

Intangible Investment and Market Concentration

Joshua Weiss*

New York University

jmw676@nyu.edu

November 11, 2019

[Click here](#) for the most recent version

Abstract

I propose a new theory to explain the recent rise in industry-level concentration and markups. I study a dynamic model of oligopolistic competition in which firms make a one-time, irreversible investment in intangible capital at entry. This effectively allows firms to commit to a higher level of production, which deters competitors and potential entrants. I take as given a change in technology that increases the importance of intangible capital in production. Large productive firms with high markups disproportionately increase investment and gain market share. In the calibrated model, a shift toward intangible capital in line with estimates of the recent rise in intangibles can explain more than half the rise in concentration and about a third of the rise in markups from 1997 to 2012. The model also quantitatively matches the observed increase in labor productivity and fall in the labor share in concentrating industries. Finally, I show that in the model, these changes result in an *increase* in welfare equivalent to a 0.3% permanent increase in consumption. High markups are inefficient in the model because they imply that large firms are *under*producing. Intangible capital allows large firms to commit to a higher level of production, undoing part of this inefficiency.

*I am indebted to Virgiliu Midrigan and Venky Venkateswaran for their continuous support and guidance on this project. I am grateful to Paula Onuchic, Dmitry Sorokin, and William Gamber for many long conversations. I would also like to thank Mark Gertler, Thomas Sargent, Ricardo Lagos, Guido Menzio, and many other NYU seminar participants for their advice and comments. All errors are my own.

1 Introduction

Many authors have documented two concurrent trends since the 1980s in the US: a rising share of sales going to the top firms in industries at the national level and increasing markups of price over marginal cost.¹ Related to the second trend is a falling labor share. A third trend that is less often associated with the first two is the rise in intangible capital.² Intangible capital captures the various investments firms can make that do not form traditional, measurable capital. Examples are research and development that affect what we might otherwise consider to be the productivity of a firm or a product; advertising and marketing that build brand value or raise awareness of a product; or investments in firm culture and organization capital. An important feature of intangible capital, which is the focus in this paper, is that it is difficult to undo or to sell. Intangible capital becomes a part of the underlying productivity of a firm or product.

I propose a new theory that links the rise in irreversible, firm-specific intangible investments to rising concentration and markups. Firms use intangible investments to commit to a higher level of production and so deter their competitors. As intangible capital becomes more important for production, large firms disproportionately take advantage and their market shares grow, generating a rise in concentration. Since large firms tend to set higher markups, this shifts sales toward higher markup firms and raises the average markup. Moreover, as these firms' sales increase, their markups rise as well, further increasing the average markup. The key premises of the theory are that irreversible, firm-specific capital became more important for production and that firms with larger market shares set higher markups.

I formalize the theory in a dynamic oligopolistic model with endogenous entry and exogenous exit. The model is a dynamic version of the oligopolistic models studied in Atkeson and Burstein (2008)

¹For empirical results on rising concentration and the falling labor share, see Autor, Dorn, Katz, Patterson, and Van Reenen (2019) and Barkai (2017). For results on increasing markups, see De Loecker, Eeckhout, and Unger (2018) and Hall (2018).

²See Crouzet and Eberly (2019), Corrado, Haskel, Jona-Lasinio, and Iommi (2012), and Ewens, Peters, and Wang (2019).

and Edmond, Midrigan, and Xu (2015) with an irreversible investment. There is a continuum of industries and finitely many firms in each industry. Firms have a fixed factor of production, intangible capital, that they choose once and for all at entry. They also have flexible factors of production, physical capital and labor, that they choose at each moment when they choose production in Cournot competition.

More productive firms depend more on intangible capital, have larger market shares, and set higher markups. When a larger firm expands, it must take a higher percentage of its competitors' revenues since they have less overall. This has two implications. First, when a firm increases its flexible factors of production, the larger it is, the more it must cut its price. As such, productive firms that tend to have large market shares set high markups. Second, when a firm increases its intangible capital, the larger it is, the more it hurts its competitors' profits. This reduction in competitor profits creates deterrence through two channels. Other incumbents anticipate lower prices, so they choose lower levels of flexible inputs and sell less. Future potential entrants anticipate lower expected profits, so they are less likely to pay their entry costs and, if they do enter, purchase less intangible capital. As a result, productive firms that anticipate large market shares depend more on intangible capital than on their flexible inputs so they can take advantage of its ability to deter.

I shock the importance of intangibles in production and track the transition of the economy to a new steady state. Following the shock, productive firms disproportionately increase their intangible investment due to their deterrence motive. Effective productivity dispersion rises and production shifts toward larger firms with higher markups. The market shares of the top firms in industries increase and the average markup rises. In the calibrated model, a shift toward intangible capital in line with estimates of the recent rise in intangibles can explain more than half the rise in the market shares of the top 8 firms in 6-digit NAICS US industries and about a third of the rise in the revenue-weighted average markup of the largest firms from 1997 to 2012. The economy converges to the new steady state relatively quickly; after 15 years, measures of concentration and markups are two-thirds of the way toward their new steady state values.

The model also quantitatively matches other features of the rise in concentration at the industry

level. I run the same industry-level regressions as Barkai (2017) and Ganapati (2019) of changes in the labor share and labor productivity on changes in the market share of the top firms. I find similar quantitative results. A one percentage point rise in the market share of the top 8 firms in an industry is associated with about a one percentage point fall in the labor share. The correlation between log market share changes and log labor productivity growth is about 0.2. More generally, the rise in concentration is associated with higher labor productivity, less use of flexible inputs (physical capital and labor), lower prices, and higher output.

This paper adds to a growing literature that seeks to document and explain the recent rise in concentration and related trends, such as rising markups, a falling labor share, a falling physical capital share, rising profits, a falling entry rate, and rising productivity dispersion. Relative to other quantitative work in this area, I add a model with strategic behavior, which introduces the main mechanism in this paper. Relative to other work that considers strategic behavior, I develop a quantitative macroeconomic model that can better fit the data, in particular industry concentration data that is important for quantifying strategic effects.

Perhaps the theory most related to the one in this paper is in Gouin-Bonenfant (2018), who studies the effects of an exogenous rise in productivity dispersion in a model with atomistic firms and a labor share that is falling in firm productivity. I show that if the ability to invest in productivity rises, firms with larger market shares will make more of these investments and productivity dispersion will increase. In this sense, I endogenize the rise in productivity dispersion. This allows us to better understand the welfare consequences of recent trends as well as the effects of policies. For example, a policy targeted toward reducing concentration, such as a size-dependent tax, may be more harmful once we consider the effects on the productivity distribution. Moreover, my theory avoids the issue that higher productivity dispersion should stimulate entry, contrary to the falling entry rate in the data. In my model, firms with high productivity are not just lucky but also have to pay for large intangible investments.

Another related paper is Aghion, Bergeaud, Boppart, Klenow, and Li (2019). They argue that technological changes such as the spread of information technology, perhaps related to the rise in

intangibles, have made it easier for firms to expand into multiple markets. In their model, more productive firms tend to set higher markups. When it becomes cheaper to enter new markets, high productivity firms spread. Concentration and the average markup rise. Yet, in their model, there are infinitely many firms and high productivity firms do not contemplate deterrence when they expand. This adds a new mechanism through which intangible capital increases concentration and markups and is important for welfare considerations.

Other theories generate rising markups by varying the elasticity of substitution across goods, such as Eggertsson, Robbins, and Wold (2018) and Jones and Philippon (2016). However, this cannot jointly match the changes in concentration and markups. If the elasticity goes down, then markups rise at the firm level but concentration falls. Moreover, consumers' love of variety increases, which encourages entry. If the elasticity goes up, then markups fall at the firm level and concentration increases. The average markup may actually rise due to a fall in entry and a shift in sales toward higher markup firms. However, this does not match the relative increases in concentration and markups in the data because it requires such a large rise in concentration to counteract the falling markups at the firm level.

Finally, theories such as those in Akcigit and Ates (2019) and Liu, Mian, and Sufi (2019) are promising, but it is not so clear how they map to the data in their current form. In particular, they rely on strategic interactions in industries with fixed revenue and only two firms. In the data, even at the 6-digit NAICS industry level, the average sales share of the top 8 firms in 2012 is 46%.

I also explore the welfare implications of the rise in concentration and analyze welfare in the model more generally. I find that following the shock and taking into account the transition path, welfare increases by the equivalent of a 0.3% permanent rise in consumption. I then solve a constrained planner's problem to help us understand the welfare consequences of the shock as well as provide a guide to optimal policies. The main intuition for the rise in welfare is as follows. Large high markup firms, like monopolists, set too high prices and sell too little. Although they produce more following the shock to deter competition, the reallocation of production toward high productivity firms brings the economy closer to the optimal distribution of productive resources.

1.1 Intangible Capital

This paper builds on a growing empirical literature that attempts to measure intangible capital.³ Perhaps the most comprehensive is Ewens, Peters, and Wang (2019), who use identifiable intangible assets and goodwill from public firm acquisitions as a measure of intangible capital and estimate how best to capitalize acquired firms' past research and development (R&D) and selling, general, and administrative (SG&A) expenses to explain intangible capital. This includes estimating what fraction of SG&A expenses are investment. They use their estimates to generate intangible capital stocks for all public firms. They find that from 1980 to 2016, intangible capital as a fraction of the total capital stock increased from 37% to 60%. I use these measures to calibrate my model and my main exercise. As further evidence that intangible capital is a factor of production like physical capital, they show that including intangible capital improves the ability of Tobin's q to predict investment. Eisfeldt and Papanikolaou (2013) show that firms with high levels of past SG&A spending tend to have higher output controlling for physical capital and labor.

Ewens, Peters, and Wang (2019) also provide insight into the nature of the component of intangible capital derived from SG&A, which they call organization capital. Firms with high levels of organization capital tend to report high levels of risk due to the potential loss of personnel or key talent and tend to have high brand rankings. This suggests that the marketing and administrative expenses in SG&A build intangible capital through brand value, human capital, and firm culture.

Beyond the similar aggregate time trends, there is evidence of a link between the rise in intangible capital and the increase in market concentration and markups. Crouzet and Eberly (2019) use similar measures of intangible capital as in Ewens, Peters, and Wang (2019) to show that industries with a higher intangible intensity – more intangible capital relative to physical capital – have higher HHIs, i.e. they are more concentrated, and higher markups. Firms with higher intangible intensities have larger industry market shares and set higher markups, both in the cross-section and over time. Along the same lines, Bessen (2017) shows that industries with more use of

³See Eisfeldt and Papanikolaou (2013), Crouzet and Eberly (2019), Peters and Taylor (2017), Corrado, Haskel, Jona-Lasinio, and Iommi (2012), and Ewens, Peters, and Wang (2019).

proprietary information technology systems tend to be more concentrated and tend to have growing concentration.

There is reason to believe that Ewens, Peters, and Wang (2019) understate the value of intangible capital. It is well documented that past sales build a form of intangible capital either through learning by doing or consumer habit.⁴ Furthermore, Bornstein (2018) argues that due to an aging population, consumer habit has become a stronger force over the same time period as the rise in concentration and markups. Of course, these types of investments affect markup decisions and complicate the analysis, so they are beyond the scope of this paper.

