
UCLA and NBER

Paul Ho
Federal Reserve Bank of Richmond

Erik Loualiche
University of Minnesota

May 19, 2020

Abstract

Episodes of booming firm creation coincide with intense speculation on financial markets leading to bubbles — increases in private market valuations and real investment followed by a crash. We provide a framework reproducing these facts and we analyze how speculation changes both the private and the social value of new firm entry. In our model, measures based on financial markets information, in contrast to measures grounded in real outcomes, indicate that speculation increases the private value of innovation and reduces negative spillovers to competing firms. We confirm these predictions empirically in the universe of U.S. patents issued between 1926 and 2010.

†Haddad: vhaddad@ad.ucla.edu. Ho: paul.ho@rich.frb.org. Loualiche: eloualic@umn.edu.
We thank Andrew Atkeson, Markus Brunnermeier, Edouard Challe, Alexandre Corhay, Emmanuel Farhi, Nicolae Garleanu , Harrison Hong, Pablo Kurlat , Deborah Lucas, Alberto Martin , Stavros Panageas, Alp Simsek, Avanidhar Subrahmanyam, Jaume Ventura , and seminar participants at Columbia University, HKUST, London Business School, Minnesota, MIT Sloan, Princeton, UCLA, University of Luxembourg, Macro Finance Society meetings, FIRS, SED, Banque de France, Bank of England, NBER Asset Pricing, and Minneapolis Fed for useful comments. The views expressed herein are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

1 Introduction

Episodes of booming innovation often coincide with intense speculation in financial markets, leading to periods that have been described as bubbles: large increases in firm creation and market values followed by a crash (e.g. Scheinkman, 2014). This paper studies the implications of speculation for the private and social value of innovation, quantities that are central for our understanding of how innovation spreads through the economy and its efficiency.

To answer this question, we develop a new framework where firm entry is endogenous and investors disagree about which firms are more likely to succeed. Because of this disagreement, different investors speculate in different firms and a bubble emerges We formulate predictions about the private and social value of innovation during these episodes, which we test using over a million patents issued between 1926 and 2010. Most importantly, in both our model and the data, we find that during bubbles *(i)* the increase in market value of a firm following the introduction of a new patent is larger; *(ii)* the relative spillover to the market value of competitors is smaller; and *(iii)* there is no commensurate change in *outcome-based* measures of value focusing on patent citations or output. As a byproduct, we show how disagreement over new innovation alters the standard results from the innovation literature about the efficiency of firm entry.

To understand the role of speculation during innovative episodes, we analyze a model in which new ideas are implemented by firms that compete with each other. We allow for a rich range of firm interactions, while remaining parsimonious enough to be tractable and yield clear testable predictions. In the first stage, firms are created by raising money on financial markets; in the second stage, competition and production occur. In our baseline analysis, competition takes a simple form: only a fixed number of the best firms get to produce. The novel ingredient of our theory is that investors agree to disagree about which firms will be more productive. Each investor picks her favorite firms for her portfolio, and therefore values her investments more than her average beliefs.[1] This mechanism increases valuations and incentivizes more firms to enter. In the second stage, not all investors can be correct about which firms produce: prices drop and many firms fail. This rise and fall as a result of speculation echoes broad narratives of bubble episodes, as well as some of their more subtle features. Disagreement is more natural in the face of new ideas so these episodes should be more likely following the introduction of a broad new technology (Brunnermeier and Oehmke, 2013; Scheinkman, 2014), or among younger firms (Greenwood, Shleifer, and You, 2018). In a dynamic extension of the model, we show that they are also related to more intense trading volume (Hong and Stein, 2007; Greenwood, Shleifer, and You, 2018).

[1] Van den Steen (2004) studies how disagreement and choice lead to optimism.

Our framework for heterogeneous beliefs overcomes the challenge of analyzing the private and social value of firms in the presence of disagreement. Market participants agree on the aggregate distribution of firms —and therefore all macroeconomic outcomes— but disagree on the relative positions of specific firms in this distribution. Each firm faces the same distribution of beliefs across investors even though the identity of these investors differs, leading to a symmetric equilibrium in the first stage. This symmetry gives rise to clear notions of private and social value defined in both market-based and outcome-based ways. The market-based approach considers prices of firms in financial markets. Under this criterion, the private value of a firm is its price, while the social value of that firm is the amount by which its introduction changes the price of all firms in the economy.[2] From an outcome-based perspective, the private value is the average output of a firm, while the social value is how much the introduction of a firm changes total output in the economy.

The model makes two predictions about the market-based value of firms. Firstly, it predicts that speculation increases the market-based private value of firms. While the aggregate productivity distribution of new firms is unchanged, investors have higher valuations of their own firms because they expect that these firms will be more productive than the average firm in the economy. Secondly, it predicts that the wedge between the market-based social and private values of a firm is attenuated by speculation. Each new firm introduces a *business-stealing effect* as it displaces other firms. With speculation, investors believe they are investing in the most productive firms in the economy, making them less concerned about being displaced by a new firm. Their valuation of their own firms is thus impacted less by any additional firm entry. In contrast, the outcome-based measures of value unchanged since speculation affects neither the aggregate productivity distribution nor the competitive structure.

An historical episode offers an illustration of our model: in 1686, William Phips secured funding in England from the 2nd Duke of Albemarle and his syndicate to search for sunken Spanish ships in the Bahamas. His expedition found 34 tons of treasure, yielding large returns to his investors and creating considerable speculation around treasure hunting technology. Seventeen patents for ways to recover underwater bounty were registered between 1691 and 1693, a large spike in patenting activity for the time. Based on each of these innovations, numerous firms were introduced on equity markets and raised large amounts of capital despite competing for what was clearly a small pool of treasures. These high valuations were consistent with our notion of high market-based value that is unaffected by competitive spillovers in presence of speculation. The boom was so large that it is sometimes credited for the emergence of developed equity markets in England. However, these expeditions only succeeded in finding a few worthless cannons, reflecting no increase in the outcome-based value of the new

[2]In our richer specifications, we also incorporate effects on other participants in the economy.

patents and firms.

To verify the model's predictions empirically, we use the universe of patents issued by public firms between 1926 and 2010 to measure the relationship between innovation and market values. We follow the empirical literature and construct empirical counterparts to the private and social value in our model. Specifically, we measure the market-based private value of new innovations using the stock returns of issuing companies in the days following the patent approval, as in Kogan et al. (2017). We estimate the market-based social value relative to private value of new innovations by using the response of firm valuations to innovations by their competitors, controlling for innovation by technologically related firms, as in Bloom, Schankerman, and Van Reenen (2013). We proxy for speculation by isolating bubbles episodes at the industry-year level following Greenwood, Shleifer, and You (2018).

During a bubble, the market-based private value of innovation increases by 30% at the patent level and between 40% to 50% at the firm level, corroborating the model's prediction that speculation increases private value. The effect is smaller for firms that are involved in more industries. Intuitively, an investor is likely to have a range of views about a firm's different products, hence reducing the effect of speculation about a given product on firm stock prices. The change in market-based private value is not accompanied by a corresponding change in outcome-based private value. In particular, we show that the increase in number of citations is small relative to the increase in market-based private value, reflecting little change in the quality of innovation in bubbles.

The data also support the prediction that speculation dampens market-based but not outcome-based measures of spillovers. In particular, the estimated coefficients suggest that during bubbles the business-stealing effect measured by asset prices completely vanishes, which occurs in our model asymptotically as the level of speculation increases. On the other hand, we do not see any empirical effect of speculation on the business-stealing effect measured through sales.

We obtain an additional set of predictions in a general equilibrium version of our model, which incorporates a decreasing returns to scale production technology that uses an input in fixed supply. The input market introduces two new spillovers—an *appropriability externality* due to the surplus from firm revenue accruing to workers and a *general equilibrium externality* from the impact of firm entry on equilibrium input prices. While the market-based measure of the appropriability externality is dampened by speculation, the general equilibrium externality is unaffected. As a result, the effect of industry characteristics on the market-based wedge between social and private value is reversed. In addition, while these industry characteristics are sufficient statistics for the wedge under agreement, with speculation the wedge depends on the microeconomic structure of the economy. Finally, speculation can reverse the sign of the wedge.

Our results also have normative implications. The market-based measure of social value coincides with the objective of a planner under the Pareto criterion, which respects the beliefs of households. On the other hand, the outcome-based measures of value capture the paternalistic criterion, under which the planner maximizes aggregate consumption net of entry costs. Data on forecast dispersion or portfolio holdings can help discipline such policy analysis by identifying and providing a finer measure of the speculation we consider in our model.

In Section 2, we introduce our model of speculation with business stealing. In Section 3, we derive predictions for private and social value, which we verify empirically in Section 4. We extend the analysis to general equilibrium in Section 5 and consider the normative implications of our model in Section 6. Finally, Section 7 concludes.

2 A Model of Disagreement and Innovation

We introduce a stylized framework to study the interplay of speculation and innovation. Innovation and competition are represented by a set of firms implementing new ideas. Speculation is represented by households who disagree about which firms will succeed. Households trade shares of the firms. The interplay between these two aspects yields equilibrium patterns that are consistent with stylized facts about bubbles. We then use the model to make testable predictions about the private and social value of innovation.

Figure XXX illustrates the overall structure of the model. At date 0, households work to create blueprints. They sell the blueprints to firm creators who implement them as firms. Households buy the claims to these firms' production. At date 1, firms compete and produce; households receive the payoffs from their positions in firms and consume. Let us detail each of these steps.

2.1 Firms

A continuum of firms, indexed by i and with total mass M_e, is created in equilibrium at date 0. At date 1, firms enter the production stage and their productivity a_i is revealed. To capture competition across these firms, we assume that only a fixed number M of the most productive firms is able to produce. Given the cumulative distribution function of productivities in the population F, only firms above a cutoff $\underline{a}$ are able to produce, with

$$\underline{a} := F^{-1}\left(1 - \frac{M}{M_e}\right). \tag{2.1}$$

The profits of a firm with productivity a are then given by

$$\pi(a) = a^{\eta} \cdot \mathbf{1}\left\{a \geq \underline{a}\right\}, \tag{2.2}$$

where η determines how differences in productivities translate into differences in generated profits, and the indicator function captures whether the firm produces or not.[3] We concentrate on situations where $M_e < M$, so that fewer firms produce than are created; the case where $M_e = M$ is straightforward.

This allocation of production slots is the key assumption to capture the notion of business-stealing. Indeed, firms do not internalize that they might take over the slot of another firm. The assumption of a fixed mass of production slots is especially plausible in industries that depend heavily on innovation. For instance, intellectual property law often provides exclusive use of a technology to its inventor.[4] We can interpret the fixed production slots as corresponding to a fixed number of M processes to produce the homogeneous good. The first firm to discover a process gets its exclusive use, and the speed of discovery is perfectly correlated with the productivity type a. Alternatively, we can assume that to produce, a firm needs one unit of an indivisible good that has not been discovered yet, and that only M of those exists in nature. Again firms with a higher type a find the ingredient faster.[5] Our assumption relies more generally on the scarcity in the ability to produce that is not internalized by individual firms. We show in Appendix C how our results hold for a wide range of models of business-stealing, some of them accommodating for endogenous variation in M.

2.2 Households

Households play two roles in the model. They work to create the blueprints for new firms. And, they speculate on wich of the new firms will succeed. Formally, there is a unit mass of households indexed by j. At date 0, household j is endowed with a fixed unit of consumption good c_0 and her share of firm creators, which we describe in the next section. In addition, each household decides how many blueprints to supply, b_j. Blueprints are produced at a convex cost $W(b_j) = f_e b_j^{\theta+1} M^{-\theta}/(\theta + 1)$, where the parameter θ controls the elasticity of supply of blueprints and f_e is the level of production costs. Finally, each household also decides on the number of shares to invest in each firm on financial market, $\{s_i^j\}_i$. Households have heterogeneous beliefs about the distribution of productivity a_i for each firm i. which we describe in detail below. We assume that they

[3] The fact that the marginal active firm collects positive profits improves tractability but is not crucial to our conclusions. We show in Appendix A MORE PRECISE REF that our results also hold in a variant of the model where the marginal firm earns zero profits.

[4] Another motivation for the incompleteness is the difficulty to establish markets for what has not been encountered yet. Indeed it is often the case that nobody owns something before it is discovered. For instance, how could we trade nuclear power before Marie Curie discovered radioactivity?

[5] Network goods or industries facing institutional constraints can face similar frictions in the allocations of productive positions that lead to a business-stealing effect—see Borjas and Doran (2012) for evidence in the context of scientific research.

can only take long positions in claims to firms.[6]

Households behave competitively and take prices as given. Hence, household j solves the problem

$$\max_{c_0, s_i^j \geq 0, b_j} \; c_0 + \mathbf{E}^j \left\{ \int s_i^j \pi_i di \right\} - W(b_j) \tag{2.3}$$

$$\text{s.t.} \;\; c_0 + \int s_i^j \; p_i \; di \leq 1 + p_b b_j + \Pi, \tag{2.4}$$

where p_i and p_b are the respective prices of firm i and blueprints, $\mathbf{E}^j$ is household j expectation and Π denotes firm creators' aggregate profits.[7]

2.3 Beliefs

To capture the notion of speculation, we assume that households disagree about which firms will be successful. Disagreement arises naturally in innovative episodes: confronted with new ideas and firms, households must rely on their priors to evaluate them. In general, it is challenging to keep track of heterogenous beliefs about an entire set of firms. We overcome this issue by imposing some structure on the distribution of beliefs. We assume that even though agents disagree about each individual firm, they agree on the population distribution of firm productivity. We assume the population distribution F follows a Pareto distribution: $F(a) = 1 - a^{-\gamma}$ for $a \geq 1$.[8]

A simple narrative for our specification of individual beliefs is as follows. Each household organizes firms into a continuum of packets containing n firms each, and believes she knows the exact ranking of productivity draws within each packet. We assume that the composition of packets and the order of firms within packets is drawn in an i.i.d. equiprobable fashion across agents and firms, and that each firm can only be in one packet. The parameter n controls the intensity of disagreement. When $n = 1$, households consider all firms to be the same, with their productivity drawn from F. As n increases, households can compare more firms in each packet, and thus have a stronger prior that the best firm in each packet will have a high productivity.

[6]Such a hard constraint facilitates the analysis, but the important assumption is some limit or cost to the ability to take short positions.

[7]The assumption of risk neutrality does not play any role in the analysis. Formally, all of our results would be exactly identical if the objective of households was: $c_0 + \mathcal{U}^{-1}\left(\mathbf{E}^j \left\{\mathcal{U}(\int s_i^j \pi_i di)\right\}\right) - W(b_j)$, with $\mathcal{U}(.)$ an increasing concave function.

[8]A motivation for focusing on disagreement across firms rather than about the aggregate is the empirical observation that high firm entry in a sector often follows disruptive innovation, either through a large technological change or the introduction of new products. New firms then conduct micro innovations to take advantage of a macro innovation (Mokyr (1992)). Empirically, these episodes appear particularly relevant to economic growth, as discussed for instance in Abernathy (1978) and Freeman (1982).

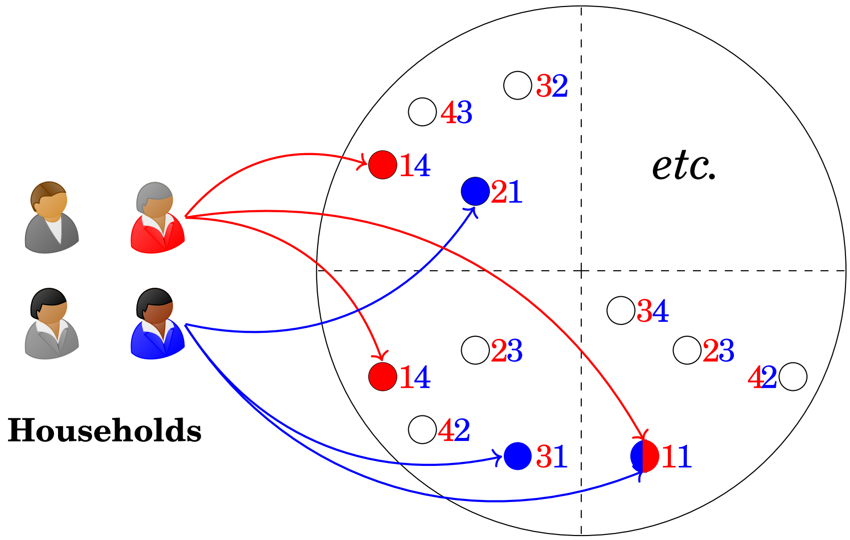

Figure 1
Firms and households beliefs for $n = 4$.[9]

In equilibrium, household j only invests in the subset of firms that she considers to be the most productive in each of their respective packets. These firms are perceived by household j to have productivity drawn from F^n, the distribution of the maximum of n independent draws from F. Since households rank firms differently, they have different beliefs about the productivity distribution of any given firm and invest in different sets of firms, as illustrated in Figure 1. Each household believes that the firms she invests in are, in expectation, more productive than the average firm in the economy.

Our analysis does not rely on the specific narrative of packet formation, even though we find it intuitively appealing. The crux of our assumptions is to reduce the distribution of beliefs to two parameters: the actual population distribution F and a single parameter n for the intensity of disagreement. Then, two features of our assumptions facilitate the analysis. First, while households disagree on which firms will succeed, they agree on the population distribution of firms. Hence, they agree on aggregate outcomes: both the threshold $\underline{a}$ and market conditions, which will play a role in the richer settings in Section 5. Second, beliefs are symmetric across households. This overcomes the issue of having to keep track of the entire distribution of beliefs and allocations.

[9]In this figure, the composition of each packet is identical across agents for exposition purposes and does not affect results.

2.4 Firm creators

Finally, firm creators connect the step of innovation and trading in financial markets. They pay households to create blueprints, and issue claims to the corresponding firms on financial market. Formally, there is a continuum of short-lived firm creators. At date 0, each firm creator can use a unit blueprint to create a new firm, which is then sold on competitive financial markets. They participate in competitive markets for blueprints and firms, taking their respective prices p_b and p_i as given.[10] The firm creator problem at time $t = 0$ is therefore:

$$\max_{c \in \{0,1\}} c \cdot (p_i - p_b). \tag{2.5}$$

2.5 Competitive equilibrium

The competitive equilibrium of the economy is defined as follows. Firm creators maximize profits from selling their firms taking prices as given. Households maximize their perceived expected utility by choosing their optimal blueprint discovery effort and a dynamic optimal portfolio allocation, taking the price of blueprints and of firms as given. Firms maximize profits given their production status. Finally, the markets for blueprints and claims to firms profits (the stock market) clear:

$$\int b_j dj = M_e, \tag{2.6}$$

$$\forall i \in [0, M_e], \qquad \int s_i^j dj = 1. \tag{2.7}$$

Combining the equilibrium conditions yields a single equation determining the quantity of firm entry M_e:

$$W'(M_e) = V^{(n)}(M_e) = \int_{\underline{a}}^{\infty} \pi(a) dF^n(a). \tag{2.8}$$

In equilibrium, the marginal cost of creating an additional firm, $W'(M_e)$, is equal to the expected profits to an investor who favors it, $V^{(n)}(M_e)$. Both of these quantities are equal to the price at which blueprints and firms trade in the economy, $p_i = p_b$. Therefore, this condition also pins down the price of new firms. We define this value, which will play a prominent role in our analysis by $V^{(n)}(M_e) = \mathcal{I}_n(M_e, \eta)$, where:

$$\mathcal{I}_n(M_e, \eta) := \int_{F^{-1}(1-M/M_e)}^{\infty} a^\eta dF^n(a). \tag{2.9}$$

[10]We assume firm creators do not have any information about the firms they create.

2.6 Bubbles

Our framework yields an equilibrium behavior consistent with features of bubble episodes related to the introduction of new technologies. Some well-documented examples include railroads, electricity, automobiles, radio, micro-electronics, personal computers, bio-technology, and the Internet. Scheinkman (2014) and Brunnermeier and Oehmke (2013) survey the field of speculation and bubbles; Janeway (2012) gives a first-person account of the relationship between innovation and speculation.

In the model, when investors disagree more, the subset of investors which value a given firm the most values it more. When n increases, the demand for firms $V^n(.)$ shifts up. This higher demand from financial market participants pushes the price of firms p_i up and also leads to more new firms M_e. This reflects the boom phase of a bubble: many new innovations are implemented and firms created, and prices on financial markets are large. Subsequently, however, not all investors can be right. Output is below what is implied by prices and subsequently prices. Formally, firms are priced at $\mathcal{I}_n$, under the distribution F^n, but the average output is only $\mathcal{I}_1$, under the population distribution F. This drop corresponds to the bust phase of a bubble. Following Greenwood, Shleifer, and You (2018), the abnormal price increase in the first phase of a boom is what we will use empirically to distinguish bubble periods in our study of the value of innovation. While they also show that conditioning on a bust helps capture more features of bubbles, the ex-post nature of this conditioning would be an issue for our empirical tests.

A more formal definition of a bubble is a situation in which asset prices exceed an asset's fundamental value (see, for example, Brunnermeier (2017)). The drop in prices at date 1 is expected by all investors — they agree on aggregate outcomes — and supports this definition. How can all investors agree that the market portfolio is overpriced? Such a situation arises from households' heterogeneous beliefs. Different households view different firms as the most valuable, and specialize their portfolios in those firms, which seem fairly priced to them. The short-sale constraint prevents each household from shorting the other firms, which she views as overpriced.[11] Chen, Hong, and Stein (2002), Diether, Malloy, and Scherbina (2002), or Yu (2011) establish empirically the link between dispersed beliefs and low future returns.[12]

[11] Van den Steen (2004) describes how disagreement combined with optimal choice leads to overvaluation and Miller (1977) first pointed out the importance of short-sale constraints in financial markets.

[12] Using investor survey data to quantify disagreement, Diether, Malloy, and Scherbina (2002) find that stocks with dispersed analysts forecasts experience low subsequent returns. Yu (2011) aggregates this measure to portfolios such as the market and finds a similar result. Using stock market positions to measure disagreement, Chen, Hong, and Stein (2002) construct the fraction of the mutual fund population investing in a given stock, a measure of the breadth of ownership for individual stocks. They find that this measure predicts low stock returns.

Our model takes a static view of speculation, but in practice investors' views change over time. In Appendix XXX, we study an extension of the model which entertains this possibility. A stronger form of overvaluation arises: the price of each firm exceeds the maximum valuation of its cash-flow by any specific investor in the economy. This difference comes from the fact that when investors change views, the current investor of a specific firm will typically not favor it anymore, and sell it to somebody else. Hence, each time households trade firms signals a change in who values them most, a mechanism similar to the models of Harrison and Kreps (1978) and Scheinkman and Xiong (2003). We show that this dynamic overvaluation is increasing in volume per period and the length of the bubble. Historically, abnormally high trading volume is seen as a hallmark of bubbles — see Scheinkman (2014) for a survey. For example, during the Roaring Twenties, daily records of share trading volume were reached ten times in 1928 and three times in 1929, with no new record set until 1968 (Hong and Stein, 2007). More recently, during the DotCom bubble. Internet stocks had three times the turnover of similar cases. Greenwood, Shleifer, and You (2018) also document that large stock price increases are more likely to end in a crash when they are accompanied by increased trading volume.

Finally, while our model does not provide an explicit premises for increases in disagreement, it is natural to expect them to happen around innovative episodes. As investors are seeing these ideas and firms for the first time, they must rely on their priors to evaluate them. In contrast, more mature industries are likely associated with more common information investors might have accumulated over time, and stronger agreement about what makes a firm successful. Consistent with this view, Greenwood, Shleifer, and You (2018) find that when the price run-up in an industry occurs disproportionately among the younger firms, crashes and low future returns are more likely.

All these observations suggest that the model is a useful representation of the episodes we are interested in. They are comforting as we turn to using the model to make new predictions, about the value of innovation.

3 The Value of Innovation

We are interested in two notions of the value of an innovation: the private value, which accrues to its investors, and the social value, which is its impact on the economy overall. Understanding these two notions is important as they shed light on how the process of innovation alters the economy, and ultimately create growth. We study two approaches to measure these values: *market-based* measures of value that use asset prices and *outcome-based* measures of value that use real outcomes such as patent citations or output. This leads us to a set of predictions on how these quantities change during bubbles, predictions we test in Section 4.

3.1 Private value

The private value of a firm is the value of that firm to its investors. Consider first the market-based measure of this value. Empirically, the change in stock price of a firm following an innovation, as measured by Kogan et al. (2017), reveals the market-based private value. In the model, the market-based private value is simply the price $p_i = \mathcal{I}_n(M_e)$ at which the firm trades. This is because because an innovation and a firm coincide for simplicity.[13] As we have discussed in the previous section, during bubble episodes (large n), this value increases.

The outcome-based counterpart to this metric is the actual effect of this innovation. In the data, it can be measured by the number of citations a patent receives, or changes in sales following the introduction of an innovation. In the model, the realized output of the firm is the outcome-based private value. On average, the output of a firm is given by $\mathcal{I}_1(M_e)$, and all investors agree about this. In the absence of disagreement, $n = 1$, market-based value and outcome-based private value coincide. However, as n increase, the two values diverge and the ratio $\mathcal{I}_n(M_e)/\mathcal{I}_1(M_e)$ increases. In other words the increase in market value due to speculation does not reflect a commensurate increase in fundamentals.

These two predictions are the two defining properties of bubbles applied to innovation. The market value of an innovation increases during a bubble. This increase is not justified by a change in outcomes. Though, as we now turn to social value, we will see that the model draws a more subtle picture than a naive theory which would simply states that all valuations are higher than usual.

3.2 Social value

The social value of a firm is the effect of introducing this firm on the whole economy. A new firm in the economy has a direct effect (its private value) and an indirect effect because it affects the value of other firms. To characterize how social value departs from the private value we introduce the measure *spillover*. We define *spillover* as the indirect effect of a new firm scaled by its private value:

$$\textit{spillover} = \frac{\text{social value} - \text{private value}}{\text{private value}} \tag{3.1}$$

Empirically this definition coincides with estimating the elasticity of firm value to the unexpected entry of another firm, closely related to Bloom, Schankerman, and Van Reenen (2013). Similarly to the measure of private value introduced above, we consider both *market-based spillover* using asset prices and *outcome-based spillover* using output.

In the model with agreement, we find no distinction between the market- and outcome-based *spillover*, since the price of claims to a firm in period 0 is equal

[13] In our empirical exercise, we will come back to the distinction between innovation and firm.

to the mean firm profits in period 1. In contrast, speculation introduces a gap between the two spillover measures, leading to novel empirical predictions.

3.2.1 Spillover in the model

In the model spillovers arise because of the business-stealing effect. When a new firm enters it might take one of the production slots and displace an existing firm.

Computing the spillover in the model is challenging because the baseline model of Section 2 does not feature unexpected entry: the number of firm is deterministic in equilibrium. We overcome this challenge by introducing the possibility that some blueprints randomly fail to be implemented. Formally, after firm creators purchase blueprints but before they introduce these blueprints to public markets, a fraction of them disappear. If M_e blueprints are created, either they all succeed or a mass Δ fails, with probability $1-\varepsilon$ and ε respectively. By focusing on the limit when ε and Δ go to zero, we trace out the equilibrium effect of the unanticipated entry of an atomistic firm on the total value of the economy.

Comparing the total market value of the economy across the two outcomes, we obtain the market-based social value of an extra firm:

$$social\ value_{market} = \lim_{\Delta\to 0} \frac{M_e V^{(n)}(M_e) - (M_e-\Delta)V^{(n)}(M_e-\Delta)}{\Delta} \tag{3.2}$$

$$= V^{(n)}(M_e) + M_e V^{(n)\prime}(M_e) \tag{3.3}$$

For the outcome-based social value, simply replace n by 1 in this expression. We recognize the two effects of introducing a new firm. The first term is the direct effect: the private value of a new firm $V^{(n)}(M_e)$. The second term is the change in the value of all other firms in response to the entry of the new firm. Each of the M_e firms in the economy sees its value decrease by $V^{(n)\prime}(M_e)$. Taking the ratio of the direct and the indirect effect, we obtain the two measures of spillover:

$$spill_{mkt} = \frac{M_e V^{(n)\prime}(M_e)}{V^{(n)}(M_e)}, \tag{3.4}$$

$$spill_{out} = \frac{M_e V^{(1)\prime}(M_e)}{V^{(1)}(M_e)}. \tag{3.5}$$

We recognize the elasticity of the value of a firm to the number of firms, $\mathcal{E}_{\mathcal{I}_n}$.[14]Again, this elasticity is negative because of the business-stealing effect. The entry of a new firm makes all other firms less likely to produce.

In the remainder of this section, we derive properties of the spillover measures. In particular, we relate them to the intensity of speculation. Before doing

[14]Throughout the paper, we denote the elasticity of quantity X to firm entry M_e by $\mathcal{E}_X = d\log(X)/d\log(M_e)$.

so, it is worth pointing out a deeper interpretation of the spillover. The spillover measures the intensity of entry externalities in the economy, and as such corresponds to a Pigouvian tax on entry. We come back to this implication at length in Section 6. For now, just note that for a non-paternalistic planner, the expression for the optimal tax rate is the opposite of the spillover: $\tau = -spill_{\text{mkt}}$.

3.2.2 Outcome-based spillovers

The real spillovers are the elasticity of average output of a firm to the number of firms in the economy. We compute explicitly $spill_{out}$ from the expression of equation (3.5), where we replace the output of a firm with $\mathcal{I}_1$:

$$spill_{out} = -\frac{\displaystyle\int_{\underline{a}}^{\infty} \pi(\underline{a}) dF(a)}{\displaystyle\int_{\underline{a}}^{\infty} \pi(a) dF(a)} = -\frac{\gamma - \eta}{\gamma}. \tag{3.6}$$

The ratio of the two integrals has an intuitive explanation. The numerator represents the expected amount of profits displaced by the introduction of a new firm. It is the product of the output of the marginal producing firm (the firm that will get displaced) with the probability of such firm does get displaced, $\int_{\underline{a}}^{\infty} dF(a)$. The denominator is the expected output of a new firm.

Displacement lowers aggregate output, thus $spill_{out}$ is negative. However, displacement only occurs when the new firm is more productive than an existing firm, thus $|spill_{out}| < 1$. There are less spillovers if the distribution of firm productivity is more spread out, lower values of γ, or when productivity differences translate into larger output differences, larger values of η.

Interestingly the outcome based spillover $spill_{out}$ does not depend on the equilibrium number of firms M_e. Technically, this result is the consequence of specifying a power production function, and a Pareto productivity distribution.[15] As the level of disagreement changes, outcome-based spillovers stays constant. This prediction provides a useful benchmark against which to evaluate the behavior of market-based spillovers. In Section 4, we confirm this prediction empirically, thereby validating our framework.

3.2.3 Market-based spillovers

Market-based spillover is the elasticity of the value of a firm to the number of firms in the economy. In the case of agreement ($n = 1$), market values and average outputs coincide; therefore market-based and outcome-based measure of spillovers are identical. However with disagreement the two measures diverge, and $spill_{\text{mkt}}$ is not constant.

[15]Under a Pareto distribution, the ratio between marginal and average productivity is independent of the lower cutoff.

Proposition 1. *Speculation lowers the intensity of market-based spillovers:*

$$spill_{mkt}(n) < spill_{mkt}(1), \quad \textit{for } n > 1 \tag{3.7}$$

Under disagreement, value spillovers are smaller despite the level of firm entry being higher. This seeming contradiction arises because the beliefs of households impact their valuation of equilibrium allocations. Under disagreement, households only invest in their favorite firms. Therefore, each household places a lower probability on being displaced by new entrants, thereby reducing the effect of the business-stealing externality on the market prices that determine the value spillovers.

More formally, we can rewrite the market-spillover from (3.4) in its integral form:

$$spill_{\mathrm{mkt}}(n) = \frac{M_e V^{(n)\prime}(M_e)}{V^{(n)}(M_e)} = -\frac{\displaystyle\int_{\underline{a}}^{\infty} \pi(\underline{a}) \frac{F_n'}{F'}(\underline{a})\, dF(a)}{\displaystyle\int_{\underline{a}}^{\infty} \pi(a) \frac{F_n'}{F'}(a)\, dF(a)}. \tag{3.8}$$

This ratio compares the value of displaced firms to the value of the firm displacing them, evaluated through the beliefs of their respective investors — which are different. Holding M_e and thus $\underline{a}$ constant, consider how this ratio changes with n. The numerator is the expected value of a displaced firm from the point of view of its owner. Disagreement affects this value through the change in probability weights F_n'/F' *at the threshold* $\underline{a}$. In contrast, the denominator is the profit of an average firm from the point of view of its owner. This quantity is affected by changes in the probability weights throughout the distribution *above the threshold*. Because increasing n corresponds to shifting the perceived distribution of productivities to the right, the change in probability weights is increasing as we move to higher productivities (see Appendix Figure A.1). Such an increase affects expected profits more strongly than the value of displaced firms, decreasing the wedge. In Appendix A, we show that this result holds more generally with minimal assumptions on the productivity distribution $F(\cdot)$ and profit function $\pi(\cdot)$.

This proposition gives rise to a clear prediction we test in the data: the market-based spillovers are lower in speculative periods. Following Bloom, Schankerman, and Van Reenen (2013), we can measure these spillovers by gauging the reaction of firms' valuation to variation in the amount of innovation by their competitors. In addition, we can also compare the market-based spillover to outcome-based spillovers. An immediate corollary of Proposition 1 is that the presence of speculation reduces market-based spillovers relative to outcome-based spillovers. These results of our model of bubbles differ from naive theories. For example, one common view is that bubbles inflate the price of all firms proportionally, either because people use too low a discount rate or overestimate the output of all firms. Then, because the market-based spillover is a ratio of valuations, it is unaffected and remains equal to the output-based spillover.

High speculation limit. We now consider the limiting case of high speculation: $n \to \infty$. This case is extreme: the total quantity of entry goes to infinity. However, it is useful because it gives rise to sharp characterizations of the spillovers, and allows us to highlight the distinction with the agreement case.

Proposition 2. *In the high disagreement limit ($n \to \infty$), the market-based spillover converges to a finite limit that depends on the sign of $\gamma\theta - \eta$:*

- *If $\gamma\theta > \eta$, the market-based spillover vanishes:* $\lim_{n\to\infty} spill_{mkt}(n) = 0$
- *If $\gamma\theta < \eta$, then τ converges to the spillover in the agreement case ($n = 1$):* $\lim_{n\to\infty} spill_{mkt}(n) = \frac{\gamma-\eta}{\gamma}$
- *In the knife-edge case of $\gamma\theta = \eta$,* $\lim_{n\to\infty} spill_{mkt}(n) = \check{spill}_{mkt} > -\frac{\gamma-\eta}{\gamma}$*, where* $\check{spill}_{mkt}$ *is defined in Appendix equation* (A.15).

Figure 2 illustrates these cases. Appendix A shows that the results hold in more general models of business stealing.

Two forces determine the asymptotic behavior of the market-based spillover. First, with more disagreement, investors increasingly believe that the firms they invest in are in the right tail of the productivity distribution. They are therefore less concerned about the risk of being displaced by new entrants, as they expect that a smaller mass of firms in their portfolio will fail to meet the entry threshold. For a given level of entry M_e, this mass converges to 0 as n goes to infinity. This is an extreme case of the result in Proposition 1.

Second, disagreement increases firm entry. For a given level of disagreement, n, as M_e converges to infinity, the M producing firms end up in the tail of both the population distribution, F, and the favorite-firm distribution, F_n. The tails of these two distributions have the same shape since $\lim_{x\to\infty} F_n'(x)/F'(x) = n$.[16] Therefore disagreement does not affect the relative position of the marginal and average valuation in the tail. This force brings the market-based spillovers back towards the level of real spillovers.

The relative strength of the forces depends of how fast firm creation increases with speculation. If $\gamma\theta > \eta$, the first force dominates. When θ is large, the marginal cost of firm creation rises more rapidly, which reduces equilibrium entry M_e and weakens the second force. As γ increases, we have a thinner tailed firm productivity distribution. The size of the tail becomes less important than the relative ordering of firms, which weakens the second force. As η decreases,

[16]This Extreme Value Theory result is not the byproduct of power distributions, but rather applies to a larger class of distributions.

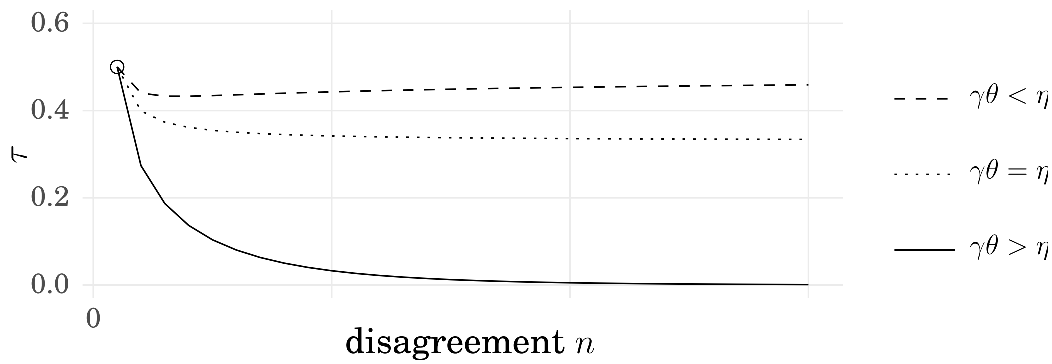

Figure 2
Entry wedge with increasing disagreement.

profits increase less with productivity, which again diminishes the importance of the second force. In the data, we find that the valuations increase faster than quantity during bubbles, consistent with the case of $\gamma\theta > \eta$.[17] The relevant prediction is therefore a decrease in spillovers towards 0.

4 Empirical Evidence

We now take the three main predictions of the model to the data. First, the market-based measure of private value of innovation increases with speculation. Second, market-based spillovers decrease in the presence of speculation. Finally, outcome-based measures respond relatively less, or not at all.

4.1 Data

To test these predictions, we combine variation in whether an industry experiences a bubble of not with measures of the private and social value of innovation at the firm and patent level. For each of these dimensions, we follow prominent work from the literature, though combining these aspects is novel to our paper. Summary statistics of the main regression variables are contained in Table 1.

Bubbles. We proxy for the presence of speculation using the empirical definition of bubbles from Greenwood, Shleifer, and You (2018). We split firms into

[17]Specifically, we find in the next section that the valuation of innovation increases by 40%during bubbles while the quantity of innovation increases only by 15%, so $\theta > 1$. Because $\eta < \gamma$ is always satisfied, we have $\gamma\theta > \eta$.

Table 1
Summary Statistics

	N	Mean	Std. Dev.	25th pct.	Median	75th pct.
Bubble						
Bubble periods dummy	2,734	0.0271	0.162	0	0	0
Value of Innovation (KPSS)						
Patent level value						
Stock Market ($ Mn)	1,171,806	14	37.8	2.32	5.46	12.7
Citations (fwd. looking)	1,171,806	12	22.6	2	5	13
Firm level value						
Stock Market ($ Mn)	47,887	232	1880	0.519	3.06	24.5
Citations (fwd. looking)	47,887	52.9	236	2.88	7.8	26.7
Firm level statistics						
Mkt. Cap. ($ Mn)	47,887	2854	14667	47.2	200	949
Segments (# NAICS-4)	53,066	1.33	0.646	1	1	2
Segments (# NAICS-6)	53,066	1.41	0.729	1	1	2
Measuring Spillovers (BSvR)						
Firm Outcomes (real or market valued)						
Sales ($ Mn)	9,382	3563	12626	135	509	2037
Tobin's q	9,382	2.46	3.09	0.86	1.49	2.71
Measures of Spillovers (Jaffe)						
Technology	9,382	9.8	1.02	9.34	9.97	10.5
Competition	9,382	7.32	2.35	6.33	7.64	9.01
Measures of Spillover (Mahalonobis)						
Technology	9,382	11.4	0.821	10.9	11.5	11.9
Competition	9,382	8.53	1.73	7.87	8.77	9.74

Note: Table 1 presents summary statistics of the main variables included in the regression specifications. The bubble dummy corresponds to bubble detected across Fama-French 49 industries according to the methodology outlined in Greenwood, Shleifer, and You (2018). The value of innovation both at the firm and patent level is directly taken from Kogan et al. (2017). The stock market value of innovation at the patent level corresponds to the appreciation in the value of a firm issuing a patent around the patent issuance date. The stock market value of innovation at the firm level corresponds to an annual aggregation of the total value of all patents issued by a firm in a given year. Both the patent and firm level citation value of a patent corresponds to its forward looking number of citation (until the end of the sample in 2010). Other firm level statistics correspond to the CRSP-Compustat merged file (for market capitalization, sales and Tobin's q) and to the Compustat segment file. Tobin's q is measured from Bloom, Schankerman, and Van Reenen (2013) as the market value of equity plus debt divided by the stock of fixed capital. We obtain measures of technological and competition spillovers from Bloom, Schankerman, and Van Reenen (2013), corresponding to the distance between the technological class of patents issued by a firm with other public firms and to the distance between the set of product market of a firm and other public competitors.

49 industries following the classification Fama and French. An industry-month is defined as being in a bubble if it satisfies simultaneously three conditions. First, the value-weighted portfolio of the corresponding industry experienced a return of 100% or more over the previous two years. Second, this industry value-weighted return also exceeds the return of the market by at least 100% over the past two years. Third, the industry value-weighted return over the past five years is larger than 50%. We aggregate back to the industry-year level — the coarseness of the remainder of our data — by considering an industry-year in a bubble if at least a month is in a bubble. In practice, multiple months are always in a bubble year due to the persistence of the bubble criteria. This approach identifies 74 industry-years in a bubble between 1962 and 2017.

Consistent with our model, Greenwood, Shleifer, and You (2018) find that "price run-ups...involving younger firms (and) having higher relative returns among the younger firms...(are) more likely to crash". These industries with young and innovative firms are likely to have scarce information, making them more susceptible to the disagreement that we model. We also confirm the validity of the bubble classification empirically and show that it predicts increases in innovation like the model. In Appendix Table E.1, we show that there are 1.4 more patents issued within a USPTO technology class during bubbles, which translates to a 15% increase in the number of patents created in a patent class during a bubble.

Private value of innovation. We use the measures of private value of innovation from Kogan et al. (2017). Their dataset combines stock market and patent data for U.S. firms for the period from 1926 to 2010. They measure the stock market response in the three-day window after a firm has a new patent issued, controlling for the return on the market portfolio during that period. This number is the direct counterpart to the market-based measure of private value of a new innovation in our framework. They also aggregate the stock market value from all the patents of a given firm every year to measure the value of innovation at the firm level. To capture the ultimate quality of each patent, Kogan et al. (2017) use the forward-looking number of citations generated by a patent —or the number of citations generated by all the patents produced by a firm in a given year. This corresponds to the outcome-based measure of the private value of a new innovation in the model.

One question is whether to focus on the patent-level or firm-level data. Our baseline model is silent on this issue because each firm corresponds to exactly one blueprint. In our empirical analysis, we consider both approaches.

Social value of innovation. For the social value of innovation, Bloom, Schankerman, and Van Reenen (2013) identify spillovers from different sources of firm interactions. They regress a firm-level outcome — log market value or log future sales — on the quantity of innovation by groups of "neighboring" firms. Because the set of close competitors and of close innovators do not coincide, this approach separates competitive spillovers arising from firms in neighboring industries and knowledge spillovers coming from firms issuing patents in the same technology USPTO class. Specifically, they first construct distances between firms in each of these two spaces. Then, for each firm-year, they compute distance-weighted stocks of innovative capital from all other firms. The resulting firm-level exposures are *spillsic* and *spilltech*, respectively, and regressions on these quantities measure competitive and innovative spillovers. Because our baseline model only considers competitive interactions — Section 5 extends to other sources of interactions — we are particularly interested in competitive spillovers. Still, we also control for innovative interactions. When the left-hand-side of the regression is log market value, this approach identifies the market-based spillover. When focusing on log sales, the coefficients identify the outcome-based spillover. Taking the log ensures that we are considering a semi-elasticity, the counterpart to the spillover in the model.

Bloom, Schankerman, and Van Reenen (2013) propose a variety of metrics for distance, and we follow their construction. Proximity between two competing firms in the product market space is the correlation of the firms' distribution of sales across their industry segments. Technology proximity is analogously defined as the correlation of patent USPTO technology classes between firms, following earlier work by Jaffe (1986). The quantity of innovation by competitors and close innovators are the stock of innovation weighted by these correlations. The stock of innovation is constructed using a perpetual inventory approach. Both measures are extended using the Mahalanobis distance to allow for flexible weighting of the correlation between firms across different technology classes or product market classes. Further details on the measures of spillovers can be found in Appendix E and in Bloom, Schankerman, and Van Reenen (2013).

4.2 The Private Value of Innovation During Bubbles

The market-based private value of innovation increases in bubbles. To study the effect of speculation on the private value of innovation at the patent level, we consider the following regression specification:

$$\log \xi_{j,t} = \beta B_{j,t} + \gamma Z_{j,t} + \varepsilon_{j,t}, \tag{4.1}$$

where $\xi_{j,t}$ is the private value of patent j issued during year t using the market-based measure (from Kogan et al. (2017)). The variable $B_{j,t}$ is an indicator for whether the firm issuing patent j was in an industry experiencing a bubble during year t. As in Kogan et al. (2017), the controls $Z_{j,t}$ include log of the number of citations for the patent, market capitalization and year dummies. We take the market capitalization lagged by one year because the value of a firm's stocks tends to rise during a bubble. The lagged market capitalization variable controls for the size of the firm without contaminating our estimate of how the private value varies with speculation. To ensure that our results are not driven by the type of firms or industries that go through bubbles, we include firm fixed effects. We run a similar regression at the firm level, replacing the private value and citation variables with their firm level analogs. We no longer control for market capitalization because the firm level measure of private innovation value is normalized by book value. Following Kogan et al. (2017), we include year fixed effects in this specification.

Table 2
Private value of innovation in bubbles

	Patent Level			Firm Level		
	(1)	(2)	(3)	(4)	(5)	(6)
Bubble dummy	0.317***	0.315***	0.306***	0.514***	0.427***	0.427***
	(0.094)	(0.094)	(0.096)	(0.114)	(0.123)	(0.080)
Log Citations		0.016***	0.023***		0.823***	0.716***
(forward looking)		(0.004)	(0.005)		(0.011)	(0.009)
Log Market Cap			0.562***			0.626***
(lagged)			(0.028)			(0.020)
Fixed Effects	Y, F	Y, F	Y, F	Y, F	Y, F	Y, F
Observations	1,171,806	1,171,806	1,169,860	47,886	47,886	47,484
R^2	0.68	0.68	0.74	0.89	0.94	0.96

Note: Table 2 presents panel regressions of the value of innovation, as measured in Kogan et al. (2017) at the patent and firm levels, on a dummy from Greenwood, Shleifer, and You (2018) that captures whether the firm is an industry that is in a bubble or not. We control for the lagged market capitalization of the firm and include fixed effects for firm F and patent grant year Y. Standard errors clustered at the grant year level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 2 shows that speculation indeed increases the private value of innovation. In particular, the presence of a bubble increases the private value of innovation by approximately 30% at the patent level and 40% to 50% at the firm level. These effects are both economically and statistically significant.

We stress that an increase in the market-based measure of the private value of innovation in a bubble is not mechanical. We measure bubbles as times of high stock market valuations, which do not necessarily coincide with a strong positive response to the news of issuing a patent. This positive response is a direct implication of our model where bubbles arise in times of high private value of innovations. In contrast, the common view that bubbles are episodes in which prices are completely disconnected from fundamentals would not make such a prediction. Under this view, valuations could be high overall, but would not be responsive to patent issuance.

Diversity dampens the effect of speculation. In our model one firm coincides with a single blueprint. In the data, we have already discussed that firms can issue multiple patents during the same year. This is why we presented results both at the patent level and the firm level. However, there is an additional source of variation not captured by our simple theory: some firms operate in multiple sectors. This variation naturally interacts with our mechanism: the effect of a bubble should be smaller for firms that have more diversified activities. Indeed, narrow firms, as opposed to firms that span multiple product lines, cater to specific investors which tend to have strong beliefs in their success. Thus, we expect a "conglomerate discount" on the value of innovation in times of bubbles: in the presence of disagreement or bubbles, multiple product firms experience a smaller increase in the value of their innovation than narrow firms.