Finally, the theory in this paper builds on an empirical and theoretical literature on endogenous sunk costs.⁵ This literature is focused on the distinctions between exogenous and endogenous sunk costs and between sunk and fixed costs. The important features of an endogenous sunk cost are that it is a choice variable for a firm and that it is irreversible. Only then does it become a strategic tool for commitment to production and deterrence. Although to some extent it is known from a comparison of standard Cournot and Stackelberg competition, I have not seen discussion in this literature of the role that endogenous sunk costs play in competition even without entry deterrence. A large sunk cost investment also deters other incumbents through their flexible cost choices. In my calibrated model, this turns out to be the dominant channel through which intangible capital is important.

1.2 Market Share and Markups

An important outcome of the model is that more productive firms have larger market shares and set higher markups. This delivers the result that as large firms become even larger, markups rise; sales shift toward high markup firms and those high markup firms set even higher markups. There

⁴See Bronnenberg, Dubé, and Gentzkow (2012), Bornstein (2018), Foster, Haltiwanger, and Syverson (2016).

⁵See Sutton (1991), Spulber (1981), Bronnenberg, Dhar, and Dubé (2011).

is evidence of a relationship between market shares and markups.⁶ In particular, Amiti, Itskhoki, and Konings (2019) show that small firms' markups don't change with their productivity and don't respond to other firms' prices, while large firms' markups increase with productivity and with other firms' prices. This confirms the mechanism in the model for the relationship between market share and markups. As a firm's productivity and market share grow, it increasingly does not pass through the lower cost into a lower price, instead letting its markup rise. Larger firms' markups are then more responsive to competitors' prices because their markups change more with market share.

The remainder of the paper is organized as follows. In section 2, I describe the model. In section 3, I describe the solution of the model. In section 4, I calibrate the model. In section 5, I discuss the numerical techniques I use to compute the solution to the model. In section 6, I show the main results. In section 7, I discuss welfare. In section 8, I conclude.

2 Model

Time is continuous and infinite. There is a representative household that consumes goods and supplies labor. There are a continuum of industries indexed by $j \in [0, 1]$. In each industry, there are finitely many, $N_{j,t}$, firms that produce differentiated goods and compete in quantities. Firms exit exogenously. Potential entrants arrive in each industry, choose whether to pay a fixed cost and enter, and make an irreversible investment. There is a representative capital producer.

⁶See Atkeson and Burstein (2008), Edmond, Midrigan, and Xu (2019), Amiti, Itskhoki, and Konings (2019).

2.1 Household Preferences and Budget Constraint

The representative household maximizes the expected discounted value of its flow utility. Its discount rate is $\rho > 0$. Flow utility at time t is

$$\ln(C_t) - L_t,$$

where C_t is final good consumption and L_t is labor. Final good consumption is Cobb-Douglas in industry consumption:

$$\ln(C_t) = \int_0^1 \ln(C_{j,t}) dj,$$

where industry consumption is a CES aggregate of differentiated good consumption within the industry:

$$C_{j,t} = \left[\sum_{i=1}^{N_{j,t}} c_{i,j,t}^{\frac{\gamma-1}{\gamma}} \right]^{\frac{\gamma}{\gamma-1}},$$

where $\gamma > 1$ is the elasticity of substitution between goods within industries.

At each time t , the household takes the prices of the differentiated goods within industries, $p_{i,j,t}$, as given. The household sells labor in a perfectly competitive spot market at a wage normalized to 1. The household can buy and sell physical capital in a competitive spot market at price $p_{K,t}$. The household owns physical capital K_t , which depreciates at rate η and which the household rents out to firms at rental rate $r_{K,t}$. The household owns all current and future firms and so receives net flow firm profits Π_t . The household has access to a risk-free bond in zero net supply that pays net interest rate r_t . As such, if the household's total wealth at time t is W_t , then it evolves according to

$$\dot{W}_t = L_t + \Pi_t + (r_{K,t} + \dot{p}_{K,t} - p_{K,t}\eta)K_t - \int_0^1 \left[\sum_{i=1}^{N_{j,t}} p_{i,j,t} c_{i,j,t} \right] dj + r_t(W_t - p_{K,t}K_t).$$

2.2 Capital Producers

There is a representative capital producer that hires labor and produces capital to maximize the expected discounted value of flow profits with the household stochastic discount factor. The production function is linear in labor with productivity Z_k . At each time t , a capital producer can hire labor in the competitive labor market to produce capital, which it sells in the competitive capital market at price $p_{K,t}$.

2.3 Industry Structure and Incumbents

In each industry, there are $N_{j,t}$ firms, indexed by $i = 1, 2, \dots, N_{j,t}$. Each firm's production is Cobb-Douglas in three factors, intangible capital, physical capital, and labor:

$$q_{i,j,t} = z_{i,j} s_{i,j}^{\alpha_S} k_{i,j,t}^{\alpha_K} l_{i,j,t}^{1-\alpha_S-\alpha_K},$$

where $z_{i,j}$ is the firm's productivity. A firm's productivity and intangible capital are fixed for the duration of its existence and physical capital and labor are rented at each moment t in competitive spot markets at prices $r_{K,t}$ and 1, respectively. As such, a firm is characterized by its effective productivity,

$$e_{i,j} = z_{i,j} s_{i,j}^{\alpha_S}.$$

A firm exits exogenously at Poisson arrival rate η . Firms maximize their expected present discounted value of flow profits with the household stochastic discount factor. When firms exit, they receive value 0.

2.4 Entry

In each industry, a potential entrant arrives at Poisson arrival rate λ . The potential entrant draws an entry cost in units of labor from distribution G with pdf g . If the firm does not pay the entry

cost, it does not enter and receives value 0. If the firm pays the entry cost, it draws its productivity from distribution F with pdf f . After observing its productivity, the firm purchases intangible capital in the competitive capital market at price $p_{K,t}$. The firm then enters its industry with its effective productivity.

2.5 Physical vs. Intangible Capital

There are two types of capital, physical and intangible. Physical capital is rented at rate $r_{K,t}$ and intangible capital is purchased at price $p_{K,t}$. A firm can freely choose its physical capital in each moment. When a firm purchases intangible capital upon entry, the capital can never be separated from the firm. When the firm exits, the intangible capital is lost forever. Physical and intangible capital are both irreversible in that neither can be converted into goods. The key feature of intangible capital relative to physical capital is that it is firm-specific so it cannot be resold. Since each type of capital is produced by capital producers at the same cost and since the depreciation rate for physical capital is the same as the firm exit rate, this firm-specificity is the only difference between physical and intangible capital.

It will be more clear later, but for now note that in steady state, we can think of all intangible capital as purchased directly from capital producers when it is produced. We need not think of intangible capital as converted from physical capital.

2.6 Equilibrium

2.6.1 Optimization

The equilibrium concept is a Markov Perfect Equilibrium. The actions of the agents in the economy can only depend on the distribution of firm effective productivity, $\{\Gamma_{j,t}\}_{j \in [0,1]}$, the aggregate

physical capital stock, K_t , and the exogenous parameters. A firm's actions can depend on its industry identity and its own effective productivity.

The household and the capital producer take as given the time paths of goods prices, the interest rate, the labor market and capital market prices, and the rental rate. The household chooses goods consumption, labor supply, net capital purchases, and capital rentals to maximize its expected present discounted flow utility subject to its budget constraint and a transversality condition. The capital producer chooses labor demand and capital sales to maximize its expected present discounted flow profits subject to its production function.

A goods producer or potential entrant takes as given the time paths of the interest rate, the labor market and capital market prices, the rental rate, goods prices outside its industry, the *policy functions* of other firms and potential entrants in its industry, and its demand function. A goods producer chooses labor, physical capital, and sales to maximize its expected present discounted flow profits subject to its production function and demand function. A potential entrant chooses whether to pay its entry cost and, conditional on doing so and its productivity draw, intangible capital to maximize its expected present discounted flow profits.

2.6.2 Market Clearing

The following market clearing conditions must be satisfied at each moment in time.

Labor Market Clearing:

$$L_t = \int_0^1 \left[\sum_{i=1}^{N_{j,t}} l_{i,j,t} + \lambda \int_0^{G^{-1}(\lambda_{j,t}^*)} c_E g(c_E) dc_E \right] dj + L_{K,t},$$

where $\lambda_{j,t}^*$ is the arrival rate of entrants who pay the entry cost in industry j and $L_{K,t}$ is labor demand from capital producers. The left-hand side is labor supply from households. The first term in brackets on the right-hand side is labor demand from goods producers for production. The

second term in brackets is labor demand from entrants for entry costs. The final term is labor demand from capital producers.

Capital Market Clearing:

$$I_t = \dot{K}_t + \eta K_t + \int_0^1 \left[\lambda_{j,t}^* \int_0^\infty s_{j,t}(z) f(z) dz \right] dj,$$

where I_t is capital production or aggregate investment and $s_{j,t}(z)$ is intangible investment of an entrant in industry j with productivity draw z .⁷ The left-hand side is capital market sales by the capital producer. The first two terms on the right-hand side are net capital market purchases by the household. The final term is intangible capital purchases by entrants.

Goods Market Clearing:

For all $j \in [0, 1]$, $i \in \{1, 2, \dots, N_{j,t}\}$,

$$c_{i,j,t} = q_{i,j,t}.$$

The left-hand side is demand for the good produced by firm i in industry j . The right-hand side is supply.

Rental Market Clearing:

$$K_t = \int_0^1 \left[\sum_{i=1}^{N_{j,t}} k_{i,j,t} \right] dj.$$

The left-hand side is capital rented out by households. The right-hand side is capital rented by goods producers.

⁷Aggregate intangible investment is usually a flow. Following the main shock considered in this paper, there will be a mass of intangible capital purchases. At this moment, if $K_{t,+}$ is capital immediately following the shock and \bar{S} is the mass of intangible purchases immediately following the shock, then

$$K_{t,+} = K_t - \bar{S}.$$

Going forward, the usual equation for \dot{K} holds with $K_{t,+}$ replacing K_t .

2.6.3 Evolution of Distributions

In each industry j at time t , the set of effective productivities is $\Gamma_{j,t}$, which has $N_{j,t}$ elements. At rate $\eta N_{j,t}$, $\Gamma_{j,t}$ transitions to $\Gamma_{j,t} \setminus e_{i,j}$, where each i is drawn with equal probability. At rate $\lambda_{j,t}^*$, $\Gamma_{j,t}$ transitions to $\{\Gamma_{j,t}, z s_{j,t}(z)^{\alpha_S}\}$, where z is drawn from distribution function F .