Table 3

Diversity and Private Value of Innovation in Bubbles

	Patent Level			Firm Level		
	(1)	(2)	(3)	(4)	(5)	(6)
Bubble x Segments	-0.612^{***}	-0.600^{***}	-0.533^{***}	-0.473^{***}	-0.402^{***}	-0.307^{***}
(NAICS 4 digits)	(0.157)	(0.155)	(0.148)	(0.109)	(0.080)	(0.069)
Bubble	1.471^{***}	1.441^{***}	1.335^{***}	1.576^{***}	1.431^{***}	1.180^{***}
	(0.198)	(0.192)	(0.205)	(0.267)	(0.323)	(0.288)
Segments	0.309^{***}	0.305^{***}	0.296^{***}	0.062	0.015	0.014
(NAICS 4 digits)	(0.097)	(0.096)	(0.100)	(0.043)	(0.045)	(0.037)
Log Citations		0.047^{***}	0.044^{***}		0.044^{***}	0.044^{***}
(forward looking)		(0.010)	(0.009)		(0.009)	(0.009)
Log Market Cap			0.154^{***}			0.286^{***}
(lagged)			(0.040)			(0.046)
Fixed Effects	Y, F	Y, F	Y, F	Y, F	Y, F	Y, F
Observations	180,636	180,636	177,911	10,426	10,426	10,256
R^2	0.72	0.72	0.72	0.88	0.93	0.94

Note: Table 3 presents panel regressions of the value of innovation, as measured in Kogan et al. (2017) at the patent and firm levels, on a bubble dummy interacted with the number of industries spanned by the different segments of the parent firm. The bubble dummy is from Greenwood, Shleifer, and You (2018) and captures whether the firm is an industry that is in a bubble or not. Compustat segments are measured at the four digit NAICS code level from the COMPUSTAT Segments file. We control for the forward looking number of citations generated by a patent (or firm) from Kogan et al. (2017), and the lagged market capitalization of the firm. We include fixed effects for firm F and patent grant year Y. Standard errors clustered at the grant year level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

To test this prediction empirically, we measure diversity with the variable $Segments_{j,t}$, which is the number of 4-digit NAICS industries that a firm is active in. We collect this information from the COMPUSTAT Segments files. We augment regression (4.1) with $B_{j,t} \times Segments_{j,t}$.[18] As predicted, we find in Table 3 that the greater the number of segments, the less the private value of innovation increases in a bubble, as indicated by the significantly negative coefficient on the $B_{j,t} \times Segments_{j,t}$ regressor. Moreover, we confirm that the effect of a bubble on the value of a patent for a firm with more than one segment is not statistically significant. Qualitatively, we find the same results whether we run the regressions at the patent or the firm level.

Innovation quality does not see a commensurate increase in bubbles. Another prediction of our theory is that the market-based private value of innovation increases during bubbles *relatively* to the outcome-based private value. In the model, this comparison is particularly sharp because firm productivity is drawn from the same distribution F in date 1 regardless of the level of disagreement, thus innovation quality does not depend on speculation. We now show empirically that there is in fact no significant increase in innovation quality, the outcome-based measure of the value of innovation, during bubbles.

In particular, we consider:

$$\log\left(1 + C_{j,t}\right) = \beta B_{j,t} + \gamma Z_{j,t} + \varepsilon_{j,t}, \tag{4.2}$$

where $C_{j,t}$ is the forward-looking number of citations received by a patent j issued in year t. The controls $Z_{j,t}$ include the lagged market value of the firm, and a year and firm fixed effect as in equation (4.1). As before, we also run the firm level analog.

Table 4 reports the results: the quality of innovation sees a slight increase during bubbles. The number of citations is 10% higher in times of bubbles. This increase arises partly from the higher number of patents issued during bubbles.

To understand why the increase in citations is small relative to the increase in the market value of innovation in bubbles from Table 2, we need to translate these two sets of estimates in the same units. We convert citations into dollars by reproducing the estimates from Kogan et al. (2017) in Appendix Table E.4. We find that within a patent-class–year (i.e. excluding the variation due to bubbles), a 10% increase in citations corresponds to a 0.1% to 1.7% increase in stock market valuation. This number is substantially smaller than the direct effect of a 30% increase in market-based private value during bubbles. Thus, the quality of innovation as measured by future citations does not experience a rise comparable to the market-based private value, in line with the model predictions.

[18]Appendix Table E.3 show that the results are robust to using a definition of different segments using 6-digit NAICS industries.

Table 4
Number of Citations for Innovation in Bubbles

	Patent Level		Firm Level	
	(1)	(2)	(3)	(4)
Bubble dummy	0.112***	0.113***	0.106***	0.107***
	(0.028)	(0.027)	(0.034)	(0.041)
Log Market Cap (lagged)		−0.034***		0.234***
		(0.008)		(0.008)
Fixed Effects	Y, F	Y, F	Y, F	Y, F
Observations	1,171,806	1,169,860	47,887	47,484
R^2	0.35	0.35	0.78	0.79

Note: Table 4 presents panel regressions of the number of citations (forward looking until the end of the sample), as measured in Kogan et al. (2017) at the patent and firm levels, on a bubble dummy interacted with the number of industries spanned by the different segments of the parent firm. The bubble dummy is measured from Greenwood, Shleifer, and You (2018) and captures whether the firm is an industry that is in a bubble state or not. We control for the lagged market capitalization of the firm and include fixed effects for firm F and patent grant year Y. Standard errors clustered at the grant year level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

4.3 The Social Value of Innovation During Bubbles

Our model predicts that market-based measures of spillovers from competitors are dampened during bubbles while outcome-based measures of spillovers are not. We test these hypotheses empirically by enriching the the specification of Bloom, Schankerman, and Van Reenen (2013) to estimate spillovers conditional on a bubble:

$$\begin{aligned}\log X_{i,t} = \; & \beta \left(B_{i,t} \times \log spillsic_{i,t}\right) + \gamma_1 \log spillsic_{i,t} \\ & + \gamma_2 \log spilltech_{i,t} + \gamma_3 B_{i,t} + \delta Z_{i,t} + \varepsilon_{i,t},\end{aligned} \tag{4.3}$$

where $X_{i,t}$ is either the market value (Tobin's q) or output (sales normalized by an industry price index) of firm i in year t. Like before, $B_{i,t}$ is an indicator of whether firm i is in an industry experiencing a bubble in year t. The controls $Z_{i,t}$ are taken from Bloom, Schankerman, and Van Reenen (2013) and include firm and year fixed effects. Taking $X_{i,t}$ to be the market value of the firm, we measure the market-based spillovers. On the other hand, for outcome-based measures of spillovers, we take $X_{i,t}$ to be firm output. Our hypotheses are that β is positive for the market value, and β is equal to 0 for firm output. Market-based spillovers disappear with speculation as investors ignore the business-stealing effect, while outcome-based spillovers are unchanged.

Table 5 shows that the regression results are consistent with the predictions of our model. The coefficient on the interaction term captures the change in spillovers accompanying an increase in speculation. The coefficient is significantly positive in the market-based spillover regressions, indicating a reduction in the business-stealing effect. Moreover, the point estimate is larger than the coefficient on $spillsic$, suggesting that the business-stealing effect vanishes completely during bubbles, consistent with our high speculation asymptotics in Proposition 2 with inelastic entry ($\gamma\theta > \eta$). In contrast, the presence of a bubble does not have any effect on outcome-based measures of spillovers. Again, this result is in line with our model.

In addition to confirming our theory, these results reject some alternative theories of bubble episodes. Common views of the episodes we assign as bubbles are that these are period of high overall expectations of fundamental (rational or not) or of low discount rates. These theories would predict proportional increase in all market valuations. Remembering that spillovers are the ratio of valuations of displaced profits from new innovation, this creates two conflicts with our findings. First, market-based spillovers would tend to be unchanged. Second, if for reasons specific to changes in the nature of innovation market-based spillovers change, they would do so in line with real spillovers.

Table 5
Social Value of Innovation and Bubbles

	Market-based Spillovers		Outcome Spillovers	
	Jaffe (1)	Mahalanobis (2)	Jaffe (3)	Mahalanobis (4)
Bubble x Spill-SIC	0.152***	0.200***	0.004	−0.000
	(0.027)	(0.037)	(0.009)	(0.013)
Spill-SIC	−0.088***	−0.103***	−0.021***	−0.021**
	(0.016)	(0.033)	(0.006)	(0.010)
Spill-Tech	0.405***	0.844***	0.175***	0.159***
	(0.145)	(0.174)	(0.025)	(0.040)
Fixed Effects	Y, F	Y, F	Y, F	Y, F
Observations	8,896	8,946	8,775	8,825
R^2	0.74	0.74	0.99	0.99

Note: Table 5 presents panel regressions of firm value (Tobin's q or log of sales) on a measure of competition from Bloom, Schankerman, and Van Reenen (2013) interacted with a dummy from Greenwood, Shleifer, and You (2018) that captures whether the firm is an industry that is in a bubble or not. We control for the technological spillover measure that corresponds to public firms that issue patent in similar technological space. We follow the specification from Tables III and V of Bloom, Schankerman, and Van Reenen (2013) for the Tobin's q and sales regressions, respectively. We also include fixed effects for firm F and patent grant year Y. Standard errors clustered at the year level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

These results highlight the importance of distinguishing between market-based and outcome-based measures of spillovers, and more generally of the value of innovation, during speculative episodes. While we focused on a specific source of spillovers, competition, we now show how to enrich our theory to make further predictions specific to other sources of spillovers. Testing these additional predictions would require a more sophisticated econometric analysis, which is beyond the scope of this paper.

5 How Other Spillovers Interact with Speculation

Competition is only one of the many ways through which the effect of an innovation are felt in the economy. When a firm innovates, it not only affects shareholders but also workers (see e.g. Kogan et al. (2020) for recent evidence). A new product alters consumers' choice over their entire consumption basket (Blanchard and Kiyotaki (1987)). Finally, others can learn from this innovation, as emphasized by Romer (1986). We incorporate these three prominent sources of spillovers in our theory. Doing so is interesting for three reasons. First, we obtain new testable predictions of our theory on the interaction of speculation and the innovation that can help discipline future empirical work. Second, by comparing across many different models, we obtain a more systematic typology of what drives the divergence between market-based and outcome-based measures due to speculation. Finally, these results highlight the flexibility of our theoretical framework; Appendix D reinforces this aspect by considering many more realistic extensions.[19]

5.1 Model with Richer Spillovers

In all the settings that follow, we maintain the same preferences, beliefs, firm creation technology, and allocation of production slots as in the model of Section 2. However, we now endogenize how profits $\pi(a)$ are determined in equilibrium, which creates new sources of spillovers. Those spillovers, while important, do not change qualitatively the aggregate behavior of the economy in response to speculation. Specifically, an increase in disagreement n always yields a bubble, and increases the market-based private value of innovation (in absolute terms and relatively to the outcome-based measure). For this reason, we focus only on the novel predictions for spillovers below.

[19]In particular, we study the role of an elastic supply of labor input, a variable number of producing firms M, a setup in which firms compete to participate, and a setting where fixed costs determine the set of producing firms, as in Melitz (2003). For all these models, we obtain simple generalizations of the wedge formula of Proposition 3 and show that the comparative statics of the wedge remain valid.

5.1.1 Labor

Worker supply the labor that is necessary to operate the firms and thus take advantage of blueprints. We formalize this in the model: households are endowed with a fixed quantity of labor L. Firms use labor to produce a homogenous good according to a decreasing returns to scale technology.[20] The production function for a given productivity level a is:

$$y(a) = a \cdot \frac{\sigma}{\sigma - 1} \ell^{\frac{\sigma-1}{\sigma}}, \tag{5.1}$$

where y is firm output, ℓ is firm labor input, and the parameter $\sigma \in [1, \infty]$ controls the returns to scale in labor. In equilibrium (see Appendix B), labor trades at a competitive wage w and the firm profit function given a productivity level a is

$$\pi(a) = \frac{1}{\sigma - 1} \cdot w^{1-\sigma} a^{\sigma}. \tag{5.2}$$

As in Section 2, profits are still isoelastic with respect to productivity a.

The presence of workers gives rise to two new sources of spillovers. First, workers capture some of the surplus from better innovations through higher wages. Investors on financial markets do not take into account this value when deciding on new firm creation. This externality is commonly known as the *appropriability effect*. Second, the competition of firms for the same source of labor leads to a second interaction between them, beyond business-stealing. When a new firm enters, it pushes the wage up. This *input price effect* makes labor for expensive for every other firm. For example, the entry of new technology firms like Facebook creates new opportunities for software engineers (the appropriability effect), bidding their wages up, and thereby making it more difficult for other firms to hire (the input price effect).

Measuring the Spillovers. We extend our measure of social value to account for the presence of labor. For the market-based social value of an extra firm, we compute the change in expected utility of households created by this new firm. This calculation includes both the market value of profits from investing in the firms and also the income from working in these firms. In our model, it is a dollar amount because of the presence of a quasi-linear good at date 0. The outcome-based social value of an extra firm is still the change in total output in the economy. Importantly both definitions coincide with the work in Section XX in the absence of labor.

[20] This perfectly competitive approach, introduced in Hellwig and Irmen (2001), differs from the most commonly used models with imperfect substitution and monopolistic competition. We study these models and the role of the demand complementarities they induce in Section 5.1.2.

This overall measure of spillover sums the effect of each type of spillover. Here these sources are the appropriability, input price, and business-stealing effects; there will be extra terms for each of the additional sources of spillover that we include below. Separation of each type of spillovers in the overall measure maps naturally to estimation strategies from the empirical literature. For example, as described in Section 4.3, Bloom, Schankerman, and Van Reenen (2013) separate the business-stealing and knowledge spillovers. One could imagine measuring the effect of firm entry on workers, the appropriability effect. Similarly, the response of firm value to innovation by other firms that compete for the same workers but operate in different product markets would identify the input price effect.[21]

Predictions. The business-stealing effect of the model of Section 2 is still present in this economy given the similar market structure. As before, this spillover, which arises because entrants displace marginal firms, has magnitude $\mathcal{E}_{\mathcal{I}_n}$. Hence, the predictions for this spillover we made in Section 3.2 still hold.

Second, consider the appropriability effect. When new firms enter, aggregate output increases, with constant elasticity $\mathcal{E}_{\mathcal{C}} = 1/\gamma$. The production function implies that workers receive a constant fraction of aggregate output given by the labor share $(\sigma - 1)/\sigma$. Therefore the social value of entry for workers is $\frac{\sigma-1}{\sigma}\mathcal{E}_{\mathcal{C}}\mathcal{C}/M_e$. Importantly this quantity does not depend directly of the amount of disagreement: all firms offer the same wage in equilibrium so beliefs are irrelevant for workers. However the market-based private value of the firm does. This value is $\mathcal{I}_n/\mathcal{I}_1 \times \sigma^{-1}\mathcal{C}/M$: the relative expected output of a favorite firm to an average firm multiplied by average profits by firm. The spillover is the ratio of these two values,

$$\text{Appropriability Spillover} = (\sigma - 1)\ \mathcal{E}_{\mathcal{C}}\ \mathcal{I}_1/\mathcal{I}_n.$$

When focusing on the outcome-based measure, we can just replace n by 1 and the spillover becomes $(\sigma - 1)\ \mathcal{E}_{\mathcal{C}}$. According to either metric, the appropriability effect is a positive spillover, larger when workers capture more of total surplus — high σ — or when entry has a stronger impact on output — high $\mathcal{E}_{\mathcal{C}}$. Novel to our model, we see that disagreement does not affect the outcome-based spillover. In contrast, the market-based spillover decreases in n, disappearing altogether in the limit. This is because the bubble inflates market value of firms, while workers realize that not all firms will be winners and their earnings do not increase as much.

Finally the input price spillover comes from the effect of firm entry on wages. Going back to equation 5.2, we see that profits have an elasticity of $1 - \sigma$ relative

[21]For clarity, we do not include multiple sectors in the model, the additivity property of total spillover with respect to each type of spillover allows us to isolate the different effects.

to the wage. In addition, the wage grows as fast as aggregate output, $\mathcal{E}_{\mathcal{C}} = \mathcal{E}_w$, because of the constant labor share and labor supply. A change in wage affects all firms proportionally, irrespective of their productivity. Hence disagreement does not impact the input price spillover. Rather, both market-based and outcome-based measures of the spillovers are constant, equal to

$$\text{Input Price Spillover} = -(\sigma - 1)\mathcal{E}_{\mathcal{C}}.$$

This negative externality is larger when firms rely more on labor — high σ — or when the economy responds more to entry — high $\mathcal{E}_{\mathcal{C}}$.[22]

5.1.2 Aggregate Demand

With goods that are not perfect substitutes, households like a consumption basket that is diversified. And, higher productivity for a particular good increases aggregate demand. This implies not only more profit for this firm but also for the firms producing the rest of the consumption basket. To formalize the role of aggregate demand, we study an economy with differentiated goods where firms operate under monopolistic competition at date 1 in the style of Dixit and Stiglitz (1977). Each firm produces a differentiated variety indexed by i and household utility over the set of goods produced is:

$$\mathcal{C} = \left(\int_0^{M_e} \int_{F^{-1}\left(1 - \frac{M}{M_e}\right)}^{\infty} c\left(a, i\right)^{\frac{\sigma-1}{\sigma}} dF(a)di \right)^{\frac{\sigma}{\sigma-1}}.$$

Firms operate a linear technology in labor and output for a firm with productivity a is $y = a\ell$. We leave our other assumptions unaltered.

At the aggregate level, the economy behaves similarly to the previous model.[23] However the microeconomics of firms' interactions is different and so are profits:

$$\pi(a) = \frac{1}{\sigma} \cdot \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} w^{1-\sigma} \cdot \mathcal{C} \cdot a^{\sigma-1}.$$

The elasticity of profits to individual firm productivity is no longer σ but rather $\sigma - 1$. However, profits are now increasing in aggregate demand, $\mathcal{C}$, because of imperfect substitution across goods.

[22] In the case of agreement, $n = 1$, the appropriability and input price spillovers exactly cancel out. It is a situation where pecuniary externalities cancel out even though the first welfare theorem does not hold because of business-stealing.

[23] The profit share is $1/\sigma$, the aggregate production function is homogeneous of degree one in the distribution of productivities, and the relative labor allocations are efficient; This result was first shown in Lerner (1934). It is the consequence of the homogeneous distortions at the firm level when markups are constant. Finally, the macroeconomic elasticities of aggregate consumption and wages to firm entry are therefore $\mathcal{E}_{\mathcal{C}} = \mathcal{E}_w = 1/\gamma$.

This role of aggregate demand gives rise to an additional source of spillovers: a demand externality as in Blanchard and Kiyotaki (1987). Similar to the role of the wage, aggregate demand affects all firms proportionally. Therefore the demand spillover is the product of the elasticity of aggregate output to entry, $\mathcal{C}$, and the elasticity of profits to aggregate output, 1:

$$\text{Demand Spillover} = \mathcal{E}_{\mathcal{C}}.$$

Demand spillovers do not depend of disagreement and are identical when measured using real outcomes or market value.

5.1.3 Knowledge Spillovers

Firms also learn from each other's innovations. We capture the role of knowledge spillovers in the style of Romer (1990) by assuming that a firm's productivity combines its own type, a, and an aggregate of all the active firms' productivity, A. We assume the aggregator is homogenous of degree one in the productivity distribution.[24] The production function extends (5.1) and becomes:

$$y = \frac{\sigma}{\sigma - 1} a^{1-\alpha} A^{\alpha} \ell^{\frac{\sigma-1}{\sigma}},$$

where α is the intensity of knowledge spillovers. Again, the macroeconomic features of the simple economy with labor are preserved.[25] However, the microeconomics of firms' interactions differ. Profits are:

$$\pi(a) = \frac{1}{\sigma - 1} \cdot w^{1-\sigma} \cdot A^{\alpha\sigma} \cdot a^{(1-\alpha)\sigma}.$$

As the role of knowledge increases – larger α – firm profits respond more to aggregate knowledge A instead of individual productivity a.

The role of aggregate knowledge for profits highlights how knowledge spillovers operate. Just like the previous two cases, the impact of knowledge on profits is the same irrespective of each individual firm's productivity. The knowledge spillover is therefore the product of the elasticity of profit to knowledge $\alpha\sigma$, and the elasticity of knowledge to entry, $1/\gamma$,[26]

$$\text{Knowledge Spillover} = \alpha(\sigma - 1)\mathcal{E}_{\mathcal{C}}.$$

This expression does not depend of disagreement, and is identical for market-based and outcome-based measures.

---

[24] In Appendix B.3 we derive the case of Hölder mean of degree q, $A = \left(M_e/M \int_{\underline{a}}^{\infty} a^q dF(a)\right)^{1/q}$, where the parameter $q < \gamma$ controls how much aggregate knowledge comes from the top firms.

[25] The labor share is $(\sigma-1)/\sigma$ and $\mathcal{E}_{\mathcal{C}} = \mathcal{E}_w = 1/\gamma$.

[26] Because both knowledge and output are determined by the distribution of firm productivity, they grow at the same pace with entry, $\mathcal{E}_A = \mathcal{E}_{\mathcal{C}}$.

5.2 Social Value and Disagreement

The previous results provide useful predictions for how each channel through which innovation affects the economy varies with disagreement. We now turn to the behavior of social value overall, that is the total effect of an innovation, which combines all sources of spillover. This quantity is interesting in its own right, but also because it represents the optimal subsidy or tax on firm entry. The following proposition shows that the spillovers we have introduced so far can be organized in three categories, each with their own response to disagreement.

Proposition 3. *For all the models of Section 5, the market-based spillover is*

$$spill_{mkt}(n) = \underbrace{\mathcal{E}_{\mathcal{I}_n}}_{\textit{bus. stealing}} + \underbrace{\mathcal{E}_{\pi}}_{\textit{general equilibrium}} + \underbrace{(\sigma-1)\mathcal{E}_{\mathcal{C}}\frac{\mathcal{I}_1}{\mathcal{I}_n}}_{\textit{appropriability effect}}. \tag{5.3}$$

The outcome-based spillover, which does not depend of disagreement, is:

$$spill_{out} = \mathcal{E}_{\mathcal{I}_1} + \mathcal{E}_{\pi} + (\sigma-1)\mathcal{E}_{\mathcal{C}}. \tag{5.4}$$

This decomposition highlights how disagreement can matter for market-based measures of spillovers. While our theoretical exercise is naturally not exhaustive, most spillovers considered in the innovation literature fit in one of our three categories.

The first category is business-stealing. As we have discussed before, disagreement dampens the effect of business-stealing. The risk of displacement is particularly acute for relatively less productive firms. However, with speculation, each investor places a relatively higher weight on her investments having high productivity. In the limit when n goes to infinity, this force can be so strong that the spillover disappears altogether.[27] Competitive interactions between firms of different productivity can take other forms that the displacement of our model. Still, the takeaway is that spillovers that are more bottom-heavy tend to be dissipated by disagreement.

This stands in contrast to our second category: general equilibrium effects. In our models, these are the effects of the wage, aggregate demand, and aggregate knowledge on firm profits. The common force across all these sources of spillovers is that they affect proportionally all firms, irrespective of their productivity. Therefore, they can be summarized by the elasticity of firm profits to firm entry, holding productivity constant, $\mathcal{E}_{\pi}$. Because the response to these general equilibrium forces does not interact with the productivity distribution, these spillovers do not depend of beliefs.

Finally, the third category of spillover are appropriability effects. Not all spillovers affect firms. In our models workers capture some of the surplus due

[27] In the model of Section 2, this occured when $\gamma\theta > \eta$, while now the condition is $\gamma\theta > 1$.

to firm entry. Because the surplus of workers is determined in the spot market for labor, it does not depend of the relative positions of firms. Unlike firm valuations, wage expectations are not affected by speculation about the relative positions of firms beyond its direct impact on entry and overall labor demand. When disagreement increases, the market-based spillover to workers disappears. This insight is not specific to workers, but rather affects all stakeholders of the innovation process. Other key stakeholders — which we could have similarly introduced in the model — are owners of production inputs in scarce supply other than labor, or consumers who enjoy some of the surplus.

5.3 Three Illustrations of the Role of Disagreement

We draw three implications from Proposition 3 that illustrate that disagreement alters fundamentally how to measure and interpret the value of innovation.

5.3.1 Macroeconomic versus Microeconomic Elasticities

When there is no disagreement, market-based and outcome-based measures of spillovers coincide because valuations are expected outcomes. In the economies we have considered, the result is actually stronger. The spillover under disagreement is the same across all specifications: with labor only, with aggregate demand and with aggregate knowledge. It takes the value:

$$spill_{out} = spill_{mkt}(1) = \sigma \mathcal{E}_{\mathcal{C}} - 1. \tag{5.5}$$

While we can derive this expression from equation 5.4 separately for each model, a simple macroeconomic argument justifies the result. The total effect of entry is the response of aggregate output to entry $d\mathcal{C}/dM_e$. Ex ante, each firm contributes an equal fraction to total output, and the value of the firm is output times the profit share $1/\sigma$. Hence, the value of a firm is $\mathcal{C}/(M_e\sigma)$. And, the spillover is given immediately by equation 5.5.

This result tells us that only two macroeconomic quantities are necessary to evaluate the total spillover, the capital share and the elasticity of aggregate output to firm entry, irrespective of the nature of firm interactions. In particular, all of our specifications lead to the same values of these two quantities. The presence of disagreement breaks this result. Because we have showed that different spillovers respond differently to disagreement, this aggregate reasoning does not work anymore.