3 Solving the Model

3.1 Household Optimization

If the household's time t value is $H_t(W_t)$, then its Hamilton-Jacobi-Bellman equation is

$$\rho H_t(W_t) = \max_{\{c_{i,j,t}\}, L_t, K_t, \dot{W}_t} \{\ln(C_t) - L_t + \dot{W}_t H'_t(W_t)\} + \dot{H}_t(W_t),$$

subject to

$$\ln(C_t) = \int_0^1 \ln \left(\left[\sum_{i=1}^{N_{j,t}} c_{i,j,t}^{\frac{\gamma}{\gamma-1}} \right]^{\frac{\gamma-1}{\gamma}} \right) dj,$$

$$\dot{W}_t = L_t + \Pi_t + (r_{K,t} + \dot{p}_{K,t} - p_{K,t}\eta)K_t - \int_0^1 \left[\sum_{i=1}^{N_{j,t}} p_{i,j,t} c_{i,j,t} \right] dj + r_t(W_t - p_{K,t}K_t),$$

and non-negativity constraints on L_t and K_t .

3.1.1 Intratemporal Optimization

The first order condition for labor supply, L_t , yields

$$H'_t(W_t, K_t) = 1$$

if labor supply is positive. The first order condition for $c_{i,j,t}$ yields the demand function

$$c_{i,j,t} = p_{i,j,t}^{-\gamma} \left[\sum_{i=1}^{N_{j,t}} p_{i,j,t}^{1-\gamma} \right]^{-1}.$$

A little algebra shows that total expenditure on all final goods is fixed at 1. This follows from the linear disutility of labor and setting the wage to be 1. Since the CES function is constant returns to scale, there is an aggregate price index, p_t , so that final good expenditure is $p_t c_t$. The marginal cost of supplying labor is 1 and the marginal benefit is the real wage, $1/p_t$, multiplied by the marginal utility of consumption, $1/c_t$. Cobb-Douglas preferences across industries implies that total expenditure on each industry is fixed at 1.

3.1.2 Intertemporal Optimization

If labor supply is always positive, then bond market clearing implies that

$$r_t = \rho.$$

The marginal utility of wealth is fixed over time by the linear disutility of labor and defining labor to be the numeraire good. As such, the only determinant of the interest rate is the household's time discount rate. Note that this holds even outside of steady state.

The first order condition for K_t yields

$$r_{K,t} + \dot{p}_{K,t} - p_{K,t}\eta = p_{K,t}r_t$$

if physical capital holdings are positive. The household must be indifferent between holding a unit of capital, which earns rent and capital gains, and holding a unit of the bond, which earns the interest rate.

3.2 Capital Producer Optimization

The capital producer's decision is purely static. At each time t , it chooses capital production to maximize profits:

$$\max_{I_t, L_{K,t}} \{p_{K,t}I_t - L_{K,t}\}$$

subject to

$$I_t = Z_K L_{K,t}$$

and a non-negativity constraint on I_t . The first order conditions yield

$$p_{K,t} = 1/Z_K$$

if investment is positive and finite. Along with the results from household optimization, this implies that the rental rate is fixed over time:

$$r_{K,t} = \frac{\rho + \eta}{Z_K}.$$

Going forward, I call the price of capital p_K and the rental rate r_K .

3.3 Incumbent Optimization

Incumbent i in industry j maximizes the expected discounted value of its flow profits with the household stochastic discount factor. If the distribution of other firms' effective productivities in industry j is $\Gamma_{j,-i,t}$ and the incumbent's value at time t is $V_t(e, \Gamma_{j,-i,t})$, then its Hamilton-Jacobi-

Bellman equation is

$$\begin{aligned}
\rho V_t(e, \Gamma_{j,-i,t}) &= \max_{q,l,k} \{p_t(q, e, \Gamma_{j,-i,t})q - l - r_K k\} \\
&+ \eta \sum_{i' \neq i}^{N_{j,t}} [V_t(e, \{\Gamma_{j,-i,t} \setminus e_{i',j}\}) - V_t(e, \Gamma_{j,-i,t})] - \eta V_t(e, \Gamma_{j,-i,t}) \\
&+ \lambda_t^*(\Gamma_{j,t}) \left[\int_0^\infty V_t(e, \{\Gamma_{j,-i,t}, e_t^*(\Gamma_{j,t}, z)\}) f(z) dz - V_t(e, \Gamma_{j,-i,t}) \right] \\
&+ \dot{V}_t(e, \Gamma_{j,-i,t}),
\end{aligned}$$

subject to

$$q = ek^{\alpha_k} l^{1-\alpha_S-\alpha_k},$$

where $\lambda_t^*(\Gamma_{j,t})$ is the arrival rate of entrants that pay their entry costs and $e_t^*(\Gamma_{j,t}, z)$ is the effective productivity of an entrant who draws productivity z .

The first term on the right-hand side is flow profits. The second line reflects the change to the incumbent's value when a firm exits – the final term is the change in value the incumbent receives when it exits. The third line reflects the change in the incumbent's value when a new firm arrives and pays the entry cost. Conditional on a new firm arriving and paying the entry cost, the incumbent takes the expected value over the possible productivity draws for the new firm. The final line is the change in the value function over time.

An incumbent's effective productivity is fixed, so it takes as given the industry effective productivity distribution, $\Gamma_{j,t}$. The only decision an incumbent makes is a static one: choosing physical capital, labor, and production to maximize profits in each moment.

3.3.1 Cost Minimization

The first order conditions for physical capital and labor show that the incumbent always combines them in the same ratio. As such, it is as if the incumbent has a single factor of production, a

variable input v , and produces according to

$$q = ev^{1-\alpha_S},$$

where v is units of the optimal variable cost bundle that consists of

$$\left(\frac{\alpha_K}{(1 - \alpha_S - \alpha_K)r_K} \right)^{\frac{-\alpha_K}{1-\alpha_S}}$$

units of labor and

$$\left(\frac{\alpha_K}{(1 - \alpha_S - \alpha_K)r_K} \right)^{\frac{1-\alpha_S-\alpha_K}{1-\alpha_S}}$$

units of physical capital. The variable input has price

$$p_V = \frac{1 - \alpha_S}{1 - \alpha_S - \alpha_K} \left(\frac{\alpha_K}{(1 - \alpha_S - \alpha_K)r_K} \right)^{\frac{-\alpha_K}{1-\alpha_S}}.$$

3.3.2 Quantity Choice and Optimal Markup

When choosing its level of production, the incumbent takes other firms' quantity choices as given. With some algebra, we can rewrite the consumer's demand to get

$$p_t(q, e, \Gamma_{j,-i,t}) = q^{\frac{-1}{\gamma}} \left[\sum_{i' \neq i}^{N_{j,t}} c_{i',j,t}^{\frac{\gamma-1}{\gamma}} + q^{\frac{\gamma-1}{\gamma}} \right]^{-1}.$$

The elasticity of revenue with respect to quantity is

$$\epsilon_t(q, e, \Gamma_{j,-i,t}) = \frac{\gamma - 1}{\gamma} (1 - s_t(q, e, \Gamma_{j,-i,t})),$$

where $s_t(q, e, \Gamma_{j,-i,t})$ is the incumbent's share of revenue in the industry, which is equal to the incumbent's revenue since total industry revenue is 1. The revenue elasticity is the product of two components. The first component $(\gamma - 1)/\gamma$ is the revenue elasticity in the standard CES model with atomistic firms. As in the standard CES model, as the incumbent's price falls, the within industry elasticity, γ , controls how quickly households substitute expenditure away from

other firms in the industry and toward the incumbent. As such, as the incumbent produces more, γ controls by how much the incumbent's price must fall and so by how much the incumbent's revenue rises – a higher elasticity γ means the price need not fall by as much and so revenue rises by more. Yet, the only revenue available for the incumbent to gain is the revenue held by other firms in the industry, which is falling as the incumbent's revenue increases. Hence, unlike in the standard CES model, as the incumbent's revenue grows, larger quantity increases are required to generate the same increase in revenue.

The revenue elasticity implies that the incumbent's optimal quantity is characterized by the optimal gross markup of price over marginal cost:

$$\frac{p_t(q, e, \Gamma_{j,-i,t})}{p_{V,t}} (1 - \alpha_S) e^{\frac{1}{1-\alpha_S}} q^{\frac{-\alpha_S}{1-\alpha_S}} = \frac{1}{\epsilon_t(q, e, \Gamma_{j,-i,t}) + 1} = \frac{\gamma}{\gamma - 1} \frac{1}{1 - s_t(q, e, \Gamma_{j,-i,t})}.$$

Although this equation only implicitly characterizes optimal quantity, we can see that optimal quantity exists and is uniquely defined and that both optimal quantity and the optimal markup are increasing in the incumbent's effective productivity. The left-hand side is the actual markup at a given level of production and the right-hand side is the “desired” markup that satisfies the firm's first order condition given the revenue elasticity at that level of production. As production rises, the incumbent's implied price falls and marginal cost rises, so its actual markup falls. The incumbent's market share rises, so the desired markup increases. Hence, the solution must be unique. A solution always exists because as production goes from 0 to infinity, the actual markup goes from infinity to 0 and the desired markup goes from $\gamma/(\gamma - 1)$ to infinity. Finally, holding quantity fixed as effective productivity increases, the only effect on the markup equation is that the actual markup rises because marginal cost falls. To bring the actual markup back in line with the desired markup, quantity must increase.⁸ As quantity rises, the desired markup rises. Hence, following the increase in effective productivity, market share and the markup are higher.

It is useful to note that the desired markup is convex in the sales share. It follows that a larger

⁸This also implies that other firms will set lower quantities in response to the higher effective productivity. Then, the demand curve shifts up, which raises the actual markup for any given quantity. This implies that quantity and the markup must rise further. More on this in the next subsection.

firm's markup is more responsive to changes in its market share.

Call the incumbent's optimal quantity and variable cost choices $q_t^*(e, \Gamma_{j,-i,t})$ and $v_t^*(e, \Gamma_{j,-i,t})$, respectively.

3.4 Entrant Optimization

A potential entrant makes two decisions. First, whether to pay the entry cost. Second, conditional on paying the entry cost and based on the subsequent productivity draw, the entrant must choose its intangible capital and so its effective productivity. The entrant's maximization problem is

$$\max_{s,e} \{V_t(e, \Gamma_{j,t}) - p_K s\}$$

subject to

$$e = z s^{\alpha_S}$$

and a non-negativity constraint on s . The solution to this problem is $s_t^*(\Gamma_{j,t}, z)$ and $e_t^*(\Gamma_{j,t}, z)$. The value of being an entrant before drawing productivity is

$$E_t(\Gamma_{j,t}) = \int_0^\infty [V_t(e_t^*(\Gamma_{j,t}, z), \Gamma_{j,t}) - p_K s_t^*(\Gamma_{j,t}, z)] f(z) dz.$$

A potential entrant pays the entry cost if and only if it is less than $E_t(\Gamma_{j,t})$. It follows that the entry rate is

$$\lambda_t^*(\Gamma_{j,t}) = \lambda G(E_t(\Gamma_{j,t})).$$

3.4.1 Incentives to Invest in Intangible Capital

The first order conditions for intangible capital and effective productivity yield

$$\alpha_S e_t^*(\Gamma_{j,t}, z)^{\frac{\alpha_S - 1}{\alpha_S}} z^{\frac{1}{\alpha_S}} V_{t,1}(e_t^*(\Gamma_{j,t}, z), \Gamma_{j,t}) = p_K,$$

where $V_{t,m}$ is the partial derivative of V_t with respect to its m^{th} argument. I use this notation throughout the remainder of the paper. We can use the envelope theorem to decompose this partial derivative by holding fixed the firm's quantity choice:

$$\begin{aligned} \rho V_{t,1}(e, \Gamma_{j,t}) &= \frac{1}{1 - \alpha_S} p_V \frac{v_t^*(e, \Gamma_{j,t})}{e} + \frac{\partial p_t(q_t^*(e, \Gamma_{j,t}), e, \Gamma_{j,t})}{\partial e} q_t^*(e, \Gamma_{j,t}) \\ &\quad + \lambda g(E_t(\{e, \Gamma_{j,t}\})) \frac{\partial E_t(\{e, \Gamma_{j,t}\})}{\partial e} \left[\int_0^\infty V_t(e, \{\Gamma_{j,t}, e_t^*(\{e, \Gamma_{j,t}\}, z)\}) f(z) dz - V_t(e, \Gamma_{j,t}) \right] \\ &\quad + \lambda_t^*(\{e, \Gamma_{j,t}\}) \int_0^\infty \frac{\partial V_t(e, \{\Gamma_{j,t}, e_t^*(\{e, \Gamma_{j,t}\}, z)\})}{\partial e_t^*} \frac{\partial e_t^*(\{e, \Gamma_{j,t}\}, z)}{\partial e} f(z) dz \\ &\quad + \eta \sum_{i=1}^{N_{j,t}} [V_{t,1}(e, \{\Gamma_{j,t} \setminus e_{i,j}\}) - V_{t,1}(e, \Gamma_{j,t})] - \eta V_{t,1}(e, \Gamma_{j,t}) \\ &\quad + \lambda_t^*(\{e, \Gamma_{j,t}\}) \left[\int_0^\infty V_{t,1}(e, \{\Gamma_{j,t}, e_t^*(\{e, \Gamma_{j,t}\}, z)\}) f(z) dz - V_{t,1}(e, \Gamma_{j,t}) \right] + \frac{\partial \dot{V}_t(e, \Gamma_{j,t})}{\partial e}. \end{aligned}$$

We can split the various effects of higher effective productivity on the right-hand side into three categories. The first line consists of effects on profits given the current industry effective productivity distribution. The second and third lines consist of effects on the evolution of the industry effective productivity distribution. The final two lines consist of effects on the change in the value function following a change in the industry effective productivity distribution; these effects simply reflect how the other effects change as the distribution changes.

3.4.2 Immediate Effects on Profits

The first term in the first line is the increase in flow profits the firm receives upon entry by producing the same quantity at lower cost. This is the only effect that would remain if the industry were unchanging over time and if other firms' actions did not respond to the entrant's decisions. If this were the only effect, the entrant would set their intangible capital as if they were choosing it with physical capital and labor at each moment in time.

The second term is the effect of the entrant's effective productivity on its demand curve. We saw in the incumbent's optimization problem that the demand curve an incumbent faces depends on

the equilibrium quantity choices of its industry competitors. These equilibrium quantity choices are functions of the distribution of effective productivities in the industry. If a firm has a higher effective productivity, then other firms believe that firm will set a higher quantity and respond by setting lower quantities. As such, if the entrant invests more to get a higher effective productivity, then it faces a higher price for any given quantity choice.

3.4.3 Effects on Industry Structure

The second and third lines are the effects of the entrant's effective productivity on future entry and intangible investment decisions, respectively. As before, if the entrant has higher effective productivity, then other firms expect it to produce more. Future entrants face lower prices and so have less incentive to invest or to pay entry costs and enter.

3.4.4 Changes in Effects Over Time

The final terms are the effects of effective productivity on the entrant's value function following changes to the industry. These effects are similar to those already described, but also capture that the entrant faces risk. The previous terms are positive and so push the entrant to invest more in intangible capital than they would if they could choose intangible capital freely at each moment in time. The effects in the final two lines, on the other hand, have an ambiguous sign that depends on whether higher effective productivity is expected to become more or less valuable over time.

4 Quantifying the Model

To proceed further, I calibrate the model and solve it numerically. A unit of time in the model is one year. Some parameters are assigned. The remaining parameters are calibrated internally;

they are chosen jointly to match a set of moments in the data. The entry cost distribution, G , is log-Normal and the productivity distribution, F , is Pareto. I normalize two parameters: the productivity of capital is $Z_K = \rho + \eta$ and the minimum draw under the Pareto productivity distribution is 1. The former normalization implies that the rental rates of capital and labor are equal at 1.

4.1 Assigned Parameters

The assigned parameters are reported in Table 1. The time discount rate, ρ is chosen to match a 4% steady state real interest rate. The exit and depreciation rate, η , is chosen to approximately match the average exit rate of firms in the US and the depreciation rate of physical capital. Even though the exit rate in the data appears to vary depending on firm size, Covarrubias, Gutiérrez, and Philippon (2019) show that 0.08 is about the rate at which a top 4 firm by sales in a 4-digit industry leaves the top 4. This is perhaps the relevant statistic for large firms in the model.

When I run the main experiment later, I shift the production function away from physical capital and into intangible capital. I want to increase the relevance of intangible capital in the production function while preserving the relative importance of total capital and labor. This is particularly important when considering changes in the labor share, so that they are not due to a change in the relative importance of labor. The labor share in production, $1 - \alpha_S - \alpha_K$, is picked to allow for that while preserving some physical capital after the shock. Finally, the arrival rate of potential entrants, λ , creates a cap on the entry rate. If it is set too low, the internal calibration is not possible.

Table 1: Assigned Parameters

Parameter	Value	Reference
ρ	0.04	Real Interest Rate
η	0.08	Exit/depreciation Rate
$1 - \alpha_S - \alpha_K$	0.7	Labor Share
λ	28	Cap on Entry Rate

4.2 Internally Calibrated Parameters

The internally calibrated parameters, reported in Table 2, are the importance of intangible capital for production, α_S , the tail parameter in the Pareto productivity distribution, the mean and standard deviation of the log entry cost distribution, and the within industry elasticity of substitution, γ .

These parameters are chosen to match the following moments: the average sales shares of the top 8 and top 50 firms in 6-digit NAICS industries in the US in 1997, the revenue-weighted markups of the top firms in the US in 1997 as estimated in Compustat by De Loecker, Eeckhout, and Unger (2018), the elasticity of entry with respect to the value of being an entrant as estimated by Gutiérrez, Jones, and Philippon (2019), and the share of total capital that is intangible among the largest firms as estimated in Compustat by Ewens, Peters, and Wang (2019).

The moments in the model are the following. Each industry in the model maps into a 6-digit industry in the data. As such, the average sales share of the top 8 and top 50 firms are the average sales shares of the top 8 and 50 firms across industries in the model. The total sales share of Compustat firms in the US is about the average sales share of top 8 firms in 6-digit US industries, so the analog of the revenue-weighted average markup across Compustat firms is the revenue-weighted average markup across the top 8 firms from each industry in the model. The same holds for the intangible capital share. The elasticity of entry with respect to the value of

Table 2: Internally Calibrated Parameters

Parameter	Value	Moment	Data	Model
Productivity tail parameter	8.4	Avg. Top 8 Sales Share	0.40	0.40
Log entry cost mean	-4	Avg. Top 50 Sales Share	0.63	0.75
Log entry cost s.d.	0.75	Entry Elasticity	1.03	0.67
α_S	0.1	Intangible Capital Share for Largest Firms	0.37	0.45
γ	15	Revenue-Weighted Markup of Largest Firms	1.37	1.40

being an entrant is measured by regressing the log of the entry rate on an intercept term and the log of the value of being an entrant in the steady state of the model.

Changes to any parameter affect all the above moments. Nonetheless, there is an intuitive mapping from some parameters to particular moments. The entry elasticity is mostly controlled by the standard deviation of the log entry cost distribution. If the standard deviation is high, then a change in the value of entering will have a small effect on the entry rate. For example, as the value of entering goes up and the entry rate rises, the marginal entrant quickly faces a higher entry cost. This is the moment about which we have the least information in the data. I plan to show how the results depend on the value for this moment that I calibrate the model to match.

The intangible capital share is mostly driven by the intangible capital share in production, α_S . As α_S rises and intangible capital replaces physical capital, the intangible capital share increases.

The revenue-weighted average markup among the top firms is mostly controlled by the within industry elasticity of substitution, γ . A firm's markup depends on its market share and the elasticity of substitution. Given the concentration related moments, a higher γ implies lower markups.

The average sales shares of the top 8 and 50 firms are mostly driven by the productivity distribution

tail parameter and the mean of the log entry cost distribution. We can match the top 8 firm sales share with a top 50 firm sales share of 1 by setting a tail parameter of infinity and a high entry cost parameter to get about 20 total firms per industry on average. In that case, each firm has a sales share of about 0.05. To match the top 50 firm sales share while continuing to match the top 8 firm sales share, we can decrease the entry cost to get more firms and decrease the tail parameter to increase the sales share of the top firms. Eventually, when the tail parameter is low enough, an increase in the number of firms can actually raise the sales share of the top 8 firms while reducing the sales share of the top 50 firms. This is because as the number of firms increases, there are two forces present. First, the top 8 firms' productivities are higher. Second, large firms' markups are more responsive to competition than are smaller firms' markups. Hence, as competition increases from a higher number of firms, large firms cut their markups and hold on to their sales shares.

5 Numerical Techniques

In this section, I cover the methods used to solve the model numerically on the computer. I mostly build on Benkard, Jeziorski, and Weintraub (2015), Ifrach and Weintraub (2017), and Doraszelski and Judd (2004), who focus on methods for solving finite firm models with a large number of firms. There are two important takeaways. First, a good approximation is to keep track of the states of the largest firms in an industry as well as some moments of the distribution of states among smaller firms. Second, continuous time dramatically reduces computational cost as compared to discrete time because the number of possible industry states to which a given state can transition is much smaller; two shocks never happen simultaneously. First I describe how to keep track of many small firms. Then I describe how to keep track of the states of the large firms.