One particularly telling example is the limit of large n, when the market-based spillover converges to the profit elasticity $\mathcal{E}_\pi$. This spillover measure is a microeconomic elasticity. It is the response of the profits of one specific firm (i.e. of given productivity) on overall entry. This implies that one needs firm-level data rather than aggregate data to estimate spillovers. Moreovers, across

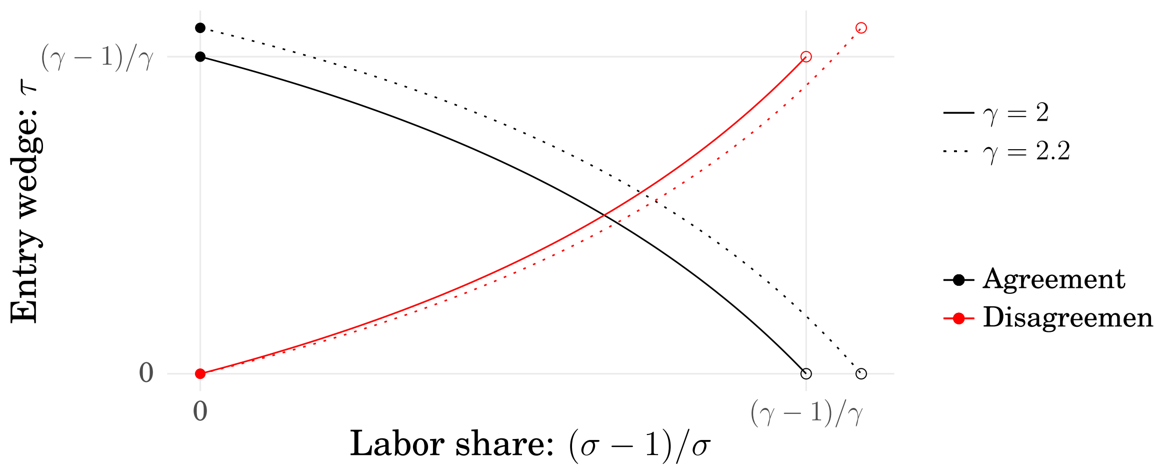

Figure 3
Comparisons of taxes with and without disagreement

our three model specifications, while the agreement spillover is identical, the spillover for large disagreement takes different values:

$$spill_{mkt}(n \to \infty) = -\frac{\sigma - 1}{\gamma} \qquad \text{with labor only,} \tag{5.6}$$

$$spill_{mkt}(n \to \infty) = -\frac{\sigma - 2}{\gamma} \qquad \text{with aggregate demand,} \tag{5.7}$$

$$spill_{mkt}(n \to \infty) = -\frac{(1-\alpha)\sigma - 1}{\gamma} \qquad \text{with aggregate knowledge.} \tag{5.8}$$

Said otherwise, the nature of microeconomic interactions matters in presence of disagreement.

5.3.2 Reversal of Comparative Statics

The divergence between market-based and outcome-based spillovers is not only quantitative but also qualitative. Key properties of the economy often have an opposite impact on the total spillover depending on whether it is measured using outcomes or market values. The following proposition highlights one such reversal, for a parameter common to all of our specification, σ.

Proposition 4. *For all models, the outcome-based spillover is increasing in the labor share. Conversely, with high disagreement ($n \to \infty$ and $\theta > 1/\gamma$), the market-based spillover is decreasing in the labor share.*

The outcome-based spillover is given by the macroeconomic formulation of equation 5.5. An economy with larger labor share has mechanically a lower capital share. Thus, the importance of social value relative to the value of one firm is larger. For the market-based spillover, the focus in on the elasticity of individual firm profits to entry. Then, a higher labor share implies higher reliance on labor and therefore stronger negative spillovers through the wage effect.

By examining our results, the reader can find more situations where reversals occur. In the model with labor, comparative statics with respect to the thickness of the tail of the productivity distribution γ are also reversed. Figure **??** illustrates these results. Appendix D has more examples.

5.3.3 Reversal of Sign of the Spillover

More strikingly, we also identify situations where the sign of the total spillover is reversed. These are cases where firm entry brings positive externalities according to market-based measures but negative externalities according to outcome-based measures, or vice-versa.

Proposition 5. *With demand externalities or knowledge spillovers, if the labor share is close to zero, the outcome-based spillover is positive and the market-based spillover is negative with large disagreement. The converse happens when the labor share is close to its upper bound.*

When the labor share is low, the labor surplus is relatively small, and the dominant force for the wedge is that firms do not internalize the aggregate decreasing returns to scale of the economy, leading to negative real spillovers. With disagreement however, since firms do not rely much on labor, the general equilibrium effect is small, hence the demand or knowledge externality dominates, leading to positive value spillovers. The sign reversal across measures of spillovers is not a knife-edge case. Reversals happen throughout the entire range of the labor share whenever $\gamma = 2$ with demand externalities or $\alpha = 1 - 1/\gamma$ with knowledge spillovers. Interestingly, the proposition also points that the sign reversal can happen in both directions: a positive spillover becoming negative or a negative spillover becoming positive.

6 Welfare Implications

One of the reasons why measuring innovative spillovers is important is because such estimates can guide the design of policies. In this section, we ask what are the welfare implications of spillover measures. We focus on a simple question: if a planner can choose a subsidy or tax on firm entry, what would it be. The presence of disagreement complicates answering this question, because the choice of planner objective is not trivial. Which beliefs should the planner use when evaluating allocations? We follow two approaches: a Pareto criterion and a paternalistic approach imposing common beliefs. For each approach we map the market-based and outcome-based measures of the private and social value we have studied so far to an optimal tax rate. These results highlight again the importance of acknowledging the reality of speculation and understanding the divergence between measurement approaches.

6.1 Non-Paternalistic Approach

A classic approach to welfare evaluation is the Pareto criterion. An allocation improves overall welfare if all agents in the economy favor it. By nature this criterion is not paternalistic: moving away from the competitive equilibrium requires the support of all agents in the economy. Said otherwise, with this approach, the optimal tax policy is one that would receive support by all agents in a vote.[28] Interestingly in our case, this approach does not take a stand on what are correct beliefs, but rather lets each individual agent evaluate their own utility.

Not having to choose beliefs resonates with the situations we study. We are interested in episodes when there is little information about new firms. Households thus rely on their priors to evaluate these firms, and there is no way to forecast who is correct. The Pareto criterion respects this difficulty. The planner is no better judge of firms' futures than any investor. In contrast, a paternalistic approach might be more appropriate when heterogenous beliefs are viewed as inherently inefficient, due to a failure of communication or the irrationality of some agent.

Our model is symmetric: despite agents having different beliefs, they each have the same evaluation of their own utility. This symmetry facilitates solving for the optimal allocations, because the problem reduces to maximizing the utility of all agents in the economy. The planner sets a tax τ for each additional firm created from blueprints, rebated as a lump sum to households.[29] This is a standard Pigouvian taxation problem:

$$\max_{\tau} \mathcal{U}_j = 1 + M_e \mathbf{E}^j \{\pi_i\} - W(M_e) \tag{6.1}$$

The planner chooses a tax rate which equalizes private incentives to create firms with their social value according to the planner's objective. When a firm creator introduces a new firm, it pays $(1-\tau)p_i$. Hence, the optimal tax rate is $\tau = 1 - \frac{\text{social value}}{p_i}$. Social value for the non-paternalistic planner is given by changes

[28] This interpretation provides a positive — as opposed to normative — view of these results, as the outcome of political decision-making.

[29] We could also consider a more general constrained efficient planner problem. In this general version, the planner allocates date-1 consumption. Consumption plans must be linear combinations of firm profits with positive coefficients. This limit reflecting our assumptions about trading: no short-selling and derivative contracts. In the model of Section 2, the allocation chosen by a planner with only a linear tax is Pareto optimal for the general constrained efficient planner. In fact, it is the most efficient allocation if we impose additionally that welfare weights are equal across agents.

in expected utility, which correspond to market prices. This yields the following simple expression for the optimal tax:

$$\tau_{pareto} = -spill_{mkt}. \tag{6.2}$$

Therefore, for a non-paternalistic planner, the market-based spillover is a sufficient statistic for the optimal tax, rather than outcome-based measures.

We have seen that the market-based spillover is often very different with agreement compared to disagreement. This implies that, in this framework, one should not use estimates from normal market conditions to draw policy conclusions during bubbles. Or, positively, this distinction explains why policies to reduce firm entry may not receive support even during bubbles, even when investors agree that there is excess entry. Further, if the only available data is on outcomes, one ought to use a framework like ours to convert these estimates into their market-based counterparts. The reversal results of the previous section highlight the importance of having a theoretical framework to do so.

6.2 Paternalistic Approach

Alternatively, one could follow a paternalistic approach: the planner knows the "true" distribution and evaluates allocations under this distribution. Brunnermeier, Simsek, and Xiong (2014) propose an improvement on this approach which avoid taking a stance on the true distribution. In their work, the planner considers efficiency across any convex combination of agents' beliefs. In our setting, all agents always agree on the distribution of firm productivity, and therefore this distribution is the only choice for the paternalistic planner. A benefit of this approach is that, at least from the perspective of social choice, the tension that not all beliefs can be right at the same time is resolved.

The planner maximizes expected utility under the population distribution F, which coincides with aggregate consumption net of entry costs, $\mathcal{C} - W(M_e)$. As before, this is a standard Pigouvian taxation problem, except that now social value is the marginal effect of entry on output rather than on households' perception of utility. Hence, the optimal tax becomes:

$$\tau_{pater} = 1 - (spill_{out} + 1)\frac{\text{private value}_{out}}{\text{private value}_{mkt}}. \tag{6.3}$$

Two aspects drive the choice of policy. First, the planner accounts for spillovers. Here the relevant measure of spillovers uses outcomes rather than market values. Symmetric observations to those we made for the Pareto planner hold. Using directly estimates of spillovers measured using market values leads to incorrect inference. If these estimates are the only ones available, one should use a theoretical framework to convert them in outcome-based spillovers.

Second, there is a wedge between the private value on markets — which shape incentives to create firms — with the private value according to the planner's beliefs. We have seen that the bubble inflates market-based private value relative to outcome-based. So, in general, this second force will push towards taxing entry in order to lean against what the planner views as excessive valuations. In particular, if outcome-based spillovers are positive, implying that there is under-entry absent disagreement, a bubble can actually push the economy towards efficiency.

Of course, in the case of agreement, the two planner objectives coincide, and lead to the same optimal tax. Then, the optimal tax is the opposite of the total spillover, market-based or outcome-based.

7 Conclusion

Speculation and innovation often coincide. A naive view of the role of speculation and the bubbles it generate is that they do not matter for the innovation process and render information from financial markets useless. In contrast, we argue that there is method to this madness. Innovation and speculation interact, and it is important to understand how. Failing to do so can distort our qualitative and quantitative understanding of the private and social value of innovation, leading to erroneous answers to both positive and normative questions about innovation.

To that end, the paper introduces a model of the interaction of innovation and speculation. Our theory makes sharp predictions for the value of innovation. We study the impact of disagreement on both the private and social value of innovation. We find that the structure of the model actually reflects the real world. During bubbles, information from financial markets and from real outcomes about the value of innovation diverge from each other; not in a random way but rather in the direction predicted by our theory. Extending the work of Kogan et al. (2017), we find that asset prices indicate increases in the private value of innovation during bubbles, with no commensurate increase in outcome-based measures. Using the method of Bloom, Schankerman, and Van Reenen (2013), financial markets suggest that competitive spillovers disappear during bubbles, while they are unchanged when looking at sales.

Thus, accounting for the link between speculation and innovation is a fruitful enterprise. We provide the first steps to dig deeper, showing how our framework can entertain many sources of innovative spillovers, and highlighting broad principles on their interaction with disagreement. We also explain how to map data to policy decisions in periods of disagreement. We look forward to seeing how far future implementations of these ideas help understand the process of innovation.

References

Abernathy, W.J. 1978. *The Productivity Dilema: Roadblock to Innovation in the Automobile Industry*. Johns Hopkins University Press, Baltimore, Md.

Blanchard, Olivier Jean and Nobuhiro Kiyotaki. 1987. "Monopolistic Competition and the Effects of Aggregate Demand." *The American Economic Review* 77 (4):647–666.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. "Identifying technology spillovers and product market rivalry." *Econometrica* 81 (4):1347–1393.

Borjas, George J. and Kirk B. Doran. 2012. "The Collapse of the Soviet Union and the Productivity of American Mathematicians." *Quarterly Journal of Economics* 127 (3):1143–1203.

Brunnermeier, Markus K. 2017. *Bubbles*. London: Palgrave Macmillan UK, 1–8.

Brunnermeier, Markus K. and Martin Oehmke. 2013. "Bubbles, Financial Crises, and Systemic Risk." In *Bubbles, Financial Crises, and Systemic Risk*, *Handbook of the Economics of Finance*, vol. 2, edited by George M. Constantinides, Milton Harris, and Rene M. Stulz, chap. 18. Elsevier, 1221 – 1288.

Brunnermeier, Markus K., Alp Simsek, and Wei Xiong. 2014. "A Welfare Criterion For Models With Distorted Beliefs." *Quarterly Journal of Economics* 129 (4):1753–1797.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein. 2002. "Breadth of ownership and stock returns." *Journal of Financial Economics* 66 (2):171 – 205. Limits on Arbitrage.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina. 2002. "Differences of Opinion and the Cross Section of Stock Returns." *Journal of Finance* 57 (5):2113–2141.

Dixit, Avinash K and Joseph E Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67 (3):297–308. URL http://www.jstor.org/stable/1831401.

Freeman, Christopher. 1982. *The Economics of Industrial Innovation*. MIT Press.

Greenwood, Robin, Andrei Shleifer, and Yang You. 2018. "Bubbles for Fama." *Journal of Financial Economics* .

Harrison, J Michael and David M Kreps. 1978. “Speculative investor behavior in a stock market with heterogeneous expectations.” *The Quarterly Journal of Economics* :323–336.

Hellwig, Martin and Andreas Irmen. 2001. “Endogenous Technical Change in a Competitive Economy.” *Journal of Economic Theory* 101 (1):1 – 39.

Hong, Harrison and Jeremy C. Stein. 2007. “Disagreement and the Stock Market.” *Journal of Economic Perspectives* 21 (2):109–128.

Jaffe, Adam. 1986. “Technological Opportunity and Spillovers of R&D: Evidence from Firms’ Patents, Profits, and Market Value.” *American Economic Review* 76 (5):984–1001.

Janeway, W.H. 2012. *Doing Capitalism in the Innovation Economy: Markets, Speculation and the State*. Cambridge University Press.

Kogan, Leonid, Dimitris Papanikolaou, Lawrence D. W. Schmidt, and Jae Song. 2020. “Technological Innovation and Labor Income Risk.” Tech. rep., NBER Working Paper.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. “Technological innovation, resource allocation, and growth.” *The Quarterly Journal of Economics* 132 (2):665–712.

Lerner, A P. 1934. “The Concept of Monopoly and the Measurement of Monopoly Power.” *Review of Economic Studies* 1 (3):157–175.

Melitz, Marc J. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica* 71 (6):1695–1725.

Miller, Edward M. 1977. “Risk, Uncertainty, and Divergence of Opinion.” *The Journal of Finance* 32 (4):1151–1168.

Mokyr, Joel. 1992. *The Lever of Riches: Technological Creativity and Economic Progress*. Oxford University Press.

Romer, Paul M. 1986. “Increasing Returns and Long-Run Growth.” *Journal of Political Economy* 94 (5):1002–1037.

Romer, Paul M. 1990. “Endogenous Technological Change.” *Journal of Political Economy* 98 (5):71–102.

Scheinkman, J.A. 2014. *Speculation, Trading, and Bubbles*. Kenneth J. Arrow Lecture Series. Columbia University Press.

Scheinkman, Jose A and Wei Xiong. 2003. “Overconfidence and speculative bubbles.” *Journal of political Economy* 111 (6):1183–1220.

Van den Steen, Eric. 2004. “Rational Overoptimism (and Other Biases).” *American Economic Review* 94 (4):1141–1151.

Yu, Jialin. 2011. “Disagreement and return predictability of stock portfolios.” *Journal of Financial Economics* 99 (1):162 – 183.

Appendix

A Derivations for Simple Model

A.1 General results

We first express the wedge in integral form, then prove Proposition 1 for general functions F and π.

A.1.1 General formulas

The wedge between the competitive equilibrium and the planner problem is

$$\tau_n(M_e) = -[M_e V^{(n)\prime}(M_e)]/[V^{(n)}(M_e)].$$

The numerator and denominator have interpretable expressions. First rewrite the denominator in the following integral form:

$$\begin{aligned} V^{(n)} &= \int_{F^{-1}}^{\infty} \pi(x) dF_n(x) \\ &= \int_{F^{-1}}^{\infty} \pi\left(x\right) \frac{F_n'}{F'}\left(x\right) dF(x), \end{aligned} \tag{A.1}$$

where we denote $F^{-1}(1 - M/M_e)$ by F^{-1} for convenience. Now the numerator can be written:

$$-M_e \frac{dV^{(n)}}{dM_e} = -\frac{M}{M_e} \cdot \pi\left[F^{-1}\right] \cdot \frac{F_n'}{F'}\left[F^{-1}\right] \tag{A.2}$$

$$= \int_{F^{-1}}^{\infty} \pi\left[F^{-1}\right] \frac{F_n'}{F'}\left[F^{-1}\right] dF(x). \tag{A.3}$$

This leads to the following formula for the wedge:

$$\tau_n = \frac{\int_{F^{-1}}^{\infty} \pi\left[F^{-1}\right] \frac{F_n'}{F'}\left[F^{-1}\right] dF(x)}{\int_{F^{-1}}^{\infty} \pi\left(x\right) \frac{F_n'}{F'}\left(x\right) dF(x)}. \tag{A.4}$$

A.1.2 Comparing the wedges

Lemma A.1. *Holding M_e constant, the wedge is larger with agreement than with disagreement.*

Proof. First recall the wedge $\tau_n(M_e) = \frac{-M_e V^{(n)\prime}(M_e)}{V^{(n)}(M_e)}$. We have the derivative:

$$V^{(n)\prime} = -\frac{1}{M_e}\frac{M}{M_e}\pi\left[F^{-1}\right] \cdot n\left(1 - \frac{M}{M_e}\right)^{n-1},$$

and we can bound $V^{(n)}$:

$$\begin{aligned} V^{(n)} &= \int_{F^{-1}}^{\infty} \pi(x) n F^{n-1}(x) dF(x) \\ &\geq \int_{F^{-1}}^{\infty} \pi(x) n F^{n-1} \left[F^{-1}\right] dF(x) \\ &\geq n\left(1-\frac{M}{M_e}\right)^{n-1} \int_{F^{-1}}^{\infty} \pi(x) dF(x). \end{aligned}$$

Therefore we are able to bound the wedge for a given M_e and n:

$$\tau_n(M_e) \leq \frac{\int_{F^{-1}}^{\infty} \pi\left[F^{-1}\right] dF(x)}{\int_{F^{-1}}^{\infty} \pi(x) dF(x)} \leq \tau_1(M_e), \tag{A.5}$$

where the second inequality comes from the definition of $\tau_1(M_e)$. ∎

A.2 Power case derivations

We now outline the derivations for the case we focus on in the main text, with $F(a) = 1 - a^{-\gamma}$ and $\pi(a) = a^{\eta} \cdot \mathbf{1}\{a \geq \underline{a}\}$.

First, define

$$\underline{a} = F^{-1}(1 - \frac{M_e}{M}) = \left(\frac{M_e}{M}\right)^{1/\gamma}.$$

The ex-ante value of a firm, $V^{(n)}(M_e)$, is:

$$V^{(n)}(M_e) = \int_{\left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}}}^{\infty} x^{\eta} \gamma n x^{-\gamma-1} \left(1 - x^{-\gamma}\right)^{n-1} dx \tag{A.6}$$

$$= \gamma n \underline{a}^{\eta-\gamma} \int_{1}^{\infty} t^{\eta-\gamma-1} \left(1 - \underline{a}^{-\gamma} t^{-\gamma}\right)^{n-1} dt, \tag{A.7}$$

The first derivative with respect to entrants is:

$$\frac{dV^{(n)}}{dM_e} = -\frac{1}{M_e} \cdot \frac{M}{M_e} \left(\frac{M_e}{M}\right)^{\frac{\eta}{\gamma}} \cdot n \left(1 - \frac{M}{M_e}\right)^{n-1}. \tag{A.8}$$

It is convenient to express $-M_e V^{(n)\prime}(M_e)$ as:

$$-M_e V^{(n)\prime}(M_e) = \underline{a}^{\eta-\gamma} \cdot n \left(1 - \underline{a}^{-\gamma}\right)^{n-1}. \tag{A.9}$$

A.2.1 Wedge and firm entry under agreement

Lemma A.2. *With agreement ($n = 1$), the wedge does not depend on the level of entry:*

$$\tau_1(M_e) = \frac{\gamma - \eta}{\gamma}. \tag{A.10}$$

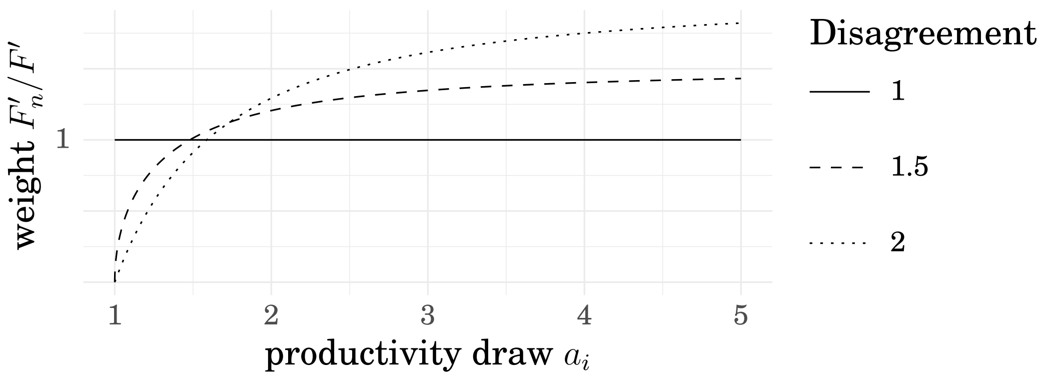

Figure A.1
Distortion of productivity weights F_n'/F'

The level of entry is:

$$\frac{M_e}{M} = \left(f_e \frac{\gamma - \eta}{\gamma}\right)^{-\frac{\gamma}{\gamma(\theta+1)-\eta}} . \tag{A.11}$$

Proof. Under agreement, $n = 1$, and we can derive an exact solution for the mass of firms entering in equilibrium, M_e. The value of a firm is:

$$\begin{aligned} V^{(1)}(M_e) &= \gamma \underline{a}^{\eta-\gamma} \int_1^\infty t^{\eta-\gamma-1} dt \\ &= \frac{\gamma}{\gamma - \eta} \underline{a}^{\eta-\gamma} = \frac{\gamma}{\gamma-\eta} \left(\frac{M_e}{M}\right)^{\frac{\eta-\gamma}{\gamma}} . \end{aligned} \tag{A.12}$$

From equation (A.9) with $n = 1$, we have the numerator of the wedge:

$$-M_e \frac{dV^{(1)}}{dM_e} = \left(\frac{M_e}{M}\right)^{\frac{\eta-\gamma}{\gamma}} , \tag{A.13}$$

which leads directly to the desired formula (A.10) for the wedge. Finally, we can rewrite equation (2.8):

$$f_e \left(\frac{M_e}{M}\right)^\theta = \frac{\gamma}{\gamma-\eta} \left(\frac{M_e}{M}\right)^{\frac{\eta-\gamma}{\gamma}} ,$$

which reduces to (A.11) as desired. ∎

A.2.2 Disagreement lowers market spillovers

A.2.3 Disagreement asymptotics

Lemma A.3. *If $\theta \geq 0$, then as disagreement increases ($n \to \infty$) the mass of entrants also increases and goes to infinity:* $\lim_{n\to\infty} M_e = \infty$.

Proof. We define $\underline{a}_n = (M_e/M)^{1/\gamma}$, where M_e now depends on n, and show that $\underline{a}_n \to \infty$.