5.1 Many Small Firms

First, consider an approximation for many small firms in a single moment in time. For firms with a small market share, their markup is well approximated by $\gamma/(\gamma - 1)$. Moreover, if M_t firms with effective productivities $\{e_1, e_2, \dots, e_{M_t}\}$ each set the same markup $\gamma/(\gamma - 1)$ and industry consumption is C , then firm i 's production (raised to the appropriate power in the CES aggregator) is

$$q_i^{\frac{\gamma-1}{\gamma}} = \left(\frac{\gamma}{\gamma-1} \frac{p_V}{1-\alpha_S} \right)^{\frac{(1-\gamma)(1-\alpha_S)}{\gamma+(1-\alpha_S)(1-\gamma)}} C^{\frac{-(1-\alpha_S)(\gamma-1)^2}{\gamma+(1-\alpha_S)(1-\gamma)}} e_i^{\frac{\gamma-1}{\gamma+(1-\alpha_S)(1-\gamma)}},$$

their revenue is

$$p_i q_i = \left(\frac{\gamma}{\gamma-1} \frac{p_V}{1-\alpha_S} \right)^{\frac{(1-\gamma)(1-\alpha_S)}{\gamma+(1-\alpha_S)(1-\gamma)}} C^{\frac{(1-\gamma)}{\gamma+(1-\alpha_S)(1-\gamma)}} e_i^{\frac{\gamma-1}{\gamma+(1-\alpha_S)(1-\gamma)}},$$

and their variable costs are

$$v_i = \left(\frac{\gamma}{\gamma-1} \frac{p_V}{1-\alpha_S} \right)^{\frac{-\gamma}{\gamma+(1-\alpha_S)(1-\gamma)}} C^{\frac{(1-\gamma)}{\gamma+(1-\alpha_S)(1-\gamma)}} e_i^{\frac{\gamma-1}{\gamma+(1-\alpha_S)(1-\gamma)}}.$$

It follows that from the perspective of the labor, capital rental, and goods markets, from the perspective of the profits received by the household, as well as from the perspective of the other firms in the industry, it is equivalent if, instead of the M_t firms, there is a single firm that sets a markup of $\gamma/(\gamma - 1)$ with effective productivity $e_0 = M_t^{\frac{\gamma+(1-\alpha_S)(1-\gamma)}{\gamma-1}} \bar{e}$, where \bar{e} is an average effective productivity:

$$\bar{e} = \left[\frac{1}{M_t} \sum_{i=1}^{M_t} e_i^{\frac{\gamma-1}{\gamma+(1-\alpha_S)(1-\gamma)}} \right]^{\frac{\gamma+(1-\alpha_S)(1-\gamma)}{\gamma-1}}.$$

Further, if each firm i gets a share $(e_i/e_0)^{\frac{\gamma-1}{\gamma+(1-\alpha_S)(1-\gamma)}}$ of the profits of the single firm, then they are indifferent between the two situations as well.

Next, consider the approximation over time. If M_t is sufficiently large and the distribution of effective productivity among the M_t firms has a sufficiently thin tail, then \bar{e} is constant as firms exit and the change in M_t over time is well approximated by $\dot{M}_t = -\eta M_t$. It follows that over

time, the single firm is equivalent to the many firms if rather than exiting at rate η , the single firm's effective productivity evolves according to

$$\dot{e}_{0,t} = -\frac{\gamma + (1 - \alpha_S)(1 - \gamma)}{\gamma - 1} \eta e_{0,t}.$$

If an entering firm is absorbed into the single firm, then the single firm's effective productivity increases accordingly. Finally, if we define $\pi_{0,t}$ to be the profits of the single firm divided by $e_{0,t}^{\frac{\gamma-1}{\gamma+(1-\alpha_S)(1-\gamma)}}$ and $V_{0,t}$ to be the expected present discounted value of $\pi_{0,t}$, then the value of an entering firm with sufficiently low effective productivity is well approximated by $V_{0,t} e_i^{\frac{\gamma-1}{\gamma+(1-\alpha_S)(1-\gamma)}}$.

To summarize, rather than keep track of the effective productivities of many small firms, we can keep track of an appropriate aggregate of their effective productivities. To determine the value of a small firm upon entering as a function of effective productivity, we only need to keep track of an appropriate average value of being a small firm.

There is a discrete set of values for the aggregate small firm effective productivity. I keep track of the small firm value, $V_{0,t}$, in each state.

5.2 A Few Large Firms

I keep track of the effective productivities of up to a finite number of firms. There is a finite set of values in which these effective productivities can fall. Since strategies can only depend on the distribution of effective productivity, it follows that we only need to keep track of the number of firms at each effective productivity value in the set rather than the effective productivity of each firm. To conserve on the number of states, I allow for fewer firms at larger values of effective productivity. This is appropriate given the Pareto distribution of productivity; higher levels of effective productivity come from higher productivity draws, which are less likely.

At the boundary of the state space, I give a firm a value of 0 if it enters at an effective productivity at which no more firms can enter. This should only further incentivize firms to enter and push the

state space to the boundary. I check in equilibrium that the economy is rarely at the boundary.

The state space is multidimensional and not a grid. For example, if there are more firms at the lowest effective productivity value, then the maximum number of firms at the higher values is smaller. Doraszelski and Judd (2004) have some discussion and references on ways to order the state space. What is most important is to preset certain maps between states. For example, a function that sends a state n to a state n' if a firm enters at the third effective productivity value. I preset functions for the entry and exit of a firm at each effective productivity value.

5.3 Solving the Entrant Problem

If a firm pays the entry cost, I randomly draw whether the firm will be large or small by drawing whether the firm's productivity is above a threshold. If the firm is small, then it chooses intangible capital s to maximize

$$V_{0,t}(zs^{\alpha s})^{\frac{\gamma-1}{\gamma+(1-\alpha s)(1-\gamma)}} - p_K s.$$

I solve the first order condition for s as a function of z . The firm receives the expected value over values of z below the threshold for being small. The aggregate small firm effective productivity in the industry grows by the average over values of z below the threshold for being small. Note that the first order condition for s is always necessary and sufficient because the value is strictly concave in s .

If the firm is large, then its value function is piecewise linear (linear between the effective productivity values in the feasible set) in $e^{\frac{\gamma-1}{\gamma+(1-\alpha s)(1-\gamma)}}$, where e is its effective productivity. This implies that the value function is concave in this transformation of effective productivity. If a firm is small, then its value function is linear in this transformation of effective productivity. As a firm's sales share and markup grow, I expect its marginal value of effective productivity to fall; it is increasingly competing with itself and does not use further increases in effective productivity to expand production. As such, a large firm's true value function should be concave in this transformation of effective productivity. Of course, this is not guaranteed because firms have many incentives for

investing in intangible capital and increasing effective productivity, but I confirm that it is true in equilibrium. Then, as for small firms, a large firm's first order condition for intangible capital is necessary and sufficient.

If a firm is large, I draw productivity from a discrete grid above the threshold necessary for being large. I find the intangible capital at which its first order condition is satisfied. If the resulting effective productivity is between two points in the feasible set, the firm goes to each of the two points with the appropriate probability so that its value function is linear in the transformation of effective productivity given above.

To check that the threshold for being large is appropriate, I check that small and large firms near the threshold end up with small sales shares.

To find the value of entering, $E_t(\Gamma_{j,t})$ – where $\Gamma_{j,t}$ now only includes the effective productivities of the largest firms and the aggregate small firm effective productivity – I take the expected value over the draws of small vs. large and productivity conditional on being large. The entry rate is $\lambda G(E_t(\Gamma_{j,t}))$ and the flow entry cost is an approximation of $\int_0^{E_t(\Gamma_{j,t})} c_E g(c_E) dc_E$.

5.4 Solving the Incumbent Problem

First, I find the optimal quantity and markup choices of firms in each industry state. These are purely static decisions that don't depend on any aggregate features of the equilibrium, so they can be solved once given the exogenous parameters.

Next, I solve for the value functions of incumbents by iterating on the Hamilton-Jacobi-Bellman equation. In each iteration, in each state, there is a guess for the value function for each value of effective productivity in the feasible set, $V_t(e, \Gamma_{j,t})$, as well as for the aggregate small firm, $V_{0,t}(\Gamma_{j,t})$. I use the value function to update the entry rate and intangible investment decisions of new entrants. Then, for a particular state and effective productivity, I take as given the value

function at other states and the updated entry rate and investment decisions. These imply a value in the current state and effective productivity, $\hat{V}_t(e, \Gamma_{j,t})$, so that the Hamilton-Jacobi-Bellman equation holds exactly:

$$\begin{aligned}
(\rho + \eta N_{j,t} + \lambda_t^*(\Gamma_{j,t}) \hat{V}_t(e, \Gamma_{j,t}) &= p_t(q_t^*(e, \Gamma_{j,t}), e, \Gamma_{j,t}) q_t^*(e, \Gamma_{j,t}) - p_V v_t^*(e, \Gamma_{j,t}) \\
&+ \eta \sum_{i' \neq i}^{N_{j,t}} V_t(e, \{\Gamma_{j,t} \setminus e_{i',j}\}) \\
&+ \lambda_t^*(\Gamma_{j,t}) \int_0^\infty V_t(e, \{\Gamma_{j,t}, e_t^*(\Gamma_{j,t}, z)\}) f(z) dz \\
&+ \dot{V}_t(e, \Gamma_{j,t}).
\end{aligned}$$

I update the value function by taking an average of the old and new value functions:

$$V_t(e, \Gamma_{j,t}) = (1 - \Delta) V_t(e, \Gamma_{j,t}) + \Delta \hat{V}_t(e, \Gamma_{j,t}).$$

The step size Δ must be small enough for the value function to converge. When the value function converges, we have a markov perfect equilibrium.

This method for finding a markov perfect equilibrium searches for the equilibrium that is the infinite horizon limit of the finite horizon equilibria found through backward induction.

5.5 Finding a Steady State

I begin at a distribution across industries that puts equal weight on each state in the state space. I use a discrete time approximation of the continuous time transition matrix to iterate the distribution forward. I take the set of most likely states. Beginning at each state in this set, I use the arrival rates of various shocks to iterate forward a hundred thousand years. I take the distribution over time, throwing away a long period at the beginning, as the steady state cross-sectional distribution.

To verify that the distribution is a steady state, I draw a large cross-section of industries from the

proposed steady state distribution. I iterate the industries forward over time and confirm that aggregate variables do not change.

5.6 Computing a Transition Path

First, note that firms only care about their industry state and the exogenous parameters. Aggregates of interest, such as the interest rate, wage, capital price, and capital rental rate do not change along transition paths. As such, I solve for value and policy functions for each set of exogenous parameters.

To compute a transition path, I draw a large cross-section of industries from the initial steady state distribution. Using the policy functions from the new set of exogenous parameters, I iterate the industries forward over time until aggregate variables converge.

6 Results

The main experiment of the paper is the following. I solve for the steady state of the model. Then, I increase α_S from 0.1 to 0.2, while holding fixed the labor share in production at 0.7. The shock is designed to match the increase among top firms in the share of capital that is intangible, as estimated by Ewens, Peters, and Wang (2019). I also multiply all realizations of productivity, z , by $(\alpha_{0,S}^{\alpha_{0,S}} \alpha_{0,K}^{\alpha_{0,K}}) / (\alpha_S^{\alpha_S} \alpha_K^{\alpha_K})$ – where $\alpha_{0,S}$ and $\alpha_{0,K}$ are the original values for α_S and α_K , respectively – to preserve the marginal cost of production if a firm could choose all factors of production freely at each moment. This is only so that the shift from physical capital toward intangible capital does not also essentially imply a mechanical change in the productivity distribution. These changes are unanticipated by agents in the model and are permanent. Table 3 shows key features of the original and the new steady states. The model is able to explain more than half the increase in average top 8 firm market share from 1997 to 2012 and about a third of the increase in the revenue-weighted

Table 3: Experiment: Change in α_S

Moment	Data	Model
Intangible Capital Share (1980 / Original Steady State)	0.37	0.45
Intangible Capital Share (2016 / New Steady State)	0.60	0.75
Top 8-firm share (1997 / Original Steady State)	40.1	39.8
Top 8-firm share (2012 / New Steady State)	45.9	43.6
Avg. markup (1997 / Original Steady State)	1.40	1.37
Avg. markup (2012 / New Steady State)	1.50	1.40

average markup among top firms.

Note that the change in intangible capital in the data is from 1980 to 2016 while the change in concentration and markups is from 1997 to 2012. I choose this time period for concentration and markups because there is a change in the concentration data between 1992 and 1997, making it difficult to compare concentration measures before 1992 with measures after 1997. Nonetheless, the authors mentioned earlier that document these trends show that the trends appear to be similar during the 1980s and early 1990s. Of course, if I considered the full time period of 1980 to 2016 for concentration measures and markups, I would match a smaller fraction of the changes.

To solve for a transition path, we must decide what happens to incumbents when the shock occurs. It is computationally difficult to keep track of two types of firms, those with α_S at the original level and those with α_S at the new level. Moreover, it is perhaps desirable to shock α_S for incumbents as well as for new firms. I take half the physical capital that incumbents were choosing in the original steady state and fix that as intangible capital following the shock. This way, the total capital stock and output do not immediately change following the shock. An alternative would be to allow incumbents to choose their new intangible capital levels. I expect that, in this case, the economy would converge more quickly toward the new steady state. In the former case that I compute, the transition path is slowed because, when the physical capital becomes fixed intangible

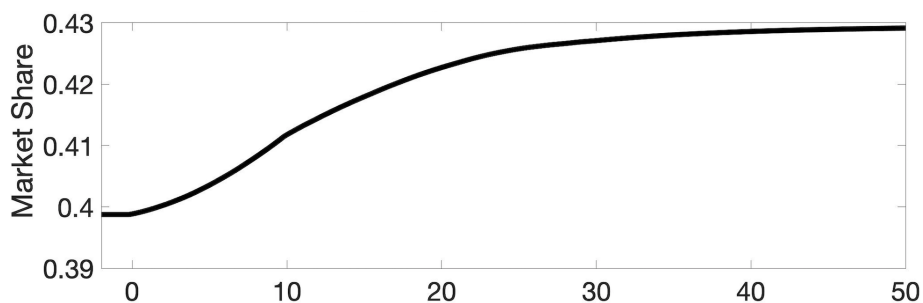


Figure 1: Average Top-8 Firm Market Share

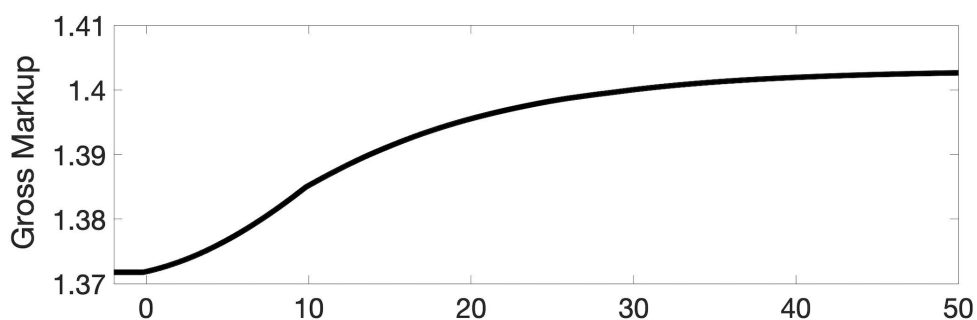


Figure 2: Revenue-Weighted Markups of Top Firms

capital, concentration and markups do not change but new entrants who will bring that change are discouraged from entering due to the increased fixedness of incumbents' output.

Figures 1 and 2 show the transition paths for the main objects of interest, concentration and markups. After 15 years, the time period in the data, concentration and markups have converged by about two-thirds toward their new steady state values. To understand what is driving this result, consider the transition paths of the entry rate, investment, labor used in production by goods producing firms, and output shown in Figure 3.

At first, as noted above, labor used in production and output do not change. Investment rises because new entrants, on average, use more of a fixed factor of production than they would if the factor were flexible. The deterrence motive outweighs any risk motive that pushes a firm to reduce

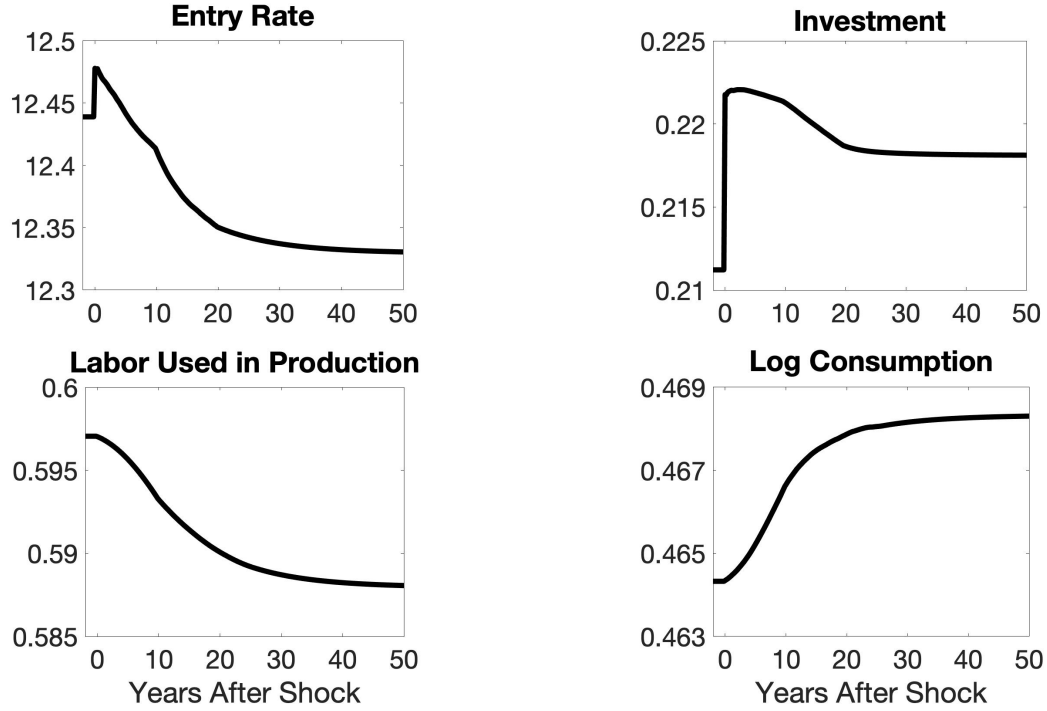


Figure 3: Evolution of Aggregates Following the Shock to α_S

investment. The entry rate rises slightly because potential entrants prefer to have the ability to deter through the fixed factor of production. They know that if they are productive, then they can take a larger share of the market. Over time, this rise in investment leads to higher output. Yet, labor used in production by final goods producers falls. Production shifts toward firms who have larger market shares, set higher markups, and rely more on effective productivity rather than labor. Over time, the entry rate and investment fall as well due to the increased competition from the higher output.

Figure 4 shows the effect of the shock on the market share of firms upon entry. The top panel shows, in the original steady state, the average market share upon entry by firms conditional on their productivity draw, z . The bottom panel shows how this market share changes following the shock. The solid line shows the ratio of the new average market share to the old immediately

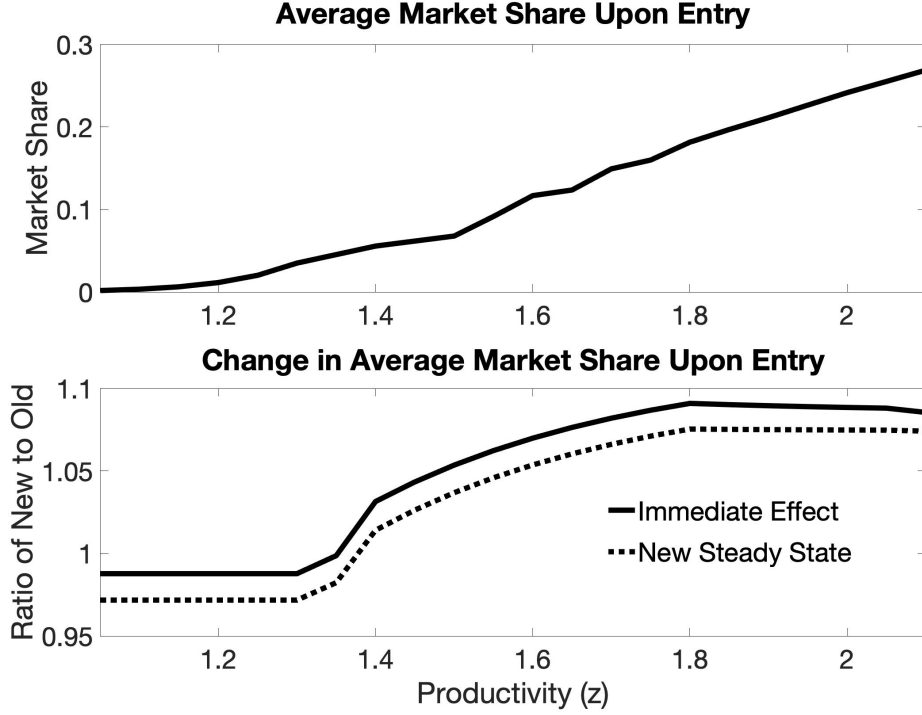


Figure 4: Market Share Upon Entry as a Function of the Productivity Draw

following the shock. The dashed line shows the same in the new steady state.

Immediately following the shock, more productive firms increase their market share by investing heavily in intangibles. Less productive firms lose market share. They are reluctant to invest because of the risk inherent in a fixed factor of production as well as because they anticipate increased output from their future competitors. Over time, as output rises, all firms' market shares fall.