Equation (2.8) implies an implicit definition of the sequence $\underline{a}_n$:

$$f_e \underline{a}_n^{\gamma(\theta+1)-\eta} = \gamma n \int_1^\infty t^{\eta-\gamma-1} \left(1 - \underline{a}_n^{-\gamma} t^{-\gamma}\right)^{n-1} dt.$$

Suppose $\underline{a}_n$ has a finite limit that is strictly larger than zero, i.e. $\underline{a}_\infty > 0$.[30] Then there exists N large enough such that $\forall n > N$, $\underline{a}_n > A = \underline{a}_\infty - \epsilon > 0$. We obtain a lower bound for the right-hand side of the implicit equation above:

$$\begin{aligned} I_n &= \gamma n \int_1^\infty t^{\eta-\gamma-1} \left(1 - \underline{a}^{-\gamma} t^{-\gamma}\right)^{n-1} dt \\ &> \gamma n \int_1^\infty t^{\eta-\gamma-1} \left(1 - A^{-\gamma} t^{-\gamma}\right)^{n-1} dt. \end{aligned}$$

Consider an arbitrary threshold T_n that depends on n and satisfies:

$$\begin{aligned} I_n &> \gamma n \int_{T_n}^\infty t^{\eta-\gamma-1} \left(1 - A^{-\gamma} t^{-\gamma}\right)^{n-1} dt \\ &> \gamma n \left(1 - A^{-\gamma} T_n^{-\gamma}\right)^{n-1} \int_{T_n}^\infty t^{\eta-\gamma-1} dt \\ &= \frac{\gamma}{\gamma-\eta} \cdot n \cdot T_n^{\eta-\gamma} \left(1 - A^{-\gamma} T_n^{-\gamma}\right)^{n-1}. \end{aligned}$$

Choose the threshold $T_n = n^{1/\gamma}$. The bound becomes:

$$\begin{aligned} I_n &> \frac{\gamma}{\gamma-\eta} \cdot n^{\frac{\eta}{\gamma}} \exp\left(-(n-1)\log(1 - A^{-\gamma} n^{-1})\right) \\ &> \frac{\gamma}{\gamma-\eta} \cdot n^{\frac{\eta}{\gamma}} \exp\left(-(n-1) A^{-\gamma} n^{-1} + \mathcal{O}(n^{-1})\right). \end{aligned}$$

Since $\gamma(\theta+1) - \eta \geq \gamma - \eta > 0$, this implies $I_n \to \infty$, contradicting $\underline{a}_\infty < \infty$. ∎

Lemma A.4 (Asymptotics for firm creation). *In the high disagreement limit ($n \to \infty$), we have the following asymptotics for the mass of firms created, M_e:*

- ***If** $\gamma\theta < \eta$, **then** $M_e/M = \left(\frac{1}{f_e}\frac{\gamma}{\gamma-\eta} \cdot n\right)^{\frac{\gamma}{\gamma(\theta+1)-\eta}}$.*
- ***If** $\gamma\theta = \eta$, **then** $\lim_{n\to\infty} M_e/M = \alpha_\infty n$, **where** α_∞ **is a constant defined below.***

Proof. Substituting $\underline{a}$ into (2.8), we have:

$$\begin{aligned} f_e &= \gamma \underline{a}^{\eta-\gamma(\theta+1)} n \int_1^\infty t^{\eta-\gamma-1} \left(1 - \underline{a}^{-\gamma} t^{-\gamma}\right)^{n-1} dt \\ &\simeq \gamma \underline{a}^{\eta-\gamma(\theta+1)} n \int_1^\infty t^{\eta-\gamma-1} \exp\left(-(n-1)\underline{a}^{-\gamma} t^{-\gamma}\right) dt, \end{aligned}$$

where we have used the fact that $\underline{a} \to \infty$ from Lemma A.4, and $\log(1-x) = -x + \mathcal{O}(x^2)$. To find a solution, we guess the asymptotics of $\underline{a}(n)$. We rewrite $\underline{a} = \alpha(n) n^{1/(\gamma(1+\theta)-\eta)}$ and show

[30] Since the mass of firms producing cannot be higher than the mass of firms created, $\underline{a}_n \geq 1$.

that $\alpha(n)$ converges to a finite limit α. The above equation becomes:

$$f_e = \gamma\alpha(n)^{\eta-\gamma(\theta+1)} \int_1^\infty t^{\eta-\gamma-1} \exp\left(-\alpha(n)^{-\gamma} \frac{n-1}{n^{\frac{\gamma}{\gamma(1+\theta)-\eta}}} t^{-\gamma}\right) dt.$$

Suppose $\gamma\theta < \eta$. Then the exponential term converges to zero and we have:

$$f_e = \gamma\alpha^{\eta-\gamma(\theta+1)} \int_1^\infty t^{\eta-\gamma-1} = \alpha^{\eta-\gamma(\theta+1)} \frac{\gamma}{\gamma-\eta},$$

such that we have the following asymptotics for firm entry:

$$\frac{M_e}{M} = \left(\frac{1}{f_e}\frac{\gamma}{\gamma-\eta} \cdot n\right)^{\frac{\gamma}{\gamma(\theta+1)-\eta}}. \tag{A.14}$$

Suppose $\gamma\theta = \eta$. Then $\underline{a}$ is defined by:

$$f_e = \gamma \underline{a}_n^{-\gamma} n \int_1^\infty t^{\eta-\gamma-1} \left(1 - \underline{a}^{-\gamma} t^{-\gamma}\right)^{n-1} dt.$$

Since $\underline{a} = (M_e/M)^{1/\gamma}$, it is sufficient to guess and verify that $\underline{a}_n = \alpha(n)^{-1/\gamma} n^{1/\gamma}$, and $\alpha(n)$ has a finite limit, α_∞ defined by:

$$\begin{aligned} f_e \;&=\; \gamma\alpha(n) \int_1^\infty t^{\eta-\gamma-1} \exp\left(-(n-1)(\alpha(n)n^{-1} + \mathcal{O}(\alpha(n)^2 n^{-2}))t^{-\gamma}\right) dt \\ &\xrightarrow[n\to\infty]{} \gamma\alpha_\infty \int_1^\infty t^{\eta-\gamma-1} e^{-\alpha_\infty t^{-\gamma}} dt, \end{aligned}$$

where we take the limit when $n \to \infty$. The wedge with agreement implies:

$$\begin{aligned} f_e &> \gamma\alpha_\infty e^{-\alpha_\infty} \int_1^\infty t^{\eta-\gamma-1} dt \\ &> \alpha_\infty e^{-\alpha_\infty} \frac{\gamma}{\gamma-\eta} \end{aligned}$$

and thus

$$\frac{\alpha_\infty e^{-\alpha_\infty}}{f_e} < \frac{\gamma-\eta}{\gamma}, \tag{A.15}$$

which implies a finite bound on α_∞. ∎

Using the asymptotics derived in Lemma A.4, we now prove Proposition 2.

Proof. **(Proposition 2)** Suppose $\gamma\theta < \eta$. Substitute the asymptotics derived in equation (A.14) into the formula for the wedge:

$$\tau_n(M_e) = \frac{\frac{M}{M_e}\left(\frac{M_e}{M}\right)^{\frac{\eta}{\gamma}} n \left(1 - \frac{M}{M_e}\right)^{n-1}}{f_e\left(\frac{M_e}{M}\right)^\theta} \tag{A.16}$$

$$\simeq \frac{1}{f_e} \cdot f_e \frac{\gamma-\eta}{\gamma}\frac{1}{n} \cdot n \left(1 - \frac{M}{M_e}\right)^{n-1} \to \frac{\gamma-\eta}{\gamma}, \tag{A.17}$$

where we have used the fact that $(1-M/M_e)^{n-1} \to 1$.[31] The wedge therefore converges to the wedge with agreement in this case.

Now suppose $\gamma\theta > \eta$. We write the wedge directly:

$$\tau_n(M_e) = \frac{n\underline{a}^{\eta-\gamma}(1-\underline{a}^{-\gamma})^{n-1}}{f_e \underline{a}^{\gamma\theta}}.$$

First suppose $\underline{a} \to \infty$. We rewrite the competitive equilibrium condition (2.8):

$$n\underline{a}^{-\gamma} = \frac{f_e \underline{a}^{\gamma\theta-\eta}}{\gamma \int_1^\infty t^{\eta-\gamma-1}(1-\underline{a}^{-\gamma}t^{-\gamma})^{n-1}dt}.$$

The denominator is bounded from above by $\gamma \int_1^\infty t^{\eta-\gamma-1}dt$, which implies $n\underline{a} \to \infty$. Using a first-order approximation, we have:

$$(1-\underline{a}^{-\gamma})^{n-1} \simeq \exp\left(-n\underline{a}^{-\gamma}\right).$$

Therefore, the wedge in the limit is:

$$\tau \simeq \frac{n\underline{a}^{-\gamma}\exp\left(-n\underline{a}^{-\gamma}\right)}{f_e \underline{a}^{\gamma\theta-\eta}} \to 0, \tag{A.18}$$

since the numerator goes to zero and the denominator goes to infinity. Suppose instead that $\underline{a}$ has a finite limit. We obtain the expression for τ:

$$\tau = \frac{n(1-\underline{a}^{-\gamma})^{n-1}}{f_e \underline{a}^{\gamma(1+\theta)-\eta}} = \frac{n\exp\left((n-1)\log(1-\underline{a}^{-\gamma})\right)}{f_e \underline{a}^{\gamma(1+\theta)-\eta}} \to 0, \tag{A.19}$$

since the denominator has a finite limit and the numerator goes to 0.

Lastly, consider the case where $\gamma\theta = \eta$. The tax expression simplifies to:

$$\tau = \frac{1}{f_e} \cdot n\underline{a}^{-\gamma}\left(1-\underline{a}^{-\gamma}\right)^{n-1}.$$

Using Lemma A.4 and the result that $\underline{a}^{-\gamma} = \alpha(n)/n$, and $\alpha(n) \to \alpha_\infty$ we have:

$$\tau \simeq \frac{1}{f_e}\alpha(n)\exp\left(-(n-1)\alpha(n)/n\right) \tag{A.20}$$

$$\simeq \frac{1}{f_e}\alpha(n)\exp\left(-\alpha(n)\right) \to \frac{1}{f_e}\alpha_\infty e^{-\alpha_\infty}. \tag{A.21}$$

Moreover using Lemma A.4, this also proves that in the limit τ is below the wedge with agreement $(\gamma-\eta)/\gamma$. ∎

[31] This follows from $(1-M/M_e)^{n-1} = \exp\left[-(n-1)\log(M_e/M)\right]$ and using the asymptotics derived above for $\gamma\theta < \eta$: $(1-M/M_e)^{n-1} = \exp\left[-(n-1)\left(f_e^{-1}\frac{\gamma}{\gamma-\eta}n\right)^{-\frac{\gamma}{\gamma(1+\theta)-\eta}}\right] \to 1$.

B Derivations for General Equilibrium Model

Recall the definition of the average of a power function in productivity under measure $F^{(n)}$:

$$\mathcal{I}_n(M_e, \sigma) = \int_{\underline{a}}^{\infty} a^{\sigma} dF^n(a).$$

The integral with no disagreement is $\mathcal{I}_1$. We will use $\mathcal{I}_n$ when the dependence of the integral to M_e or σ is unambiguous. Under the Pareto distribution with parameter γ, we have the following result:

$$\mathcal{I}_1 = \frac{\gamma}{\gamma - \sigma} \cdot \left(\frac{M_e}{M}\right)^{\frac{\sigma}{\gamma} - 1}. \tag{B.1}$$

B.1 Model with decreasing returns to scale

B.1.1 Equilibrium

The firm optimization problem given the production function and the competitive input price w is:

$$\max_{\ell(a)} \pi(a) = a \cdot \frac{\sigma}{\sigma - 1} \ell(a)^{\frac{\sigma - 1}{\sigma}},$$

The first-order condition leads to demand for labor at the firm level:

$$\ell(a) = \left(\frac{w}{a}\right)^{-\sigma}.$$

Output and profit at the firm level are:

$$y(a) = \frac{\sigma}{\sigma - 1} \cdot w^{1-\sigma} \cdot a^{\sigma}$$
$$\pi(a) = \frac{1}{\sigma - 1} \cdot w^{1-\sigma} \cdot a^{\sigma}.$$

Market clearing on the input market yields:

$$L = M_e \cdot w^{-\sigma} \int_{\underline{a}}^{\infty} a^{\sigma} dF(a) = M_e \cdot w^{-\sigma} \cdot \mathcal{I}_1, \tag{B.2}$$

which, given (B.1), leads to the following wage in equilibrium under a Pareto distribution for F:

$$w = \left(\frac{\gamma}{\gamma - \sigma}\right)^{\frac{1}{\sigma}} \cdot \left(\frac{M_e}{L}\right)^{\frac{1}{\sigma}} \cdot \left(\frac{M_e}{M}\right)^{\frac{1}{\gamma} - \frac{1}{\sigma}}.$$

Given the equilibrium quantities, we can decompose aggregate output into the profit and labor shares. First, observe that aggregate ouput is:

$$\mathcal{C} = M_e \cdot \int_{\underline{a}}^{\infty} y(a) dF(a) = M_e \cdot \frac{\sigma}{\sigma - 1} w^{1-\sigma} \mathcal{I}_1,$$

From this expression we immediately conclude that:

$$w^{1-\sigma} \cdot \mathcal{I}_1 = \frac{\sigma - 1}{\sigma} \cdot \frac{\mathcal{C}}{M_e},$$

and we are able to simplify the ex-ante valuation of firms:

$$\begin{aligned} V^{(n)}(M_e) &= \int_{\underline{a}}^{\infty} \pi(a) dF^n(a) = \frac{1}{\sigma - 1} \cdot w^{1-\sigma} \cdot \mathcal{I}_n(M_e) \\ &= \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\mathcal{I}_n(M_e)}{\mathcal{I}_1(M_e)}. \end{aligned}$$

Finally, we express the wage as function of the equilibrium mass of firms:

$$w = \left(\frac{\gamma - \sigma}{\gamma} L \right)^{-\frac{1}{\sigma}} \cdot M^{\frac{\gamma - \sigma}{\gamma\sigma}} \cdot M_e^{\frac{1}{\gamma}}.$$

The equilibrium condition that determines entry in equilibrium is:

$$W'(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\mathcal{I}_n(M_e)}{\mathcal{I}_1(M_e)}. \tag{B.3}$$

B.1.2 Entry wedge

The optimal entry tax is such that the level of entry in the competitive equilibrium with a tax τ is the same as in the planner problem. The planner maximizes expected consumption while respecting the individual household's belief. Consumption for household j is the product of labor income and profits from its investment:

$$\mathcal{C}_j = \underbrace{\frac{\sigma - 1}{\sigma} \cdot \mathcal{C}}_{\text{labor income: } wL} + \underbrace{\frac{1}{\sigma} \cdot \frac{\mathcal{I}_n}{\mathcal{I}_1} \cdot \mathcal{C}}_{\text{firm profits: } V^{(n)}}.$$

Hence the planner optimization sets the marginal cost of labor to equal its effect on perceived aggregate output $\mathcal{C}_j$:

$$W'(M_e) = \frac{1}{\sigma} \cdot \frac{d\left(\mathcal{I}_n / \mathcal{I}_1 \cdot \mathcal{C}\right)}{dM_e} + \frac{\sigma - 1}{\sigma} \cdot \frac{d\mathcal{C}}{dM_e}. \tag{B.4}$$

To find the tax formula we take the ratio of the expression for the planner entry to the competitive equilibrium entry and find:

$$1 - \tau_n = \frac{M_e}{\mathcal{C}} \frac{\mathcal{I}_1}{\mathcal{I}_n} \cdot \frac{d\left(\mathcal{I}_n / \mathcal{I}_1 \cdot \mathcal{C}\right)}{dM_e} + (\sigma - 1) \cdot \frac{M_e}{\mathcal{C}} \frac{\mathcal{I}_1}{\mathcal{I}_n} \cdot \frac{d\mathcal{C}}{dM_e}.$$

This leads us immediately to the general optimal entry tax formula:

$$\tau_n = -\mathcal{E}_{\mathcal{I}_n} + (1 + \mathcal{E}_{\mathcal{I}_1} - \mathcal{E}_{\mathcal{C}}) - (\sigma - 1)\mathcal{E}_C \cdot \frac{\mathcal{I}_1}{\mathcal{I}_n}. \tag{B.5}$$

The asymptotic environment is similar to that of Section 2. We start by studying the asymptotics of the business stealing effect.

Lemma B.1 (Asymptotics for business stealing distortion)**.** *In the high disagreement limit ($n \to \infty$), the business stealing distortion converges to a limit that depends on the marginal cost of firm creation θ:*

- *If* $\theta\gamma < 1$, *then* $\lim_{n\to\infty} \mathcal{E}_{\mathcal{I}_n} = \mathcal{E}_{\mathcal{I}_1} = \frac{\sigma}{\gamma} - 1$.
- *If* $\theta\gamma > 1$, *then* $\lim_{n\to\infty} \mathcal{E}_{\mathcal{I}_n} = 0$.
- *If* $\theta\gamma = 1$, *then* $\lim_{n\to\infty} \mathcal{E}_{\mathcal{I}_n} = \alpha_\infty e^{-\alpha_\infty}/f_e$.

Proof. The free entry condition equation (B.3) leads to:

$$(\sigma - 1)\left(\frac{\gamma - \sigma}{\gamma}L\right)^{\frac{1-\sigma}{\sigma}} \cdot f_e = M^{\theta + \frac{\sigma-\gamma}{\gamma}\frac{\sigma-1}{\sigma}} \cdot M_e^{\frac{1-\sigma}{\gamma} - \theta} \cdot \mathcal{I}_n$$

We recast the free entry condition using $\underline{a}$ to be able to use the asymptotic results from Lemma A.4

$$\text{constant} = \underline{a}^{1-\sigma-\gamma\theta} \int_{\underline{a}}^{\infty} x^\sigma dF_n(x).$$

Writing $\tilde{\theta} = \theta + (\sigma - 1)/\gamma$ and $\tilde{\eta} = \sigma$, we recognize the first-order condition from equation Lemma A.4 and use Proposition 2. ∎

For the labor surplus term, we study the behavior of $\mathcal{I}_1/\mathcal{I}_n$.

Lemma B.2 (Asymptotics for labor surplus distortion)**.** *In the high disagreement limit ($n \to \infty$), the labor surplus distortion disappears:*

$$\lim_{n\to\infty} (\sigma - 1)\mathcal{E}_{\mathcal{C}} \frac{\mathcal{I}_1}{\mathcal{I}_n} = 0$$

Proof. Since $\tilde{\theta} > 0$, Lemma A.3 gives $\lim_{n\to\infty} M_e = \infty$. The proof of Lemma A.3 implies $\lim_{n\to\infty} \mathcal{I}_n = \infty$. Finally, because $\sigma < \gamma$,

$$\mathcal{I}_1 = \frac{\gamma}{\gamma - \sigma}\left(\frac{M_e}{M}\right)^{\frac{\sigma-\gamma}{\gamma}} \to 0$$

as $n \to \infty$. Therefore $\mathcal{I}_1/\mathcal{I}_n$ converges to 0. ∎

B.2 Differentiated goods

B.2.1 Date 1 economy

The introduction of differentiated goods in Section **??** changes the production stage. We therefore focus on the equilibrium conditions in date 1.

Firms produce a mass M of differentiated goods, indexed by (a, i), where a is firm productivity and i indexes the firms. We drop the i index when unambiguous. Household utility aggregates consumption of these goods with constant elasticity of substitution σ across goods.

At date 1, household j with total expenditure E_j solves:

$$\mathcal{C}(E_j) = \max_{\{c(a,i)\}} \left(\int_0^{M_e} \int_{F^{-1}\left(1-\frac{M}{M_e}\right)}^{\infty} c\,(a,i)^{\frac{\sigma-1}{\sigma}}\, dF(a)di \right)^{\frac{\sigma}{\sigma-1}}$$

$$\text{s.t.} \int_0^{M_e} \int_{F^{-1}\left(1-\frac{M}{M_e}\right)}^{\infty} p\,(a,i)\, c\,(a,i)\, dF(a)di \leq E_j.$$

For reasons that will soon be clear, we denote by $1/\mathcal{P}$ the Lagrange multiplier on the budget constraint. Because the objective function is homogeneous of degree one in consumption and the budget constraint is linear, $\mathcal{C}(E_j)$ is linear in E_j. Thus we have $\mathcal{C}(E_j) = E_j/\mathcal{P}$. Therefore $\mathcal{P}$ is the price of one unit of the consumption basket. We use this consumption basket as the numeraire at date 1 by normalizing $\mathcal{P} = 1$. The linearity also implies that to aggregate individual demands, it is sufficient to know the aggregate expenditure in the economy, and not the whole distribution of individual expenditures. The first-order condition in the problem above implies the demand curve:

$$c(p) = \mathcal{C}p^{-\sigma}.$$

Output for a firm with productivity a is $y = a\ell$. Firms face monopolistic competition. They maximize profits by setting prices, taking as given the demand curve from each household:

$$\max_{p(a)} \; p(a)y\,(p\,(a)) - \frac{wy\,(p\,(a))}{a} = \mathcal{C}\left[p(a)^{1-\sigma} - \frac{w}{a}p(a)^{-\sigma}\right].$$

The optimal price is therefore

$$p(a) = \frac{\sigma}{\sigma-1}\frac{w}{a}.$$

Firms charge a markup $\sigma/(\sigma-1)$ over their marginal cost w/a.

We can then compute output y, revenue py, labor expenditure $w\ell$ and profits π as functions of productivity:

$$\begin{aligned}
y &= \mathcal{C}w^{-\sigma}a^{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \\
py &= \mathcal{C}w^{1-\sigma}a^{\sigma-1}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \\
w\ell &= \frac{\sigma-1}{\sigma}\mathcal{C}w^{1-\sigma}a^{\sigma-1}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \\
\pi &= \frac{1}{\sigma}\mathcal{C}w^{1-\sigma}a^{\sigma-1}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}
\end{aligned}$$

We see that labor expenditure is a fraction $(\sigma-1)/\sigma$ of revenues, and profits make up the remaining $1/\sigma$ share.

Labor market clearing gives $\mathcal{C}(\sigma-1)/\sigma = wL$. In equilibrium, aggregate expenditure is equal to aggregate consumption, so we have:

$$\begin{aligned}\mathcal{C} &= \mathcal{C}\left(\frac{\sigma}{\sigma-1}w\right)^{1-\sigma} M_e \mathcal{I}_1(M_e, \sigma-1) \\ &= M_e^{\frac{1}{\sigma-1}}\left(\frac{\gamma}{\gamma-(\sigma-1)}\right)^{\frac{1}{\sigma-1}}\left(\frac{M_e}{M}\right)^{\frac{(\sigma-1)-\gamma}{(\sigma-1)\gamma}} \cdot L \\ &= \left(\frac{\gamma}{\gamma-(\sigma-1)}\right)^{\frac{1}{\sigma-1}} M^{\frac{1}{\sigma-1}-\frac{1}{\gamma}} M_e^{\frac{1}{\gamma}} \cdot L.\end{aligned}$$

Therefore we have $\mathcal{E}_\mathcal{C} = \mathcal{E}_w = 1/\gamma$. Alternatively, notice that the labor allocation is efficient with monopolistic competition and the aggregate production function is homogenous of degree 1 in the distribution of productivity. Because an increase in M_e increases all productivities with an elasticity $1/\gamma$, this results in an elasticity of aggregate consumption of $1/\gamma$.

B.2.2 Entry wedge

All arguments behind Proposition 3 apply, so the proposition is still valid, but with $\mathcal{I}_1$ and $\mathcal{I}_n$ now evaluated with parameter $\sigma - 1$.

With agreement, because the aggregate consumption elasticity is unchanged, the entry wedge is unchanged: $\tau_1 = (\gamma - \sigma)/\gamma$.

With speculation, the free entry condition is:

$$\left(\frac{M_e}{M}\right)^\theta = \frac{1}{\sigma}\mathcal{C}\left(\frac{\mathcal{C}}{L}\right)^{1-\sigma}\mathcal{I}_n,$$

which we can rewrite as:

$$K M_e^{\theta-(1-(\sigma-1))/\gamma} = \mathcal{I}_n,$$

where K does not depend on M_e and n. This is again the same condition as the homogeneous goods model, with σ replaced by $\sigma - 1$. The condition for the convergence of $\mathcal{E}_{\mathcal{I}_n}$ from Lemma B.1, still applies as well. In the high disagreement limit with $\theta > 1/\gamma$, the tax becomes:

$$\tau_\infty = 1 + \mathcal{E}_{\mathcal{I}_1} - \mathcal{E}_\mathcal{C} = \frac{\sigma-2}{\gamma}. \tag{B.6}$$

The upper panel of Figure B.2 shows the sign reversal of the entry wedge, as described in Proposition 5.

B.3 Knowledge externalities

B.3.1 Date 1 economy

We again focus on the date 1 economy and return to a setting with decreasing return to scale.