Figure 5 decomposes an entrant's incentive to invest in intangible capital in the new steady state. To generate the figure, I do the following. For the top panel, I add a new firm at various effective productivities and map out the quantity index of its competitors (on the left axis), $\left[\frac{N_{j,t}}{\sum_{i' \neq i} c_{i,j,t}^{\frac{\gamma}{\gamma-1}}} \right]^{\frac{\gamma-1}{\gamma}}$,

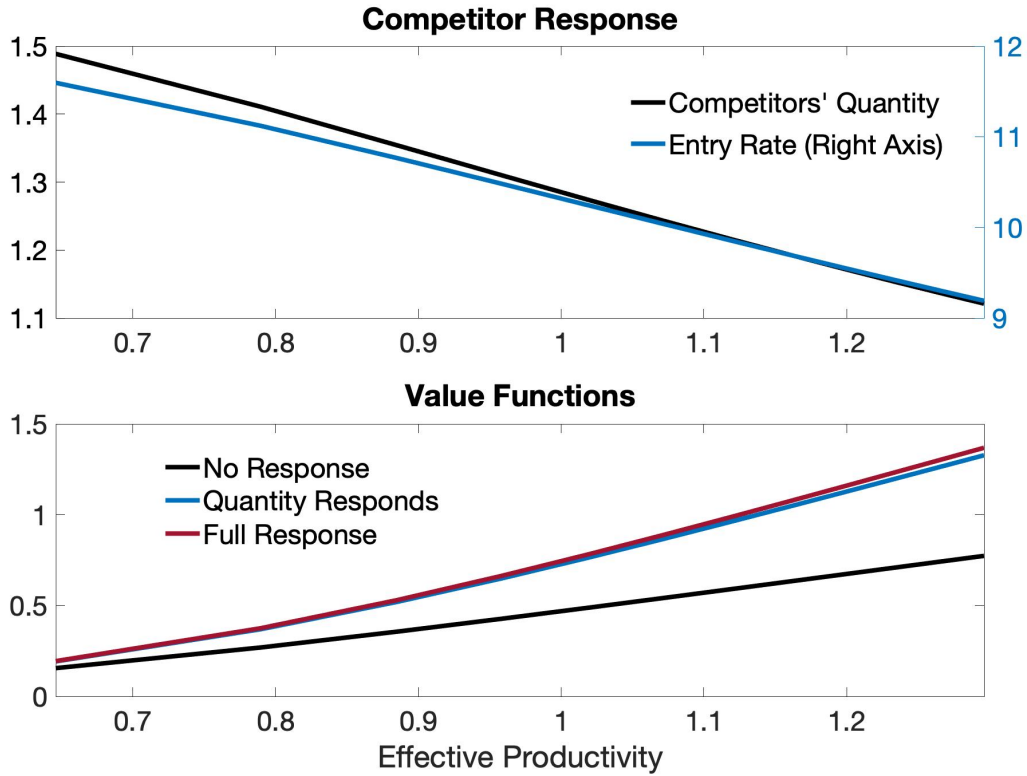


Figure 5: The top panel shows the responses of incumbents and potential entrants in an industry upon the entry of a new firm, as a function of the new firm's effective productivity. The bottom panel shows the entrant's value function taking into account various degrees of responsiveness by other firms.

and the entry rate (on the right axis). I take the average of these values across the steady state distribution of industries. For the bottom panel, I look at the value function of this entering firm, again averaging over the steady state distribution. The black line shows the value the new firm would receive if it entered and all its competitors' actions – variable cost, entry, and intangible investment decisions – did not respond. The blue line shows the value the new firm would receive if it entered and its competitors' entry and intangible investment decisions did not respond, but their variable cost decisions did. The red line is the actual value the new firm receives in equilibrium.

The black line captures the incentive a firm faces to increase a flexible factor of production because

other firms cannot respond to this choice. The difference between the black and blue lines is the value the new firm receives from the response of its competitors' variable cost choices to its entry. The top panel shows that both entry and variable cost choices are responsive. The responsiveness of variable costs has a large effect on the incentive of the entrant to invest in intangible capital – the blue line is about twice as steep as the black line – as well as to pay the entry cost – the blue line is above the black line. The responsiveness of the entry rate and of future entrants' investment, on the other hand, have almost no effect on the entrant's intangible capital choice or entry decision. Firms simply exit too quickly to care much about deterring future entrants.

6.1 Other Features of the Rise in Concentration

Related to the rise in markups in the data is the fall in the labor share. Barkai (2017) shows that industries in which concentration rose also experienced falls in their labor shares. They regress the five year change in the labor share in a 6-digit NAICS industry on the five year change in the market share of the top 8 firms. There are three five year changes between 1997 and 2012. They find a coefficient of -0.108 on the change in the market share. I run the same regression over the first 15 years of the transition path and find a coefficient of -0.172.

Ganapati (2019) shows that industries in which concentration rose also experienced increases in labor productivity. They standardize log labor productivity (real output per unit of labor) and the log of the top 4 firm market share in 6-digit NAICS industries by subtracting the means and dividing by the standard deviations. As above, they regress five year changes in standardized log labor productivity on five year changes in standardized log market shares. They find a coefficient of 0.206 between 1997 and 2012. I run the same regression over the first 15 years of the transition path and find a coefficient of 0.278. The relationship between concentration and labor productivity helps distinguish the mechanism in this paper from other proposed drivers of increased concentration. For example, if the rise in concentration were due to higher entry costs, then labor productivity would be constant.

7 Welfare

Households own all profits, pay all labor costs, and consume all final goods, so their present discounted utility is the measure of welfare. Flow utility is the log of final good output minus the labor used in final good production, in capital production, and on entry costs. Following the shock, we can compute the change in welfare by computing the changes in the present discounted values of log final good output and total labor supply. I find that welfare increased by the equivalent of a 0.3% permanent increase in consumption or, equivalently, a permanent rise in log final good output of 0.003 or a permanent fall in total labor supply of 0.003. As discussed above, I expect that if we allowed incumbents to choose intangible capital following the shock, then the transition path would be faster. In that case, the increase in welfare would likely be larger because the increase in steady state welfare is larger than the increase when taking into account the transition path.

By looking at Figure 3, we can see where the increase in welfare comes from. More labor is used to create capital, but less is used on entry costs and in final good production. Final good output is higher. If investment rose and replaced labor at the firm level, output and welfare would not rise because the physical capital to labor ratio is optimal from the perspective of minimizing cost. Output and welfare rise because the increase in investment embodies a shift in production from firms with lower productivity to firms with higher productivity. The firms that gain market share are able to produce more with less labor.

Figure 6 compares the incentives of a new entrant for profit maximization vs. for welfare maximization. Similar to Figure 5, the right panel shows the value for the firm (in black) and welfare (in blue) that result from a firm entering the average industry (in the new steady state following the shock) at various levels of effective productivity if other firms' actions and future entrants' actions are taken as given. The left panel shows the value for the firm and welfare that result from a firm entering in the actual equilibrium, taking into account that other firms' actions and future entrants' actions depend on the entering firm's effective productivity.

As in Figure 5, the right panel illustrates the incentives for profit maximization vs. for welfare

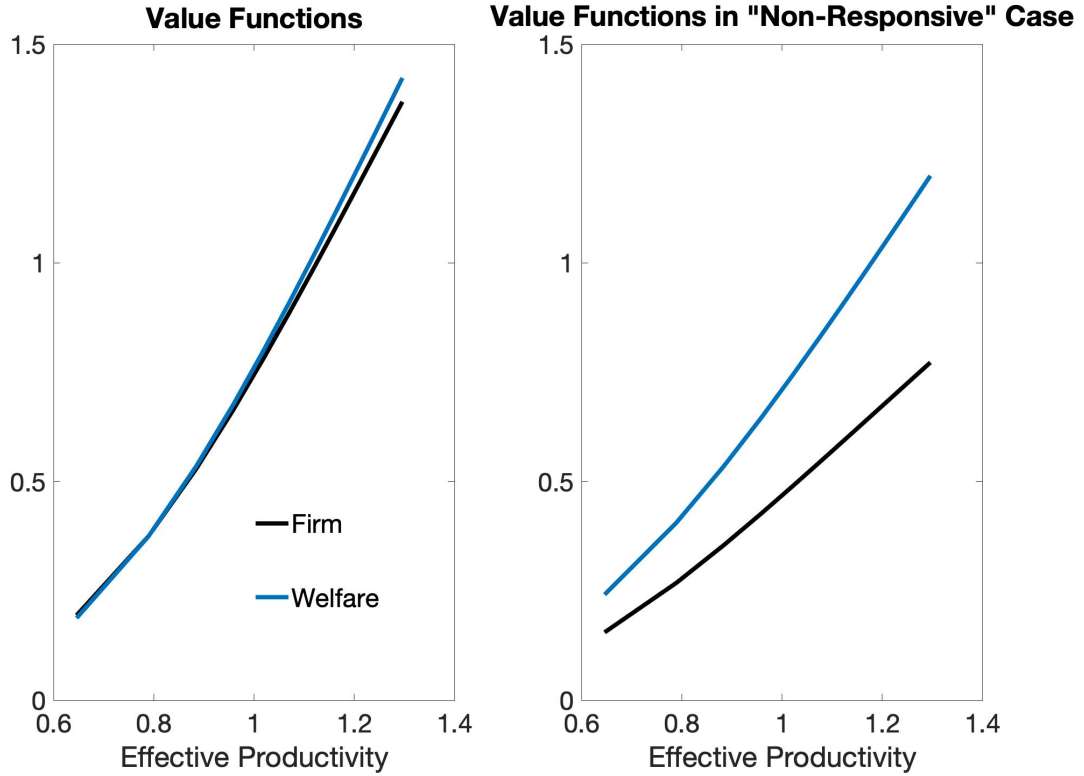


Figure 6: The left panel shows expected discounted profits and welfare in equilibrium. The right panel shows expected discounted profits and welfare if other firms do not respond to the entrant.

maximization when choosing a flexible factor of production. The line for welfare is steeper than for profits, which suggests that, from a welfare perspective, the firm underproduces. As the firm produces more, it is increasingly competing with itself; each new unit of output the firm produces lowers the price of all units of output produced by other firms as well as by itself.

Taking into account the responses of competitors, in the left panel, the line for profits is quite similar to the line for welfare. There are many factors at work here. For all values of effective productivity, the profit line in the left panel is higher than the profit line in the right panel. The firm always gains profits from the response of other firms to its entry and investment decision.

For low values of effective productivity, the welfare line in the left panel is lower than in the right panel, and for high values of effective productivity, the welfare line in the left panel is higher than in the right panel. The reallocation toward a low productivity entrant is excessive from a welfare perspective and the reallocation toward a high productivity entrant improves welfare. Finally, even though the two lines are similar in the left panel, the welfare line is below the profit line for low values of effective productivity and is above the profit line for high values of effective productivity. The welfare line is still steeper than the profit line. Given the costly reallocation toward low productivity firms, their profits overstate their value from a welfare perspective. On the other hand, even though high productivity firms invest to deter competitors, they still underinvest and their profits understate their value from a welfare perspective. The responses of other firms and the incentives they create ultimately serve to bring a firm's private incentive to invest into closer alignment with the social incentive to invest.

7.1 Constrained Planner Problem

It is useful to analyze the following planner problem. When a potential entrant arrives, the planner chooses whether the entrant pays the entry cost. Moreover, if the firm enters, the planner chooses the intangible investment of the new firm. The planner makes these decisions to maximize expected present discounted welfare. Incumbents choose their variable costs to maximize profits. Taking into account the transition path, welfare increases by the equivalent of a 0.1% permanent rise in final good consumption.

Figures 7 and 8 show the transition paths of key aggregates following the takeover of the planner starting from the steady state with $\alpha_S = 0.2$ (after the shock). In line with the comparison of the incentives for welfare maximization vs. profit maximization in Figure 6, the planner reduces entry and increases intangible investment conditional on entry. Since most firms have low productivity, the fact that profits overstated the value of low productivity entrants to the planner pushes the planner to reduce entry.