Now firm productivity depends on the productivity of other firms producing. In particular,

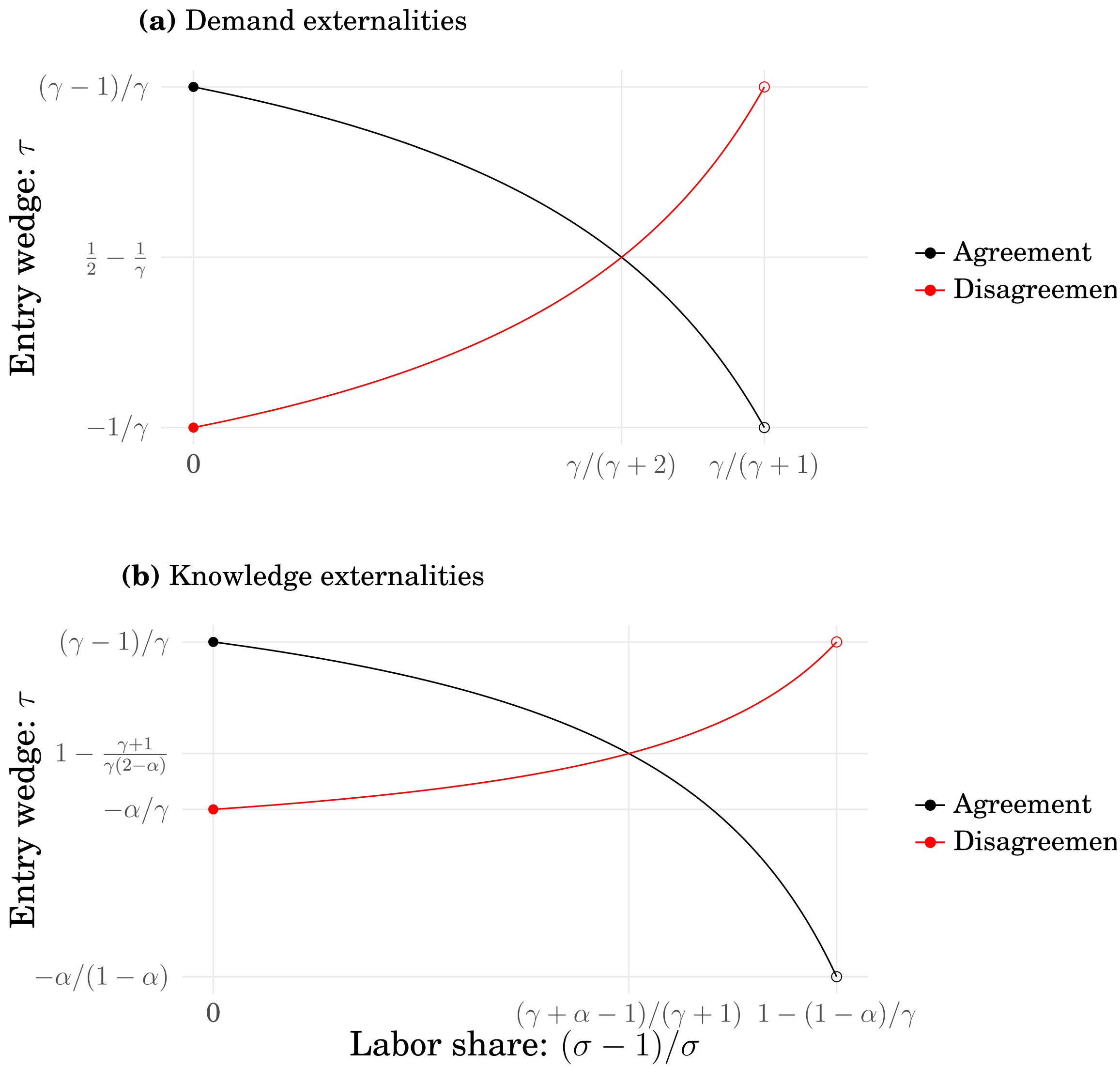

Figure B.2
Comparisons of taxes with and without disagreement

consider the production function:

$$y=\frac{\sigma}{\sigma-1}a^{1-\alpha}A^{\alpha}\ell^{\frac{\sigma-1}{\sigma}},$$

where a is firm productivity and A is an aggregator of all producing firms' productivities. $\alpha>0$ captures the intensity of knowledge spillovers. We use a Hölder mean of the productivity of all firms producing:

$$A=\left(\frac{M_e}{M}\int_{(\frac{M_e}{M})^{\frac{1}{\gamma}}}a^q dF(a)\right)^{\frac{1}{q}}.$$

Imposing $q < \gamma$ so that the integral is well defined, we have:

$$A = \left(\frac{\gamma}{\gamma - q}\right)^{\frac{1}{q}} \left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}}.$$

Our results generalize to any aggregator that is homogeneous of degree one in the productivity distribution of producing firms. Such aggregators similarly yield an elasticity $1/\gamma$ with respect to M_e.

Firms maximize their profits taking the wage as given:

$$\max_{\ell} \frac{\sigma}{\sigma - 1} a^{1-\alpha} A^{\alpha} \ell^{\frac{\sigma-1}{\sigma}} - w\ell.$$

The demand for labor is therefore

$$\ell = \left(\frac{w}{a^{1-\alpha} A^{\alpha}}\right)^{-\sigma},$$

and we have:

$$\begin{aligned}
y(a) &= \frac{\sigma}{\sigma - 1} \left(a^{1-\alpha} A^{\alpha}\right)^{\sigma} w^{1-\sigma} \\
w\ell(a) &= \left(a^{1-\alpha} A^{\alpha}\right)^{\sigma} w^{1-\sigma} = \frac{\sigma - 1}{\sigma} y(a) \\
\pi(a) &= \frac{1}{\sigma - 1} \left(a^{1-\alpha} A^{\alpha}\right)^{\sigma} w^{1-\sigma} = \frac{1}{\sigma} y(a)
\end{aligned}$$

The labor share is still $(\sigma - 1)/\sigma$.

The market clearing condition for labor is:

$$\begin{aligned}
L &= w^{-\sigma} A^{\alpha\sigma} M_e \mathcal{I}_1 \left(M_e, (1-\alpha)\sigma\right) \\
&= \left(\frac{\gamma}{\gamma - q}\right)^{\frac{\alpha\sigma}{q}} \left(\frac{M_e}{M}\right)^{\frac{\alpha\sigma}{\gamma}} M_e \frac{\gamma}{\gamma - (1-\alpha)\sigma} \left(\frac{M_e}{M}\right)^{\frac{(1-\alpha)\sigma}{\gamma} - 1} w^{-\sigma} \\
&= \left(\frac{\gamma}{\gamma - q}\right)^{\frac{\alpha\sigma}{q}} \frac{\gamma}{\gamma - (1-\alpha)\sigma} M \left(\frac{M_e}{M}\right)^{\frac{\sigma}{\gamma}} w^{-\sigma} \\
w &= \left(\frac{M}{L}\right)^{\frac{1}{\sigma}} \left(\frac{\gamma}{\gamma - q}\right)^{\frac{\alpha}{q}} \left(\frac{\gamma}{\gamma - (1-\alpha)\sigma}\right)^{\frac{1}{\sigma}} \left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}}.
\end{aligned}$$

We still have $(\sigma - 1)/\sigma\mathcal{C} = wL$, and the same elasticities: $\mathcal{E}_w = \mathcal{E}_{\mathcal{C}} = \mathcal{E}_A = 1/\gamma$.

B.3.2 Entry wedge

Proposition 3 and Lemma B.1 still apply, with $\mathcal{I}_1$ and $\mathcal{I}_n$ evaluated with parameter $(1-\alpha)\sigma$. The wedge with agreement is unchanged: $\tau_1 = (\gamma - \sigma)/\gamma$. The wedge in the high disagreement limit with $\theta > 1/\gamma$ becomes:

$$\tau_\infty = 1 + \mathcal{E}_{\mathcal{I}_1} - \mathcal{E}_{\mathcal{C}} = \frac{(1-\alpha)\sigma - 1}{\gamma}. \tag{B.7}$$

The lower panel of Figure B.2 shows the sign reversal of the entry wedge, as described in Proposition 5.

C Extensions to Simple Model

C.1 Generalizing the business-stealing effect

We now consider more general functions for the business-stealing effect. In particular, suppose the expected profit of a firm with productivity a is:

$$\pi(a) = a^{\eta}\delta\left(r(a, M_e)\right)$$

where $r(a, M_e) \equiv (1 - F(a))\, M_e$ is the ranking of the firm, or the mass of firms with productivity greater than a. We can interpret δ as being the probability of producing conditional on a firm's ranking r. The main text focused on the special case of $\delta(r) = \mathbf{1}\left\{r \leq M\right\}$.

We continue to focus on the case with $F(a) = 1 - a^{-\gamma}$.

C.1.1 Wedge under agreement

Lemma C.1. *With agreement ($n = 1$) and $\delta(M_e) = 0$, the wedge does not depend on the level of entry:*

$$\tau_1(M_e) = \frac{\gamma - \eta}{\gamma}. \tag{C.1}$$

Proof. Under agreement, $n = 1$, and we can derive an exact solution for the mass of firms entering in equilibrium, M_e. Integrating by parts, the value of a firm is:

$$\begin{aligned} V^{(1)}(M_e) &= \int_1^{\infty} \gamma x^{\eta-\gamma-1}\delta(M_e x^{-\gamma})dx \\ &= \frac{\gamma}{\gamma - \eta}\left[\delta(M_e) - \gamma M_e \int_1^{\infty} x^{\eta-2\gamma-1}\delta'(M_e x^{-\gamma})dx\right]. \end{aligned} \tag{C.2}$$

In addition, we have:

$$-M_e\frac{dV^{(1)}}{dM_e} = -\gamma M_e \int_1^{\infty} x^{\eta-2\gamma-1}\delta'(M_e x^{-\gamma})dx \tag{C.3}$$

Recalling that $\tau_1 = -M_e \frac{dV^{(1)}}{dM_e}/V^{(1)}$, we have the desired formula (C.1) for the wedge. ∎

C.1.2 Disagreement asymptotics with multiple cutoffs

We now consider the generalization of δ to allow for multiple cutoffs. In particular, suppose we have cutoffs $\underline{a}_1 < ... < \underline{a}_K$, with $\underline{a}_k \equiv F^{-1}(1 - \frac{M_k}{M_e})$, and constants $\Delta_1, ..., \Delta_K$ so that

$$\delta(r) = \sum_{k=1}^{K} \Delta_k \mathbf{1}\left\{r \leq M_k\right\} \tag{C.4}$$

Notice that this imples that $V^{(n)} = \sum_{k=1}^{K} \Delta_k V_k^{(n)}$, where $V_k^{(n)} \equiv \int_{\underline{a}_k}^{\infty} x^\eta dF_n(x)$ and

$$-M_e \frac{dV^{(n)}}{dM_e} = -\sum_{k=1}^{K} \left(\Delta_k \frac{M_k}{M_e} \cdot a^\eta \cdot \frac{F_n'}{F'}(\underline{a}_k) \right)$$

For convenience, we normalize the cost of producing blueprints so that $W(b) = f_e b^{\theta+1} M_K^{-\theta}/(\theta+1)$.

Lemma C.2. *Holding M_e constant, the wedge is larger with agreement than with disagreement.*

Proof. Apply the proof for Proposition 1 for each k. ∎

Theorem C.3 (Asymptotics for the wedge, τ, with multiple cutoffs)**.** *With business-stealing of the form (C.4), in the high disagreement limit ($n \to \infty$), the wedge between the planner problem and the competitive equilibrium converges to a finite limit.*

- *If $\gamma\theta < \eta$, then $\tau \to (\gamma - \eta)/\gamma$.*
- *If $\gamma\theta > \eta$, then $\tau \to 0$.*
- *If $\gamma\theta = \eta$, then $\tau \to \frac{1}{f_e} \sum_{k=1}^{K} \left[\Delta_k \left(\frac{M_K}{M_k} \right)^{\frac{\eta-\gamma}{\gamma}} \alpha_\infty \exp\left(-\frac{M_k}{M_K} \alpha_\infty \right) \right]$.*

Proof. Suppose $\gamma\theta < \eta$. As before, conjecture that we can write $\underline{a}_K = \alpha(n) n^{1/(\gamma(1+\theta)-\eta)}$. This yields

$$f_e \simeq \sum_{k=1}^{K} \Delta_k \left(\frac{M_k}{M_K} \right)^{\frac{\gamma-\eta}{\gamma}} \alpha^{\eta-\gamma(\theta+1)} \frac{\gamma}{\gamma-\eta} \tag{C.5}$$

Therefore, we have the asymptotics for firm entry:

$$\frac{M_e}{M_K} = \left[\frac{1}{f_e} \sum_{k=1}^{K} \Delta_k \left(\frac{M_k}{M_K} \right)^{\frac{\gamma-\eta}{\gamma}} \frac{\gamma}{\gamma-\eta} \cdot n \right]^{\frac{\gamma}{\gamma(\theta+1)-\eta}}. \tag{C.6}$$

Substituting this into the formula for the wedge, we have:

$$\tau_n = \frac{\sum_{k=1}^{K} \Delta_k \left(\frac{M_k}{M_e} \right)^{\frac{\gamma-\eta}{\gamma}} n \left(1 - \frac{M_k}{M_e} \right)^{n-1}}{f_e \left(\frac{M_e}{M_K} \right)^{\theta}} \to \frac{\gamma-\eta}{\gamma} \tag{C.7}$$

as desired.

Now suppose $\gamma\theta > \eta$. Then we have

$$\begin{aligned} \tau_n &= \frac{n \sum_{k=1}^{K} \underline{a}_k^{\eta-\gamma} \left(\underline{a}_k^{-\gamma} \right)^{n-1}}{f_e \underline{a}_K^{\gamma\theta}} \\ &= \frac{n \sum_{k=1}^{K} \left(\frac{M_K}{M_k} \right)^{\frac{\eta-\gamma}{\gamma}} \underline{a}_K^{-\gamma} \left(1 - \frac{M_k}{M_K} \underline{a}_K^{-\gamma} \right)^{n-1}}{f_e \underline{a}_K^{\gamma\theta-\eta}} \end{aligned} \tag{C.8}$$

Suppose $\underline{a}_K \to \infty$. Then we can write the first-order condition for firm creation as:

$$\frac{f_e \underline{a}_K^{\gamma\theta-\eta}}{n\underline{a}_K^{-\gamma}} = \sum_{k=1}^{K} \Delta_k \left(\frac{M_K}{M_k}\right)^{\frac{\eta-\gamma}{\gamma}} \gamma \int_1^\infty t^{\eta-\gamma-1}\left(1-\underline{a}_k^{-\gamma}t^{-\gamma}\right)^{n-1} dt \tag{C.9}$$

Since the integral on the right-hand side is bounded from above by $\int_1^\infty t^{\eta-\gamma-1}dt$, we have that $n\underline{a}_K^{-\gamma} \to \infty$, which implies that we can use a similar approximation to the proof of Proposition 2 to show that

$$\tau_n \simeq \frac{n\sum_{k=1}^{K}\left(\frac{M_K}{M_k}\right)^{\frac{\eta-\gamma}{\gamma}} \underline{a}_K^{-\gamma}\exp\left(-n\frac{M_k}{M_K}\underline{a}_K^{-\gamma}\right)}{f_e \underline{a}_K^{\gamma\theta-\eta}} \to 0 \tag{C.10}$$

Finally, suppose $\gamma\theta = \eta$. We then have

$$f_e = \sum_{k=1}^{K} \Delta_k \gamma n \left(\frac{M_K}{M_k}\right)^{\frac{\eta-\gamma}{\gamma}} \underline{a}_K^{-\gamma} \int_1^\infty t^{\eta-\gamma-1}\left(1-\frac{M_k}{M_K}\underline{a}_K^{-\gamma}t^{-\gamma}\right) \tag{C.11}$$

As before, we conjecture that $\underline{a}_K = \alpha(n)^{-1/\gamma}n^{1/\gamma}$. Then

$$\begin{aligned} f_e &\simeq \sum_{k=1}^{K} \Delta_k \left(\frac{M_K}{M_k}\right)^{\frac{\eta-\gamma}{\gamma}} \gamma\alpha(n) \int_1^\infty t^{\eta-\gamma-1}\exp\left[-(n-1)\alpha(n)n^{-1}\frac{M_k}{M_K}t^{-\gamma}\right] dt \\ &\to \sum_{k=1}^{K} \Delta_k \left(\frac{M_K}{M_k}\right)^{\frac{\eta-\gamma}{\gamma}} \gamma\alpha_\infty \int_1^\infty t^{\eta-\gamma-1}\exp\left[-\frac{M_k}{M_K}\alpha_\infty t^{-\gamma}\right] dt \end{aligned} \tag{C.12}$$

By analogous reasoning to the proof in Proposition 2, we can obtain a finite bound on α_∞. We therefore have the wedge

$$\begin{aligned} \tau_n &= \frac{1}{f_e}\sum_{k=1}^{K}\left[\Delta_k n \left(\frac{M_K}{M_k}\right)^{\frac{\eta-\gamma}{\gamma}} \underline{a}_K^{-\gamma}\left(1-\frac{M_k}{M_K}\underline{a}_K^{-\gamma}\right)\right] \\ &\to \frac{1}{f_e}\sum_{k=1}^{K}\left[\Delta_k \left(\frac{M_K}{M_k}\right)^{\frac{\eta-\gamma}{\gamma}} \alpha_\infty\left(1-\frac{M_k}{M_K}\alpha_\infty\right)\right] \end{aligned} \tag{C.13}$$

as desired. ∎

C.1.3 Disagreement asymptotics with continuous business-stealing

We now consider a continuous function for the business-stealing effect

$$\delta(r) = \begin{cases} 1 & \text{if } r < 1 \\ r^{-\zeta} & \text{if } r \geq 1 \end{cases} \tag{C.14}$$

so that ζ parameterizes the business stealing effect for low productivity firms with $r \geq 1$. Larger ζ implies that low productivity firms have a lower probability of producing, with $\zeta = 0$ corresponding to the case with no business-stealing effect. With $\zeta \to \infty$, this converges to the benchmark step function business-stealing effect with $M = 1$.

We now normalize the cost of producing blueprints so that $W(b) = f_e b^{\theta+1}/(\theta+1)$, and define $\underline{a} \equiv M_e^{\frac{1}{\gamma}}$ to be the cutoff above which $\delta(a, M_e) = 1$, i.e. firms produce with probability one.

It will be convenient to consider the decomposition $V^{(n)} = V_L^{(n)} + V_U^{(n)}$, where

$$V_L^{(n)} \equiv \gamma n M_e^{-\zeta} \int_1^{\underline{a}} x^{\eta-\gamma(1-\zeta)-1}(1-x^{-\gamma})^{n-1}dx$$
$$V_U^{(n)} \equiv \gamma n \int_{\underline{a}}^{\infty} x^{\eta-\gamma-1}(1-x^{-\gamma})^{n-1}dx$$

capture the expected profit conditional on having productivity below and above $\underline{a}$ respectively. We can write

$$V_L^{(n)} = \gamma n \underline{a}^{\eta-\gamma(\theta+1)} \int_{\underline{a}^{-1}}^{1} t^{\eta-\gamma(1-\zeta)-1}(1-\underline{a}^{-\gamma}t^{-\gamma})^{n-1}dt \tag{C.15}$$

$$V_U^{(n)} = \gamma n \underline{a}^{\eta-\gamma(\theta+1)} \int_{1}^{\underline{a}^{-1}} t^{\eta-\gamma-1}(1-\underline{a}^{-\gamma}t^{-\gamma})^{n-1}dt \tag{C.16}$$

Moreover, since

$$\begin{aligned}
\frac{dV_L^{(n)}}{dM_e} &= -\zeta M_e^{-1} V_L^{(n)} + \gamma n M_e^{-\zeta} M_e^{\frac{\eta-\gamma(1-\zeta)-1}{\gamma}} \left(1-M_e^{-1}\right)^{n-1} \\
&= -\zeta M_e^{-1} V_L^{(n)} - \frac{dV_U^{(n)}}{dM_e}
\end{aligned}$$

we have that

$$-M_e \frac{dV^{(n)}}{dM_e} = \zeta V_L^{(n)} \tag{C.17}$$

Theorem C.4 (Asymptotics for the wedge, τ, with continuous business-stealing)**.** *Suppose we have business-stealing of the form (C.14) and $\zeta > \frac{\gamma-\eta}{\gamma}$. In the high disagreement limit ($n \to \infty$), the wedge between the planner problem and the competitive equilibrium converges to a finite limit.*

- *If $\gamma\theta < \eta$, then $\tau \to (\gamma-\eta)/\gamma$.*
- *If $\gamma\theta \geq \eta$, then $\tau \to 0$.*

Proof. Suppose $\gamma\theta < \eta$. Conjecture that $\underline{a} = \alpha(n) n^{1/(\gamma(1+\theta)-\eta)}$. We have from the proof of Proposition 2 that $V_U^{(n)} \to \alpha^{\eta-\gamma}\frac{\gamma}{\gamma-\eta}$. In addition, we have from (C.15) that $V_U^{(n)} \to \alpha^{\eta-\gamma}\frac{\gamma}{\eta-\gamma(1-\zeta)}$. Therefore, we have

$$f_e = \left(\frac{\gamma}{\gamma-\eta} - \frac{\gamma}{\gamma(1-\zeta)-\eta}\right)\alpha^{\eta-\gamma(\theta+1)} \tag{C.18}$$

which verifies the conjecture. We thus have the asymptotic wedge

$$\tau_n = \zeta \frac{V_L^{(n)}}{V^{(n)}} \to \frac{\gamma-\eta}{\gamma} \tag{C.19}$$

as desired.

Suppose $\gamma\theta > \eta$. Suppose $\underline{a} \to \infty$. Then we can rewrite (2.8) as

$$n\underline{a}^{-\gamma} = \frac{f_e \underline{a}^{\gamma\theta-\eta}/\gamma}{\int_{\underline{a}^{-1}}^{1} t^{\eta-\gamma(1-\zeta)-1}(1-\underline{a}^{-\gamma}t^{-\gamma})^{n-1}dt + \int_1^{\infty} t^{\eta-\gamma-1}(1-\underline{a}^{-\gamma}t^{-\gamma})^{n-1}dt} \tag{C.20}$$

The two terms in the denominator of the right-hand side are bounded from above by $\int_0^1 t^{\eta-\gamma(1-\zeta)-1}dt$ and $\int_1^{\infty} t^{\eta-\gamma-1}dt$ respectively, which implies that $n\underline{a}^{-\gamma} \to \infty$. Using the approximation $(1 - \underline{a}^{-\gamma})^{n-1} \simeq \exp(-n\underline{a}^{-\gamma})$, we have that

$$\tau_n = \frac{\zeta}{f_e \underline{a}^{\gamma\theta}} V_L^{(n)} \to 0. \tag{C.21}$$

If $\underline{a}$ has a finite limit, we can show that $V_L^{(n)} \to 0$ since $n(1-\underline{a}^{-\gamma}t^{-\gamma})^{n-1} \to 0$ for $t \in (\underline{a}^{-1}, 1)$, which implies that $\tau_n \to 0$ as well.

Suppose $\gamma\theta = \eta$. Using an analogous proof to Lemma A.4, we can show that $\underline{a}_n = \alpha(n)^{-1/\gamma} n^{1/\gamma}$ where $\alpha(n)$ has a finite limit α_∞. Since we can bound $V_L^{(n)}$ from above by:

$$\begin{aligned} V_L^{(n)} &= \gamma\alpha(n) \int_{\underline{a}^{-1}}^{1} t^{\eta-\gamma(1-\zeta)-1}(1-\underline{a}^{-\gamma}t^{-\gamma})^{n-1}dt \\ &\leq \gamma\alpha(n) \int_{\underline{a}^{-1}}^{1} t^{\eta-\gamma(1-\zeta)-1}dt \to \frac{\gamma\alpha_\infty}{\eta-\gamma(1-\zeta)} \end{aligned}$$

we have that $\tau_n = \frac{\zeta V_L^{(n)}}{f_e \underline{a}^{\gamma\theta}} \to 0$. ∎

C.2 Results with a zero cutoff profit condition

Our results are robust to an extension in which the marginal firm earns zero profits. Our baseline model specifies that the M most productive firms will be allowed to produce, which allows for tractability but results in the marginal firm earning positive profits, $\pi(\underline{a}) > 0$. We now augment the model with an intermediate stage where firms, after entering the market, compete to be among one of the M firms producing. The competition stage ensures that the business-stealing externality remains. We keep the belief and production structure of the model intact, and show that the main features of the entry wedge remain unchanged despite the introduction of a zero cutoff profit (ZCP) condition for the marginal firm.

In the new intermediate decision stage, firms can use some of their production as advertisement to reach consumers, a deadweight loss. Only the M firms that spend the most on advertising produce in equilibrium. Formally, each firm chooses how much of its production to use on advertisement, $h_i \leq \pi(a_i)$. In doing so, firms take as given the equilibrium level $\underline{h}$ of advertising necessary to attract consumers. Their profit function is therefore $\pi(a_i)\mathbf{1}\{h_i \geq \underline{h}\} - h_i$. The optimal advertisement choice is $h_i = \underline{h}$ if $\pi(a_i) \geq \underline{h}$ and 0 otherwise. The equilibrium value of $\underline{h}$ is such that exactly M firms choose to spend on advertisement. Keeping the definition of the production cutoff $\underline{a}$ from earlier, this implies

$$\underline{h} = \pi(\underline{a}). \tag{C.22}$$

Firms must spend the profits of the marginal firm to be able to produce, resulting in zero profits for the marginal firm.

C.2.1 General derivations

Firm value in this model is modified to account for the cost of advertisement:

$$\widetilde{V}^{(n)}(M_e) = \int_{\underline{a}}^{\infty} (\pi(a) - \pi(\underline{a}))\, dF^n(a), \tag{C.23}$$

We can define the corresponding integral $\widetilde{\mathcal{I}}_n$. With this new definition of firm value, the remainder of the competitive equilibrium and the planner problem are unchanged. In particular, the entry wedge is $\tau = -\mathcal{E}_{\widetilde{\mathcal{I}}_n}$.