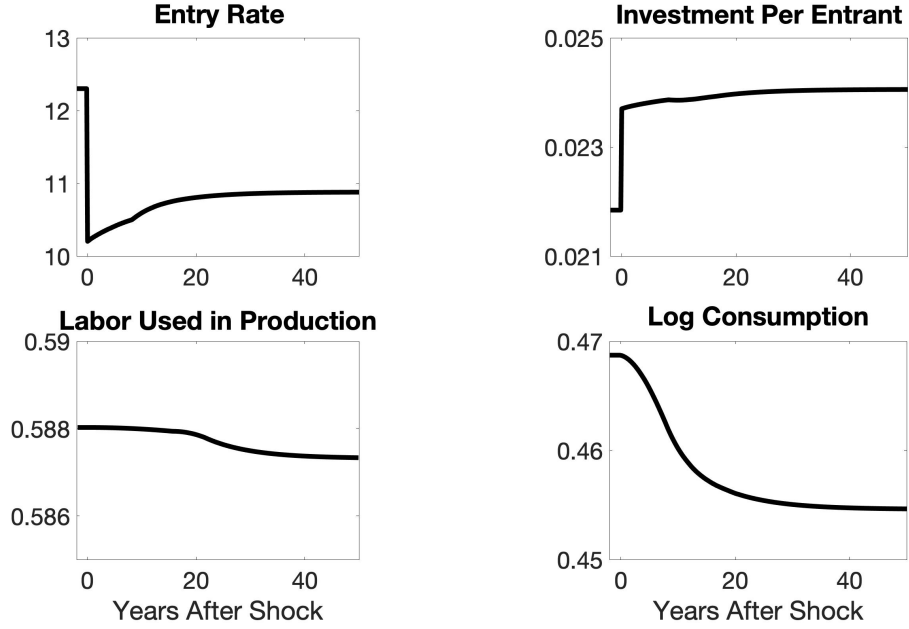


Figure 7: Evolution of Aggregates under the Constrained Planner Starting from the Steady State with High α_S

The fall in entry under the planner reduces labor productivity, output, and concentration. This is because less entry means fewer high productivity firms. Since the markup is convex in market share, as more firms enter, large productive firms cut their markups and resist falls in their market shares while small unproductive firms leave their markups unchanged and let their market shares fall. As such, high productivity firms mostly take market share from low productivity firms, not from each other. Fewer high productivity firms then means more market share is held by low productivity firms.

The revenue-weighted average markup of the top firms rises under the planner. The remaining high productivity firms are larger and set higher markups because they face less competition due to lower entry and because they invested more heavily in intangible capital.

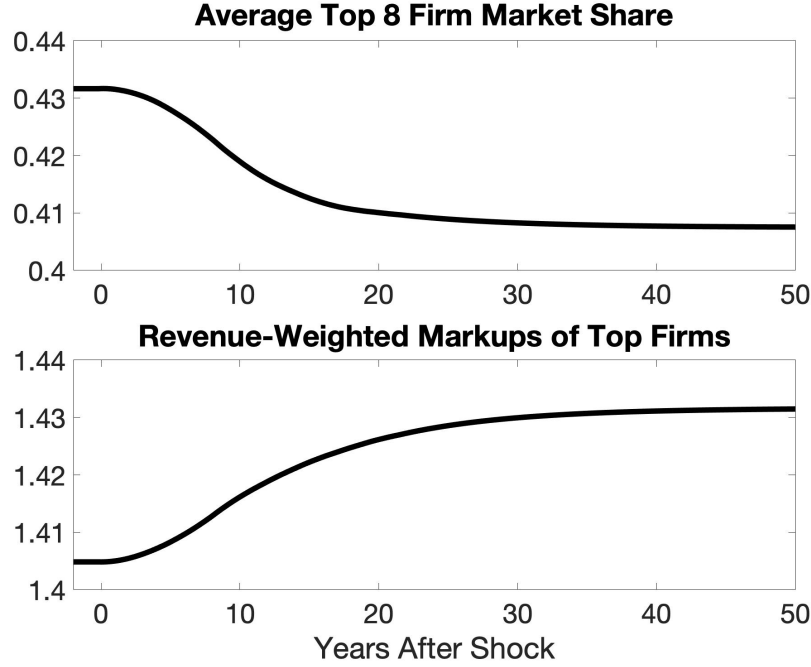


Figure 8: Evolution of Concentration and Markups under the Constrained Planner Starting from the Steady State with High α_S

8 Conclusion

In this paper, I show that an increase in the importance of irreversible intangible capital in line with estimates from the literature can quantitatively explain a meaningful portion of recent changes in industry concentration, markups, and the labor share. Large firms use intangible capital to commit to higher production, which deters competition through other firms' flexible input choices as well as future entrants' entry and investment decisions. The shift in technology toward intangible capital improves welfare because it reallocates production away from small low productivity firms toward large high productivity firms. The shift effectively acts as a size-dependent subsidy, which is the optimal policy in the presence of markup dispersion in a variety of models, such as in Edmond, Midrigan, and Xu (2019).

I solve a constrained planner problem and find that the planner wants to reduce entry and increase intangible investment. This suggests that, despite their “anti-competitive” motives, large firms are *underinvesting* in intangible capital. These results provide evidence for the argument that we should be hesitant to design policies that limit investment or production by large firms with high markups. Indeed, small unproductive firms’ profits overstate their value to the planner and large productive firms’ profits understate their value to the planner. This suggests that if mergers and acquisitions were allowed in the model, firms would undertake *too few* relative to what the planner would choose. This warrants further exploration.

The main results also suggest a new consideration when evaluating welfare and developing policy. Rising markups may reflect increased irreversible intangible investments, which is good for welfare. Creating markets for otherwise illiquid capital or forcing firms to relinquish that capital, such as proprietary data, may only serve to reallocate production to less productive firms. This may be an important consideration when evaluating policies such as a firing tax that limit firms’ abilities to adjust their inputs. Yet, it is important to keep in mind that reducing flexibility likely has diminishing returns and may eventually reduce welfare. The responsiveness of other firms’ flexible input choices is necessary for a new productive firm to be able to invest heavily in intangible capital and take a large market share.

References

- Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li.** 2019. “A Theory of Falling Growth and Rising Rents.” Working paper.
- Akcigit, Ufuk and Sina Ates.** 2019. “Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory.” NBER Working Paper 25755.
- Amiti, Mary, Oleg Itskhoki, and Jozef Konings.** 2019. “International Shocks, Variable Markups, and Domestic Prices.” *The Review of Economic Studies*, 86(6): 2356-2402.
- Atkeson, Andrew and Ariel Burstein.** 2008. “Pricing-to-Market, Trade Costs, and International Relative Prices.” *American Economic Review*, 98(5): 1998-2031.
- Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen.** 2019. “The Fall of the Labor Share and the Rise of Superstar Firms.” NBER Working Paper 23396.
- Barkai, Simcha.** 2017. “Declining Labor and Capital Shares.” Chicago Booth mimeo.
- Benkard, C. Lanier, Przemyslaw Jeziorski, and Gabriel Y. Weintraub.** 2015. “Oblivious Equilibrium for Concentrated Industries.” *The RAND Journal of Economics*, 46(4): 671-708.
- Bessen, James E.** 2017. “Industry Concentration and Information Technology.” Boston Univ. School of Law, Law and Economics Research Paper No. 17-41.
- Bornstein, Gideon.** 2018. “Entry and Profits in an Aging Economy: The Role of Consumer Inertia.” Working Paper.

- Bronnenberg, Bart J., Sanjay K. Dhar, and Jean-Pierre H. Dubé.** 2011. “Endogenous Sunk Costs and the Geographic Differences in the Market Structures of CPG Categories.” *Quantitative Marketing and Economics*, 9(1): 1-23.
- Bronnenberg, Bart J., Jean-Pierre Dubé, and Matthew Gentzkow.** 2012. “The Evolution of Brand Preferences: Evidence from Consumer Migration.” *American Economic Review*, 102(6): 2472-2508.
- Corrado, Carol, Jonathan Haskel, Cecilia Jona-Lasinio, and Massimiliano Iommi.** 2012. “Intangible Capital and Growth in Advanced Economies: Measurement Methods and Comparative Results.” IZA Discussion Paper Series 6733.
- Covarrubias, Matias, Germán Gutiérrez, and Thomas Philippon.** 2019. “From Good to Bad Concentration? U.S. Industries Over the Past 30 Years.” NBER Working Paper 25983.
- Crouzet, Nicolas and Janice C. Eberly.** 2019. “Understanding Weak Capital Investment: the Role of Market Concentration and Intangibles.” NBER Working Paper 25869.
- De Loecker, Jan, Jan Eeckhout, and Gabriel Unger.** 2018. “The Rise of Market Power and the Macroeconomic Implications.” Working paper.
- Doraszelski, Ulrich and Kenneth L. Judd.** 2004. “Solving Continuous-Time, Jump Process Games with Discrete States.” Mimeo.
- Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu.** 2015. “Competition, Markups, and the Gains from International Trade.” *American Economic Review*, 105(10): 3183-3221.
- Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu.** 2019. “How Costly Are Markups?” NBER Working Paper 24800.

- Eggertsson, Gauti B., Jacob A. Robbins, and Ella Getz Wold.** 2018. “Kaldor and Piketty’s Facts: the Rise of Monopoly Power in the United States.” Working paper.
- Eisfeldt, Andrea L. and Dimitris Papanikolaou.** 2013. “Organization Capital and the Cross-Section of Expected Returns.” *The Journal of Finance*, 68(4): 1365-1406.
- Ewens, Michael, Ryan H. Peters, and Sean Wang.** 2019. “Acquisition Prices and the Measurement of Intangible Capital.” NBER Working Paper 25960.
- Foster, Lucia, John Haltiwanger, and Chad Syverson.** 2016. “The Slow Growth of New Plants: Learning about Demand?” *Economica*, 83(329): 91-129.
- Ganapati, Sharat.** 2019. “Oligopolies, Prices, Output, and Productivity.” Working paper.
- Gouin-Bonenfant, Emilien.** 2018. “Productivity Dispersion, Between-Firm Competition, and the Labor Share.” Working paper.
- Gutiérrez, Germán, Callum Jones, and Thomas Philippon.** 2019. “Entry Costs and the Macroeconomy.” NBER Working Paper 25609.
- Hall, Robert E.** 2018. “Using Empirical Marginal Cost to Measure Market Power in the US Economy.” NBER Working Paper 25251.
- Ifrach, Bar and Gabriel Y. Weintraub.** 2017. “A Framework for Dynamic Oligopoly in Concentrated Industries.” *The Review of Economic Studies*, 84(3): 1106-1150.
- Jones, Callum and Thomas Philippon.** 2016. “The Secular Stagnation of Investment?” Working paper.

- Liu, Ernest, Atif Mian, and Amir Sufi.** 2019. “Low Interest Rates, Market Power, and Productivity Growth.” NBER Working Paper 25505.
- Peters, Ryan H. and Lucian A. Taylor.** 2017. “Intangible Capital and the Investment-q Relation.” *Journal of Financial Economics*, 123(2): 251-272.
- Spulber, Daniel F.** 1981. “Capacity, Output, and Sequential Entry.” *American Economic Review*, 71(3): 503-514.
- Sutton, John.** 1991. *Sunk Costs and Market Structure*. The MIT Press.