Decompose firms' valuations into the revenue (from (C.23)) and advertising cost components:

$$V^{(n)}(M_e) = \int_{F^{-1}\left(1-\frac{M}{M_e}\right)}^{\infty} \pi(a) dF_n(a) - \left(\frac{M}{M_e}\right)^{1-\frac{\eta}{\gamma}} \cdot n \left(1 - \frac{M}{M_e}\right)^{n-1}. \tag{C.24}$$

The first derivative of $V^{(n)}$ is:

$$-M_e \cdot \frac{dV^{(n)}(M_e)}{dM_e} = \frac{\eta}{\gamma} \left(\frac{M_e}{M}\right)^{\frac{\eta}{\gamma}} \cdot \left[1 - \left(1 - \frac{M}{M_e}\right)^n\right].$$

Using the free-entry condition, $V^{(n)}(M_e) = W'(M_e)$, we have following formula for the wedge between planner problem and competitive equilibrium:

$$\tau_n(M_e) = \frac{\eta}{\gamma} \cdot \frac{1}{f_e} \cdot \left(\frac{M_e}{M}\right)^{\frac{\eta}{\gamma}-\theta} \cdot \left[1 - \left(1 - \frac{M}{M_e}\right)^n\right]. \tag{C.25}$$

C.2.2 Wedge and firm entry

Lemma C.5. *In the model with a ZCP condition and agreement ($n = 1$), the wedge between the competitive equilibrium and the planner problem is:*

$$\tau = \frac{\gamma - \eta}{\gamma}.$$

Proof. The free-entry condition with $n = 1$ gives us:

$$\left(\frac{M_e}{M}\right)^{\frac{\gamma(1+\theta)-\eta}{\gamma}} = \frac{1}{f_e} \cdot \frac{\eta}{\gamma - \eta}.$$

Given the derivation of the wedge in (C.25), we have:

$$\tau_1(M_e) = \frac{\eta}{\gamma} \cdot \frac{1}{f_e} \cdot \left(\frac{M_e}{M}\right)^{\frac{\eta}{\gamma}-\theta} \cdot \frac{M}{M_e} = \frac{\gamma - \eta}{\gamma}, \tag{C.26}$$

where we have used our equilibrium solution for M_e/M. ∎

Lemma C.6. *In the high disagreement limit ($n \to \infty$), the mass of entrants also increases and goes to infinity:* $\lim_{n\to\infty} M_e = \infty$.

Proof. We adapt the proof from Lemma A.3, again defining the sequence $\underline{a}_n = (M_e/M)^{1/\gamma}$ and showing that $\underline{a}_n \to \infty$. Equation (C.24) implies the implicit definition of the sequence $(\underline{a}_n)_n$ in this case:

$$f_e \underline{a}_n^{\gamma(\theta+1)-\eta} = \gamma n \int_1^\infty t^{\eta-\gamma-1} \left(1 - \underline{a}_n^{-\gamma} t^{-\gamma}\right)^{n-1} dt - n(1 - \underline{a}_n^{-\gamma})^{n-1}.$$

Assume that $\underline{a}_n$ has a finite limit that is strictly larger than zero, $\underline{a}_\infty > 0$. Then there exists N large enough such that $\forall n > N, \underline{a}_n > A = \underline{a}_\infty - \epsilon > 0$. For any arbitrary threshold T_n we have

$$I_n > n \left[\frac{\gamma}{\gamma - \eta} \cdot T_n^{\eta-\gamma} \left(1 - A^{-\gamma} T_n^{-\gamma}\right)^{n-1} - 1 \right].$$

As in Lemma A.3, we conclude by considering the threshold $T_n = n^{1/\gamma}$. ∎

Lemma C.7 (Asymptotics for firm creation)**.** *In the high disagreement limit ($n \to \infty$), we have the following asymptotics for the mass of firms created M_e:*

- *If $\gamma\theta < \eta$, then $\lim_{n\to\infty} M_e/M = \alpha_\infty^\gamma n^{\frac{\gamma}{\gamma(1+\theta)-\eta}}$.*
- *If $\gamma\theta = \eta$, then $\lim_{n\to\infty} M_e/M = \alpha_\infty^{-1} n$.*

In each case, α_∞ is a constant defined below.

Proof. We adapt the proof from Lemma A.4. Starting from (2.8), and using $\underline{a}$:

$$f_e \simeq \gamma \underline{a}^{\eta-\gamma(\theta+1)} n \int_1^\infty (t^\eta - 1)\, t^{-\gamma-1} \exp\left(-(n-1)\underline{a}^{-\gamma} t^{-\gamma}\right) dt,$$

where we have used the fact that $\underline{a} \to \infty$ and $\log(1-x) = -x + \mathcal{O}(x^2)$. To find a solution, we guess that asymptotically $\underline{a} \simeq \alpha(n) n^{\frac{1}{\gamma(1+\theta)-\eta}}$ and show that $\alpha(n)$ converges to a finite limit α. The first-order condition becomes:

$$f_e \simeq \gamma \alpha(n)^{\eta-\gamma(\theta+1)} \int_1^\infty (t^\eta - 1)\, t^{-\gamma-1} \exp\left(-\alpha(n)^{-\gamma} \frac{n-1}{n^{\frac{\gamma}{\gamma(1+\theta)-\eta}}} t^{-\gamma}\right) dt.$$

If $\gamma\theta < \eta$, then the exponential term converges to zero and we have:

$$\alpha_\infty = \left(\frac{1}{f_e} \cdot \frac{\eta}{\gamma - \eta}\right)^{\frac{1}{\gamma(1+\theta)-\eta}}. \tag{C.27}$$

If $\gamma\theta = \eta$, then $\underline{a}$ is defined by:

$$f_e = \gamma \underline{a}_n^{-\gamma} n \int_1^\infty t^{\eta-\gamma-1} \left(1 - \underline{a}^{-\gamma} t^{-\gamma}\right)^{n-1} dt.$$

We guess and verify that $\underline{a}_n = \alpha(n)^{-1/\gamma} n^{1/\gamma}$, and $\alpha(n)$ has a finite limit, α_∞:

$$f_e = \gamma \alpha_\infty \int_1^\infty (t^\eta - 1)\, t^{-\gamma-1} e^{-\alpha_\infty t^{-\gamma}} dt,$$

where we took the limit when $n \to \infty$. Moreover we are able to bound the wedge above the wedge with agreement using a bound on α_∞:

$$\frac{\alpha_\infty e^{-\alpha_\infty}}{f_e} < \frac{\gamma - \eta}{\eta}, \tag{C.28}$$

which verifies that α_∞ is finite. ∎

We now show that, despite the presence of the ZCP, we have the same result as in Proposition 2.

Theorem C.8. *In the high disagreement limit ($n \to \infty$), the wedge between the planner problem and the competitive equilibrium converges to a finite limit.*

- *If $\gamma\theta < \eta$, then $\tau \to (\gamma - \eta)/\gamma$.*
- *If $\gamma\theta > \eta$, then $\tau \to 0$.*
- *If $\gamma\theta = \eta$, then $\tau \to \frac{\eta}{\gamma}\frac{1}{f_e}e^{-\alpha_\infty}$.*

Proof. If $\gamma\theta > \eta$, then given equation (C.25), we use that $M_e \to \infty$ to conclude that $\lim_{n\to\infty} \tau = 0$.

If $\gamma\theta < \eta$, then we can use the asymptotics from C.7 and the formula for the wedge from (C.25):

$$\tau_n(M_e) \simeq \frac{\eta}{\gamma}\frac{1}{f_e}\alpha_\infty^{\eta-\theta\gamma} \cdot n^{\frac{\eta-\theta\gamma}{\gamma(1+\theta)-\eta}} \cdot \left[1 - \left(1 - \alpha_\infty^{-\gamma} n^{\frac{-\gamma}{\gamma(1+\theta)-\eta}}\right)^n\right] \tag{C.29}$$

$$\simeq \frac{\eta}{\gamma}\frac{1}{f_e}\alpha_\infty^{\eta-\theta\gamma} \cdot n^{\frac{\eta-\theta\gamma}{\gamma(1+\theta)-\eta}} \cdot \left[1 - \exp\left(-\alpha_\infty^{-\gamma} n^{\frac{\gamma\theta-\eta}{\gamma(1+\theta)-\eta}}\right)\right] \tag{C.30}$$

$$\simeq \frac{\eta}{\gamma}\frac{1}{f_e}\alpha_\infty^{\eta-\theta\gamma} \cdot n^{\frac{\eta-\theta\gamma}{\gamma(1+\theta)-\eta}} \cdot \alpha_\infty^{-\gamma} n^{\frac{\gamma\theta-\eta}{\gamma(1+\theta)-\eta}} \to \frac{\eta}{\gamma}\frac{1}{f_e}\alpha_\infty^{\eta-\gamma(1+\theta)}. \tag{C.31}$$

Using the definition of α_∞ from the proof above, we conclude $\lim_{n\to\infty} \tau_n(M_e) = (\gamma - \eta)/\gamma$.

In the knife-edge case with $\gamma\theta = \eta$, we have:

$$\tau_n(M_e) \simeq \frac{\eta}{\gamma}\frac{1}{f_e} \cdot \left[1 - \left(1 - \alpha_\infty n^{-1}\right)^n\right] \to \frac{\eta}{\gamma}\frac{1}{f_e} \cdot e^{-\alpha_\infty}. \tag{C.32}$$

We can bound the wedge in the limit: $\lim_{n\to\infty} \tau_n(M_e) < \alpha_\infty^{-1} \cdot (\gamma - \eta)/\gamma$. ∎

Our conclusions are therefore robust to including competition to enter. Intuitively, marginal firms drive the externality in both settings. In our baseline, the externality operates at the extensive margin: more entry displaces the profits of excluded marginal firms. In this model, the externality is on the intensive margin: firm entry increases the productivity of the marginal firm and thus advertisement costs for all producing firms.

D Extensions to General Equilibrium Model

D.1 Elastic labor supply

D.1.1 Setting and equilibrium

We now consider the case of variable labor supply. Households can provide labor L by exerting an effort cost $S(L)$. We assume

$$S'(L) = f_l \left(\frac{L}{L_0}\right)^{1/\kappa},$$

where κ is the Frisch elasticity of labor supply. As κ converges to 0, the model converges to a constant labor supply L_0. The remainder of the model is unchanged.

In equilibrium, we have $S'(L) = w$, which implies that $\mathcal{E}_L = \kappa\mathcal{E}_w$. Noting that we still have a constant labor share ($(\sigma-1)/\sigma\mathcal{C} = wL$), we also have $\mathcal{E}_C = \mathcal{E}_w + \mathcal{E}_L = (1+\kappa)\mathcal{E}_w$.

Market clearing for labor yields:

$$L = M_e \cdot w^{-\sigma} \int_{\underline{a}}^{\infty} a^{-\sigma} dF(a) = M_e \cdot w^{-\sigma} \cdot \mathcal{I}_1$$

Which given the expression for $\mathcal{I}_1$ in (B.1) under a Pareto leads to the following restriction:

$$wL^{1/\sigma} = \left(\frac{\gamma}{\gamma-\sigma}\right)^{\frac{1}{\sigma}} \cdot M_e^{\frac{1}{\sigma}} \cdot \left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}-\frac{1}{\sigma}}$$

Using the labor supply equation, we obtain:

$$\mathcal{E}_w = \frac{1}{\gamma}\frac{\sigma}{\sigma+\kappa}$$
$$\mathcal{E}_\mathcal{C} = \frac{1}{\gamma}\frac{\sigma+\kappa\sigma}{\sigma+\kappa}$$

As we increase the elasticity of labor supply, the wage becomes less elastic to firm entry and consumption becomes more elastic to firm entry as it becomes less costly to expand labor.

D.1.2 Entry wedge

Asymptotics. To study the asymptotic behavior of the wedge, recall the free entry condition:

$$f_e \left(\frac{M_e}{M}\right)^{\theta} = \frac{1}{\sigma-1} w^{1-\sigma} \mathcal{I}_n.$$

We define

$$\tilde{\theta} = \theta + \frac{\sigma-1}{\gamma}\frac{\sigma}{\kappa+\sigma} \geq 0$$

and recognize the free entry condition from the pure business stealing model. This guarantees that $\lim_{n\to\infty} \mathcal{I}_1/\mathcal{I}_n = 0$. In addition, $\mathcal{E}_{\mathcal{I}_n}$ converges to 0 if:

$$\gamma\tilde{\theta} > \sigma \tag{D.1}$$

$$\Leftrightarrow \gamma\theta > 1 + (\sigma - 1)\frac{\kappa}{\kappa + \sigma} \tag{D.2}$$

Planner problem and wedge. The planner objective is now:

$$\max_{M_e} \frac{1}{\sigma}\mathcal{C}\frac{\mathcal{I}_n}{\mathcal{I}_1} + \frac{\sigma - 1}{\sigma}\mathcal{C} - S(L) - W(M_e)$$

The first-order condition is:

$$W'(M_e) = \frac{1}{\sigma}\mathcal{C}\frac{\mathcal{I}_n}{\mathcal{I}_1}\frac{\mathcal{I}_n'}{\mathcal{I}_n} + \frac{1}{\sigma}\mathcal{C}'\frac{\mathcal{I}_n}{\mathcal{I}_1} - \frac{1}{\sigma}\mathcal{C}\frac{\mathcal{I}_n}{\mathcal{I}_1}\frac{\mathcal{I}_1'}{\mathcal{I}_1} + \frac{\sigma - 1}{\sigma}\mathcal{C}' - \underbrace{S'(L)L}_{\frac{\sigma-1}{\sigma}\mathcal{C}}\frac{1}{L}\frac{dL}{dM_e}.$$

The wedge is similar to Proposition 3. However, the appropriability term now accounts for the utility cost of expanding the labor supply:

$$\tau_n = -\mathcal{E}_{\mathcal{I}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - (\sigma - 1)\underbrace{(\mathcal{E}_{\mathcal{C}} - \mathcal{E}_L)}_{\mathcal{E}_w}\frac{\mathcal{I}_1}{\mathcal{I}_n}.$$

With agreement, we have:

$$\tau_1 = 1 - \mathcal{E}_{\mathcal{C}} - (\sigma - 1)\mathcal{E}_w = \frac{\gamma - \sigma}{\gamma}. \tag{D.3}$$

With high speculation and large θ, we have:

$$\tau_\infty = (\sigma - 1)\mathcal{E}_w = \frac{\sigma - 1}{\gamma}\frac{\sigma}{\kappa + \sigma}. \tag{D.4}$$

The wedge with agreement is unchanged from the baseline model. The wedge with high disagreement is increasing in σ and decreasing in γ, as is the case in Proposition 4. The elasticity of labor supply does not affect the wedge with agreement, but with disagreement a higher κ lowers the tax. As labor supply becomes more elastic, the wage becomes less responsive to entry, and firms have less influence on each other through general equlibrium effects. With perfectly elastic labor supply, there are no general equilibrium effects, and the economy with disagreement is efficient, i.e. $\tau_\infty = 0$.

D.2 Variable number of participating firms

D.2.1 Setting and equilibrium

We study a model where the number of participating firms, M, responds to firm creation M_e, which can be interpreted as households' consumption bundles becoming more or less concentrated as more firms enter the economy. We assume that M varies exogenously with the level of firm entry M_e:

$$M = \frac{1}{M_0^{\chi-1}} \cdot M_e^{\chi},$$

where χ is the elasticity of firms producing to firms created and M_0 a normalization constant. We assume that $\chi \leq 1$ such that we always have $M \leq M_e$.

The cost of creating a firm only depends on M_0 through:

$$W'(M_e) = f_e \left(\frac{M_e}{M_0}\right)^{\theta}.$$

The productivity threshold to produce is now:

$$\underline{a} := F^{-1}\left(1 - \frac{M}{M_e}\right) = \left(\frac{M_e}{M_0}\right)^{\frac{1-\chi}{\gamma}}.$$

The model still features a constant labor share and firm profits are still isoelastic in the productivity:

$$\pi(a) = \frac{1}{\sigma - 1} \cdot w^{1-\sigma} \cdot a^{\sigma} = \frac{1}{\sigma} \frac{\mathcal{C}}{M_e} \cdot \frac{a^{\sigma}}{\mathcal{I}_1},$$

where we have redefined the integrals $\mathcal{I}_1$ and $\mathcal{I}_n$ to adjust for the new expressions for the productivity threshold $\underline{a}$:

$$\mathcal{I}_n(\chi) = \int_{\left(\frac{M_e}{M_0}\right)^{\frac{1-\chi}{\gamma}}}^{\infty} a^{\sigma} dF_n(a)$$

$$\mathcal{I}_1(\chi) = \frac{\gamma}{\gamma - \sigma} \cdot \left(\frac{M_e}{M_0}\right)^{(\chi-1)\frac{\gamma-\sigma}{\gamma}}.$$

The market-clearing condition $L = M_e w^{-\sigma} \mathcal{I}_1$ implies the equilibrium wage:

$$w = \left(\frac{\gamma}{\gamma - \sigma}\right)^{\frac{1}{\sigma}} \cdot L^{-\frac{1}{\sigma}} M_0^{(1-\chi)\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)} \cdot M_e^{\frac{\chi}{\sigma} + \frac{1-\chi}{\gamma}},$$

so that the labor elasticity is :

$$\mathcal{E}_w = \frac{1}{\gamma} + \chi \cdot \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right).$$

We obtain aggregate consumption by aggregating individual output $\mathcal{C} = M_e \sigma/(\sigma-1) w^{1-\sigma} \mathcal{I}_1$, which yields equilibrium aggregate consumption and elasticity:

$$\mathcal{C} = \frac{\sigma}{\sigma - 1} \frac{\gamma}{\gamma - \sigma} \cdot L^{\frac{\sigma-1}{\sigma}} \cdot M_0^{(1-\chi)\left(\frac{1-\sigma}{\sigma} + \frac{\gamma-1}{\gamma}\right)} \cdot M_e^{\frac{1}{\gamma} + \chi \cdot \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)}$$

$$\mathcal{E}_{\mathcal{C}} = \mathcal{E}_w = \frac{1}{\gamma} + \chi \cdot \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)$$

D.2.2 Entry wedge

Given the constant labor share and isoelastic profits, we can apply Proposition 3 and obtain the wedge:

$$\tau_n = -\mathcal{E}_{\mathcal{I}_n(\chi)} + 1 + \mathcal{E}_{\mathcal{I}_1(\chi)} - \mathcal{E}_{\mathcal{C}} - (\sigma-1)\mathcal{E}_C \cdot \frac{\mathcal{I}_1(\chi)}{\mathcal{I}_n(\chi)}. \tag{D.5}$$

From the expression for $\mathcal{I}_n$ we have the following change in the elasticities:

$$\mathcal{E}_{\mathcal{I}_n(\chi)} = (1-\chi)\mathcal{E}_{\mathcal{I}_n(\chi=0)} = (1-\chi)\mathcal{E}_{\mathcal{I}_n} \tag{D.6}$$

Asymptotics. We now turn to the high disagreement limit. The first-order condition for firm creation is:

$$\begin{aligned} f_e\left(\frac{M_e}{M_0}\right)^\theta &= \frac{1}{\sigma-1}w^{1-\sigma}\mathcal{I}_n(\chi) \\ \iff \quad \text{constant} &= \underline{a}^{-\theta\frac{\gamma}{1-\chi}+(1-\sigma)\frac{\gamma}{1-\chi}\left(\frac{1}{\gamma}+\chi\left(\frac{1}{\sigma}-\frac{1}{\gamma}\right)\right)}\int_{\underline{a}}^{\infty} a^\sigma dF_n(a). \end{aligned}$$

We define:

$$\tilde{\theta} = \frac{1}{1-\chi}\left(\theta + \frac{\sigma-1}{\gamma} + \chi(\sigma-1)\left(\frac{1}{\sigma}-\frac{1}{\gamma}\right)\right),$$

and recognize the entry condition of the baseline model. We apply our previous results, changing the condition for $\mathcal{E}_{\mathcal{I}_n(\chi)} \to 0$ to $\gamma\tilde{\theta} > \sigma$, which reduces to:

$$\gamma(\theta+\chi) > 1 + \chi\left(\frac{\gamma}{\sigma}-1\right). \tag{D.7}$$

If this condition is satisfied, then $\lim_{n\to\infty}\mathcal{E}_{\mathcal{I}_n(\chi)} = 0$. When the inequality is reversed, $\lim_{n\to\infty}\mathcal{E}_{\mathcal{I}_n(\chi)} = \mathcal{E}_{\mathcal{I}_1(\chi)} = (1-\chi)(\sigma-\gamma)/\gamma$. With equality, the elasticity admits a finite limit between these two values.

Behavior of the wedge. The wedge with agreement is

$$\tau_1 = 1 - \sigma\mathcal{E}_{\mathcal{C}} = \frac{\gamma-\sigma}{\gamma} - \chi\sigma\left(\frac{1}{\sigma}-\frac{1}{\gamma}\right). \tag{D.8}$$

The wedge with high disagreement, when θ is large enough, is:

$$\tau_\infty = 1 + \mathcal{E}_{\mathcal{I}_1(\chi)} - \mathcal{E}_{\mathcal{C}} = \frac{\sigma-1}{\gamma} + \chi(\sigma-1)\left(\frac{1}{\sigma}-\frac{1}{\gamma}\right) \tag{D.9}$$

As long as $|\chi| < 1$, Proposition 4 still holds with the generalized version of $\mathcal{I}_n$. A higher elasticity of firm participation with respect to firm entry leads to opposite results with and without disagreement. A large elasticity χ dampens the wedge with agreement because it diminishes business stealing. However it leads to a higher wedge with disagreement: in response to firm entry, labor demand responds at the intensive margin with more productive firms, and at the extensive margin with more participating firms.

D.3 Advertising to participate

We augment the previous model with an intermediate stage after market entry when firms compete to be one of M firms producing, as in Appendix C.2.

D.3.1 Setting and equilibrium

In the new intermediate decision stage, firms choose to spend on advertisement to reach consumers. We assume that only the M firms that spend the most on advertising produce in equilibrium. Formally if a firm with productivity a_i spends h_i in advertising, its profit is: $\pi(a_i)\mathbf{1}\{h_i \geq \underline{h}\} - h_i$. Hence there is a threshold level of advertising, $\underline{h}$, below which firms cannot reach any consumers and above which firms do produce. Firms take the threshold as given and decide on their choice of advertising. Thus the advertising equilibrium is such that the threshold matches the profit of the marginal firm: $\underline{h} = \pi(\underline{a})$.

Profits are modified with respect to the standard model of Section B.1 to incorporate the advertisement payments:

$$\pi(a) = \frac{1}{\sigma} w^{1-\sigma} \left(a^\sigma - \underline{a}^\sigma\right),$$

The ex-ante firm valuation is therefore:

$$V^{(n)}(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1},$$

where we define the adjusted integral $\widetilde{\mathcal{I}}_n$ as

$$\widetilde{\mathcal{I}}_n(M_e) = \int_{\left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}}}^{\infty} \left(a^\sigma - \left(\frac{M_e}{M}\right)^{\frac{\sigma}{\gamma}}\right) dF_n(a).$$

The entry condition in the competitive equilibrium is now:

$$W'(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1}.$$

D.3.2 Entry wedge

The planner maximizes expected utility:

$$\max_{M_e} \; - W(M_e) + \frac{1}{\sigma}\mathcal{C}\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1} + \frac{\sigma - 1}{\sigma}\mathcal{C}$$

The corresponding optimality condition is:

$$W'(M_e) = \frac{1}{\sigma}\mathcal{C}\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\widetilde{\mathcal{I}}_n'}{\widetilde{\mathcal{I}}_n} + \frac{1}{\sigma}\mathcal{C}'\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1} - \frac{1}{\sigma}\mathcal{C}\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\mathcal{I}_1'}{\mathcal{I}_1} + \frac{\sigma - 1}{\sigma}\mathcal{C}'.$$

The wedge is therefore

$$\tau_n = -\mathcal{E}_{\widetilde{\mathcal{I}}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - (\sigma - 1)\mathcal{E}_{\mathcal{C}}\frac{\mathcal{I}_1}{\widetilde{\mathcal{I}}_n}. \tag{D.10}$$

With agreement, the wedge is:

$$\tau_1 = 1 - \mathcal{E}_{\mathcal{C}} - (\sigma - 1)\mathcal{E}_{\mathcal{C}}\frac{\gamma}{\sigma} = \frac{1}{\sigma} - \frac{1}{\gamma}, \tag{D.11}$$

where we have used that, as in the baseline model, $\mathcal{E}_{\mathcal{C}} = \mathcal{E}_w = 1/\gamma$. However, the profit share is now lower because of advertisement costs. A fraction $(\sigma-1)/\sigma$ of output accrues to labor but only a fraction $1/\gamma$ ($< 1/\sigma$) is collected as profits. The larger importance of labor relative to profits gives rise to a lower tax with agreement. As in the standard model, τ_1 is decreasing in σ and increasing in γ.

To derive the wedge with high disagreement, notice that $\mathcal{I}_1/\widetilde{\mathcal{I}}_n \to 0$ and therefore, if $\theta > 1/\gamma$, then τ_∞ is unchanged from the standard model:

$$\tau_\infty = \frac{\sigma - 1}{\gamma}. \tag{D.12}$$

Therefore, Proposition 4 still holds.

D.3.3 Advertising with demand or knowledge spillovers.

Introducing advertisement costs to the aggregate demand and knowledge externalities models from Sections B.2 and B.3 yields the same formula for τ_n as in Equation (D.10), so that we have the limit with high disagreement $\tau_\infty = \sigma/\gamma - 1/\gamma$.

The main difference between wedge in the standard model and the models in Section B.2 and B.3 arises through the differences in the profit function, which affects the ratio $\mathcal{I}_1/\widetilde{\mathcal{I}}_n$. With aggregate demand externalities, the ratio is $\gamma/(\sigma-1)$ and the wedge with agreement is:

$$\tau_1^{\text{AD}} = -\frac{1}{\gamma}. \tag{D.13}$$

For the model with knowledge spillovers the ratio is $\gamma/(1-\alpha)\sigma$ and the wedge with agreement is:

$$\tau_1^{\text{KS}} = 1 - \frac{1}{\gamma} - \frac{\sigma - 1}{(1-\alpha)\sigma}. \tag{D.14}$$

D.4 Participation costs in the baseline model

Another way to ensure the marginal firm makes zero profits is to assume firms invest in infrastructure to produce. In particular, suppose that upon entry all firms can participate on the goods market, but firms must buy one unit of infrastructure to reach all of their customers. Households produce infrastructure competitively at a cost of effort Φ. In an equilibrium with M producing firms, the price of infrastructure is:

$$\Phi'(M) = \varphi(M) = \varphi_0 \cdot M^\nu$$

with $\nu > 0$, so that the cost of infrastructure is increasig in the mass of producing firms M.

D.4.1 Participating firms

Given M_e and M, profits before the infrastructure costs are unchanged from the standard model with decreasing returns to scale:

$$\pi(a) = \frac{1}{\sigma - 1} \cdot w^{1-\sigma} \cdot a^{\sigma}.$$

The equilibrium wage is also unchanged:

$$w = \left(\frac{\gamma}{\gamma - \sigma}\right)^{\frac{1}{\sigma}} \cdot \left(\frac{M_e}{L}\right)^{\frac{1}{\sigma}} \cdot \left(\frac{M_e}{M}\right)^{\frac{1}{\gamma} - \frac{1}{\sigma}}.$$

The marginal firm has productivity $\underline{a}$ and spends all of its profit in infrastructure. Therefore, we have the zero cutoff profit condition $\Phi'(M) = \pi(\underline{a})$, which implies:

$$M^{\nu + \frac{1}{\gamma} + \frac{\sigma-1}{\sigma}} = \frac{1}{\varphi_0} \frac{1}{\sigma - 1} \left(\frac{\gamma - \sigma}{\gamma}\right)^{\frac{\sigma-1}{\sigma}} L^{\frac{\sigma-1}{\sigma}} \cdot M_e^{\frac{1}{\gamma}},$$

where we use the fact that $\underline{a} = (M_e/M)^{1/\gamma}$. In Section D.2, we specified an exogenous set of producing firms $M = M_e^{\chi}/M_0^{\chi-1}$. This arises endogenously through our cost of infrastructure with

$$\chi = \frac{1}{\gamma}\left(\nu + \frac{1}{\gamma} + \frac{\sigma - 1}{\sigma}\right)^{-1}$$

$$M_0^{1-\chi} = \left(\frac{1}{\varphi_0} \frac{1}{\sigma - 1} \left(\frac{\gamma - \sigma}{\gamma}\right)^{\frac{\sigma-1}{\sigma}} L^{\frac{\sigma-1}{\sigma}}\right)^{\left(\nu + \frac{1}{\gamma} + \frac{\sigma-1}{\sigma}\right)^{-1}},$$

where the exponent satisfies $\chi \leq 1$.

We can also compute the elasticity $\mathcal{E}_{\mathcal{C}}$:

$$\mathcal{E}_{\mathcal{C}} = \frac{1}{\gamma} + \chi\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right) = \frac{1}{\gamma} \cdot \frac{1 + \nu}{1 + \nu + 1/\gamma - 1/\sigma} = \chi \cdot (1 + \nu).$$

D.4.2 Equilibrium

The equilibrium condition in the competitive equilibrium is:

$$W'(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1},$$

where we define the modified $\widetilde{\mathcal{I}}_n$ integral to account for the infrastructure expenditures of the firm:

$$\widetilde{\mathcal{I}}_n(M_e, \chi) = \int_{\left(\frac{M_e}{M_0}\right)^{\frac{1-\chi}{\gamma}}}^{\infty} \left(a^{\sigma} - \left(\frac{M_e}{M_0}\right)^{\sigma \frac{1-\chi}{\gamma}}\right) dF_n(a).$$

With $n = 1$, we have:

$$\widetilde{\mathcal{I}}_1(M_e, \chi) = \frac{\sigma}{\gamma - \sigma} \cdot \left(\frac{M_0}{M_e}\right)^{(1-\chi)\frac{\gamma-\sigma}{\gamma}} = \frac{\sigma}{\gamma} \cdot \mathcal{I}_1(M_e, \chi).$$

Aggregate profits therefore represent a fraction σ/γ of aggregate revenue after labor costs, while aggregate infrastructure costs account for the other $(\gamma - \sigma)/\gamma$. Therefore aggregate profits represent a share $1/\gamma$ of consumption and aggregate infrastructure costs $1/\sigma - 1/\gamma$.

D.4.3 Entry wedge

The planner maximizes expected utility:

$$\max_{M_e} \; - W(M_e) + \frac{1}{\sigma}\mathcal{C}\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1} + \frac{\sigma - 1}{\sigma}\mathcal{C} + \underbrace{\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)\mathcal{C}}_{\substack{\text{consumption from} \\ \text{infrastructure}}} - \underbrace{\Phi(M)}_{\substack{\text{cost of} \\ \text{infrastructure}}} \; .$$

The corresponding optimality condition is:

$$\begin{aligned} W'(M_e) = &\frac{1}{\sigma}\mathcal{C}\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\widetilde{\mathcal{I}}'_n}{\widetilde{\mathcal{I}}_n} + \frac{1}{\sigma}\mathcal{C}'\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1} - \frac{1}{\sigma}\mathcal{C}\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\mathcal{I}'_1}{\mathcal{I}_1} \\ &+ \frac{\sigma - 1}{\sigma}\mathcal{C}' + \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)\mathcal{C}' - \underbrace{\Phi'(M)M}_{\left(\frac{1}{\sigma}-\frac{1}{\gamma}\right)\mathcal{C}} \cdot \frac{1}{M}\frac{dM}{dM_e}. \end{aligned}$$

The wedge is therefore:

$$\tau_n = -\mathcal{E}_{\widetilde{\mathcal{I}}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - (\sigma - 1)\mathcal{E}_{\mathcal{C}} \cdot \frac{\mathcal{I}_1}{\widetilde{\mathcal{I}}_n} - \underbrace{\frac{\gamma - \sigma}{\gamma}(\mathcal{E}_{\mathcal{C}} - \mathcal{E}_M) \cdot \frac{\mathcal{I}_1}{\widetilde{\mathcal{I}}_n}}_{\substack{\text{surplus from} \\ \text{participation costs}}},$$

where the last term accounts for the surplus from infrastructure creation.

In particular, the wedge with agreement ($n = 1$) is:

$$\tau_1 = 1 - \mathcal{E}_{\mathcal{C}} - \left[(\sigma - 1)\mathcal{E}_{\mathcal{C}} + \left(1 - \frac{\sigma}{\gamma}\right)(\mathcal{E}_{\mathcal{C}} - \chi)\right]\frac{\gamma}{\sigma}.$$

Using the values of $\mathcal{E}_{\mathcal{C}}$ and χ, we obtain:

$$\tau_1 = 1 - \chi\gamma\left(1 + \nu - \frac{1}{\sigma} + \frac{1}{\gamma}\right) = 0, \tag{D.15}$$

given the formula above for χ. Participation is now a good traded on a competitive market, hence the first welfare theorem applies and the competitive equilibrium is efficient with agreement.

Now we apply the reasoning from Section D.1 to find the condition for convergence when

θ is large. The condition for convergence of $\mathcal{E}_{\widetilde{\mathcal{I}}_n}$ is the same as for $\mathcal{E}_{\mathcal{I}_n}$:

$$\gamma(\theta+\chi) > 1 + \chi\left(\frac{\gamma}{\sigma} - 1\right) \tag{D.16}$$

$$\iff \quad \gamma\theta > 1 + \frac{1}{\nu + \frac{1}{\gamma} + \frac{\sigma-1}{\sigma}}\left(\frac{1}{\sigma} - \frac{1}{\gamma} - 1\right) \tag{D.17}$$

As $n \to \infty$, we have that $\widetilde{\mathcal{I}}_n \to \infty$ and $\mathcal{I}_1 \to 0$, and therefore $\mathcal{I}_1/\widetilde{\mathcal{I}}_n \to 0$.

For the high disagreement wedge we have:

$$\tau_\infty = (\sigma-1)\cdot\mathcal{E}_w = \frac{\sigma-1}{\gamma} + (\sigma-1)\chi\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right) \tag{D.18}$$

$$= \frac{\sigma-1}{\gamma}\cdot\left(\frac{1+\nu}{1+\nu+1/\gamma-1/\sigma}\right). \tag{D.19}$$

Again, Proposition 4 holds. Moreover τ_∞ is decreasing in ν since $1/\gamma < 1/\sigma$. As the cost of producing infrastructure becomes steeper, the sensitivity of firm participation to firm creation is smaller, and the wedge is less responsive. In the limit with $\nu \to \infty$, a fixed number of firms produces, we are back to our baseline $\tau_\infty = (1-\sigma)/\gamma$. Finally, τ_∞ is increasing in σ.

D.5 Melitz (2003) model: participation costs and Dixit-Sitglitz

We now introduce Dixit-Stiglitz preferences to the above model, as in Melitz (2003).

D.5.1 Participating firms

Given M_e and M, profits before the infrastructure costs are unchanged from the standard model with decreasing returns to scale:

$$\pi(a) = \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\cdot\mathcal{C}\cdot w^{1-\sigma}\cdot a^{\sigma-1}.$$

The equilibrium consumption is also unchanged:

$$\frac{\mathcal{C}}{L} = \left(\frac{\gamma}{\gamma-(\sigma-1)}\right)^{\frac{1}{\sigma-1}} M^{\frac{1}{\sigma-1}-\frac{1}{\gamma}} M_e^{\frac{1}{\gamma}}.$$

The marginal firm has productivity $\underline{a}$ and spends all of its profit in infrastructure. Therefore, we have the zero cutoff profit condition $\Phi'(M) = \pi(\underline{a})$, which implies:

$$M^{\nu+\frac{1}{\gamma}+\frac{\sigma-2}{\sigma-1}} = \frac{1}{\varphi_0}\frac{1}{\sigma}\left(\frac{\gamma-(\sigma-1)}{\gamma}\right)^{\frac{\sigma-2}{\sigma-1}} L\cdot M_e^{\frac{1}{\gamma}},$$

where we use the fact that $\underline{a} = (M_e/M)^{1/\gamma}$. In Section D.2, we specified an exogenous set of producing firms $M = M_e^{\chi}/M_0^{\chi-1}$. This arises endogenously through our cost of infrastructure with

$$\chi = \frac{1}{\gamma}\left(\nu + \frac{1}{\gamma} + \frac{\sigma-2}{\sigma-1}\right)^{-1},$$

$$M_0^{1-\chi} = \left(\frac{1}{\varphi_0}\frac{1}{\sigma}\left(\frac{\gamma-(\sigma-1)}{\gamma}\right)^{\frac{\sigma-2}{\sigma-1}} L\right)^{\left(\nu+\frac{1}{\gamma}+\frac{\sigma-2}{\sigma-1}\right)^{-1}},$$

where the exponent satisfies $\chi \leq 1$ if and only if $\nu + \frac{\sigma-2}{\sigma-1} \in (-\infty, -1/\gamma) \cup [0, \infty)$. Otherwise, all firms participate as M_e grows to infinity.

Finally, we derive the elasticity $\mathcal{E}_{\mathcal{C}}$:

$$\mathcal{E}_{\mathcal{C}} = \frac{1}{\gamma} + \chi\left(\frac{1}{\sigma-1} - \frac{1}{\gamma}\right) = \frac{1}{\gamma} \cdot \frac{1+\nu}{1+\nu+1/\gamma-1/(\sigma-1)} = \chi \cdot (1+\nu).$$

D.5.2 Equilibrium

The equilibrium condition in the competitive equilibrium is:

$$W'(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1}.$$

Aggregate profits represent a fraction $(\sigma-1)/\gamma$ of aggregate revenue after labor costs, and aggregate infrastructure costs account for the other $(\gamma-(\sigma-1))/\gamma$. Therefore aggregate profits represent a share $(\sigma-1)/(\sigma\gamma)$ of consumption and aggregate infrastructure costs $(\gamma-(\sigma-1))/(\sigma\gamma)$.

D.5.3 Entry wedge

The planner maximizes expected utility:

$$\max_{M_e} \; - W(M_e) + \frac{1}{\sigma}\mathcal{C}\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1} + \frac{\sigma-1}{\sigma}\mathcal{C} + \left(\frac{\gamma-(\sigma-1)}{\sigma\gamma}\right)\mathcal{C} - \Phi(M).$$

The corresponding optimality condition is

$$\begin{aligned} W'(M_e) = {} & \frac{1}{\sigma}\mathcal{C}\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\widetilde{\mathcal{I}}_n'}{\widetilde{\mathcal{I}}_n} + \frac{1}{\sigma}\mathcal{C}'\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1} - \frac{1}{\sigma}\mathcal{C}\frac{\widetilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\mathcal{I}_1'}{\mathcal{I}_1} \\ & + \frac{\sigma-1}{\sigma}\mathcal{C}' + \left(\frac{\gamma-(\sigma-1)}{\sigma\gamma}\right)\mathcal{C}' - \Phi'(M)M\frac{1}{M}\frac{dM}{dM_e}. \end{aligned}$$

The wedge is therefore

$$\tau_n = -\mathcal{E}_{\widetilde{\mathcal{I}}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - \left[(\sigma-1)\mathcal{E}_{\mathcal{C}} + \left(1 - \frac{\sigma-1}{\gamma}\right)(\mathcal{E}_{\mathcal{C}} - \chi)\right]\frac{\mathcal{I}_1}{\widetilde{\mathcal{I}}_n}. \tag{D.20}$$

In particular, the wedge with agreement ($n = 1$):

$$\tau_1 = 1 - \mathcal{E}_{\mathcal{C}} - \left[(\sigma - 1)\mathcal{E}_{\mathcal{C}} + \left(1 - \frac{\sigma - 1}{\gamma}\right)(\mathcal{E}_{\mathcal{C}} - \chi)\right]\frac{\gamma}{\sigma - 1} \tag{D.21}$$

$$= -\frac{1}{\sigma - 1} \cdot \frac{1 + \nu}{1 + \nu + 1/\gamma - 1/(\sigma - 1)}. \tag{D.22}$$

Now we apply the reasoning from Section D.1 to find the condition for convergence when θ is large. The condition for convergence of $\mathcal{E}_{\widetilde{\mathcal{I}}_n}$ is the same as for $\mathcal{E}_{\mathcal{I}_n}$:

$$\gamma(\theta + \chi) > 1 + \chi\left(\frac{\gamma}{\sigma - 1} - 1\right) \tag{D.23}$$

$$\iff \quad \gamma\theta > 1 + \frac{1}{\nu + \frac{1}{\gamma} + \frac{\sigma-2}{\sigma-1}}\left(\frac{1}{\sigma - 1} - \frac{1}{\gamma} - 1\right) \tag{D.24}$$

As $n \to \infty$, we have $\widetilde{\mathcal{I}}_n \to \infty$ and $\mathcal{I}_1 \to 0$, and therefore $\mathcal{I}_1/\widetilde{\mathcal{I}}_n \to 0$.

For the high disagreement wedge we have:

$$\tau_\infty = (\sigma - 1) \cdot \mathcal{E}_w - \mathcal{E}_{\mathcal{C}} = \frac{\sigma - 2}{\gamma} \cdot \left(\frac{1 + \nu}{1 + \nu + 1/\gamma - 1/(\sigma - 1)}\right). \tag{D.25}$$

When $\nu \to \infty$, there is a fixed supply of infrastructure and thus a fixed number of firms, which implies:

$$\tau_1 = -\frac{1}{\sigma - 1},$$
$$\tau_\infty = \frac{\sigma - 2}{\gamma}.$$

E Data Appendix

E.1 Data construction details

Bubbles. Following Greenwood, Shleifer, and You (2018), we identify bubbles as episodes in which stock prices of an industry have increased over 100% in terms of both raw and net of market returns over the previous two years, followed by a decrease in absolute terms of 40% or more. Industries are classified according to the Fama French 49 industry scheme and the data begin in 1928.

Value of Innovation. We use the stock market value of patents at the patent and at the firm level directly from Kogan et al. (2017), as well as the number of citations that accrue to a patent.[32]

Compustat Segments. We merge the Compustat funda file with the Compustat segments file. We estimate the number of segments with different industry codes. The Compustat segment file provides both a six and a four digit industry code, which gives two measures of the number of different types of industries within a public firm.

Value of Spillovers. We obtain information on the quantity of competition spillovers (variable spillsic) as well as technological spillovers (variable spilltec) from the replication files in Bloom, Schankerman, and Van Reenen (2013). The exposure to spillovers from product market, *spillsic*, is defined as the correlation of the sales across two firms' Compustat segments. If we consider the vector of average sales share across each industry for a given firm i, $\mathbf{S}_i$. Product market proximity between firm i and j is defined by the uncentered correlation: $\text{SIC}_{ij} = \mathbf{S}_i\mathbf{S}_j' / \left(\sqrt{\mathbf{S}_i\mathbf{S}_i'}\sqrt{\mathbf{S}_j\mathbf{S}_j'}\right)$. The product market spillover is the average stock of R&D that are in the product market proximity of firm i:

$$spillsic_i = \sum_{j \neq i} \text{SIC}_{ij} G_j,$$

where G_j is the stock of R&D for firm j. The exposure to knowledge spillovers is constructed the same way, where we define for firm i a vector of share of patents across technology classes from the USPTO as $\mathbf{T}_i$. The uncentered correlation of technology between firm i and j is: $\text{TECH}_{ij} = \mathbf{T}_i\mathbf{T}_j' / \left(\sqrt{\mathbf{T}_i\mathbf{T}_i'}\sqrt{\mathbf{T}_j\mathbf{T}_j'}\right)$. The product market spillover is the average stock of R&D that are in the product market proximity of firm i:

$$spilltech_i = \sum_{j \neq i} \text{TECH}_{ij} G_j,$$

To look at the effect of spillovers we use sales item from Compustat funda file and Tobin's q (market-to-book ratio) from the CRSP-Compustat merged file.

[32]We thank Dimitris Papanikolaou for graciously sharing his data with us.

E.2 Supplementary Tables

Table E.1
Innovation in times of Bubbles

	Patents (#)	Log Patents (#)			
	(1)	(2)	(3)	(4)	(5)
Bubble	1.385**	0.148***	0.154**	0.169***	0.178**
	(0.578)	(0.056)	(0.066)	(0.063)	(0.075)
Lagged Citations	0.982***				
	(0.026)				
Lagged Log Citations		0.824***	0.795***	0.827***	0.799***
		(0.008)	(0.008)	(0.007)	(0.008)
Fixed Effects	C	–	C	Y	C, Y
Observations	106,176	106,278	106,176	106,278	106,176
R^2	0.91	0.67	0.68	0.68	0.68

Note: Table E.1 presents panel regressions of the quantity of innovation, measured by the number of patents issued at the USPTO three digit class level, on a dummy from Greenwood, Shleifer, and You (2018) that captures whether the firm is an industry that is in a bubble or not. We control for the lagged number of patents for column one and lagged logarithm for columns two to five. Depending on the specification, we include fixed effects for the patent class level C and patent grant year Y. Standard errors clustered at the grant year level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table E.2
Private value of innovation in bubbles

	Patent Level			Firm Level		
	(1)	(2)	(3)	(4)	(5)	(6)
Bubble dummy	0.317***	0.289***	0.277***	0.514***	0.429***	0.431***
	(0.094)	(0.090)	(0.083)	(0.114)	(0.123)	(0.080)
Log Market Cap			0.543***			0.625***
(lagged)			(0.026)			(0.020)
Fixed Effects	Y, F	Y, F	Y, F	Y, F	Y, F	Y, F
Interaction	NA	Ind x Cite	Ind x Cite	NA	Ind x Cite	Ind x Cite
Observations	1,171,806	1,118,675	1,116,740	47,886	47,886	47,484
R^2	0.68	0.69	0.74	0.89	0.94	0.96

Note: Table E.2 presents panel regressions of the value of innovation, as measured in Kogan et al. (2017) at the patent firm levels, on a dummy from Greenwood, Shleifer, and You (2018) that captures whether the firm is an industry that is in a bubble or not. We control for the forward looking number of citations generated by a patent (or firm) from Kogan et al. (2017), and the lagged market capitalization of the firm. We include firm fixed effects F and patent grant year fixed effects Y. Depending on the specification, we also use industry fixed effects (from the Fama-French 49 industry classification) interacted with the log number of forward looking citations to allow for different slopes in the relation between private valuation and the patent quality across industries. Standard errors clustered at the grant year level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table E.3
Diversity and private value of innovation in bubbles

	Patent Level			Firm Level		
	(1)	(2)	(3)	(4)	(5)	(6)
Bubble x Segments	-0.562^{***}	-0.550^{***}	-0.488^{***}	-0.388^{**}	-0.370^{***}	-0.295^{***}
(NAICS 6 digits)	(0.178)	(0.175)	(0.164)	(0.155)	(0.094)	(0.072)
Bubble	1.456^{***}	1.425^{***}	1.329^{***}	1.459^{***}	1.389^{***}	1.167^{***}
	(0.232)	(0.227)	(0.235)	(0.317)	(0.342)	(0.289)
Segments	0.122	0.122	0.112	0.010	-0.031	-0.026
(NAICS 6 digits)	(0.096)	(0.095)	(0.101)	(0.044)	(0.048)	(0.041)
Log Citations		0.049^{***}	0.047^{***}		0.047^{***}	0.047^{***}
(forward looking)		(0.010)	(0.009)		(0.009)	(0.009)
Log Market Cap			0.156^{***}			0.287^{***}
(lagged)			(0.042)			(0.046)
Fixed Effects	Y, F	Y, F	Y, F	Y, F	Y, F	Y, F
Observations	180,636	180,636	177,911	10,426	10,426	10,256
R^2	0.71	0.71	0.72	0.88	0.93	0.94

Note: Table E.3 presents panel regressions of the value of innovation, as measured in Kogan et al. (2017) at the patent and firm levels, on a bubble dummy interacted with the number of industries spanned by the different segments of the parent firm. The bubble dummy is from Greenwood, Shleifer, and You (2018) and captures whether the firm is an industry that is in a bubble or not. Compustat segments are measured at the six digit NAICS code level from the Compustat segment file. We control for the forward looking number of citations generated by a patent (or firm) from Kogan et al. (2017) and the lagged market capitalization of the firm. We also include fixed effects for firm F and patent grant year Y. Standard errors clustered at the grant year level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table E.4
Forward Citations and Patent Market Values

	Table II from Kogan et al. (2017)					Firm-Year Fixed Effects		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log citations	0.174***	0.099***	0.054***	0.013***	0.004***	0.023***	0.019***	0.012***
	(0.017)	(0.010)	(0.005)	(0.001)	(0.001)	(0.004)	(0.003)	(0.002)
Controls								
Firm Size	—	✓	✓	✓	✓	—	✓	✓
Volatility	—	—	✓	✓	—	—	—	✓
Fixed Effects	CxY	CxY	CxY	CxY, F	CxY, FxY	Y, F	Y, F	Y, F
Observations	1,801,301	1,801,301	1,801,301	1,801,301	1,801,301	1,801,301	1,801,301	1,801,301
R^2	0.20	0.71	0.79	0.93	0.95	0.82	0.91	0.92

Note: Table E.4 presents panel regressions of the stock market value of innovation on the logarithm of the number of citations received by a patent until the end of the sample in 2010. Depending on the specification, we control for the logarithm of firm size and volatility. Depending on the specification, we include patent grant times year fixed effects, CxY, year fixed effects F, and firm fixed effects Y. Columns (1) to (5) reproduce Table II from Kogan et al. (2017), while columns (6) to (8) only include Year and Firm fixed effects to be comparable with the Tables above. Standard errors clustered at the grant year level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.