Dynamics of Markups, Concentration and Product Span

Elhanan Helpman and Benjamin Niswonger*
Harvard University

This version: November 24, 2020

Abstract

We develop a model with a finite number of multi-product firms that populate an industry together with a continuum of single-product firms and study the dynamics of this industry that arise from investments in the invention of new products. Consistent with the available evidence, the model predicts rising markups and concentration and a declining labor share. We then examine the dynamics of market shares and product spans in response to improvements in the technologies of the multi-product and single-product firms, and the impact of these changes on the steady state distribution of market shares and product spans. Our model predicts the possibility of an inverted-U relationship between labor productivity and product span in the cross-section of firms, for which we provide suggestive evidence. It also predicts that rising entry costs of single-product firms may flatten the relationship between labor productivity and market shares of the large multi-product firms. In addition to analyzing market outcomes, we characterize the optimal dynamic allocation. In the steady state of this allocation, there is a monotonic relationship between labor productivity and product span in the cross-section of firms; more productive firms have larger product spans. This highlights a distinct facet of the role of market power in distorting the economy.

Keywords: single- and multi-product firms, firm dynamics, industry dynamics, markup, market share, product span

JEL Classification: L11, L13, L25, D43

*Helpman: Department of Economics, Harvard University, Cambridge, MA 02138, NBER and CEPR (e-mail: ehelpman@harvard.edu); Niswonger: Department of Economics, Harvard University, Cambridge, MA 02138 (e-mail: niswonger@g.harvard.edu). We thank Thomas Sampson for useful comments.

1 Introduction

A number of recent studies have investigated the evolution of markups and the growth of concentration in U.S. industries, finding that both markups and concentration have increased; see Autor *et al.* (2020); De Loecker *et al.* (2020). These studies find that the ascent in average markups was driven by rising markups of the largest firms and market share reallocation from low- to high-markup firms. Contemporaneously, the labor share declined. We develop a model of firm dynamics that generates these patterns, as well as rich predictions about the unfolding of the cross-section of firms. Our theory focuses on the evolution of a sector rather than on long-run growth of the entire economy.[1]

An industry has a continuum of varieties of a differentiated product and it is populated by a continuum of single-product firms and a finite number of large multi-product firms. While the turnover of single-product firms is very high, the large multi-product firms have long life spans. Large firms lose some products over time, but they can invest in innovation in order to replenish or expand the range of their products. Free entry of the single-product firms, which engage in monopolistic competition, creates a competitive fringe that impacts the oligopolistic competition of the large firms. The interaction between the single- and multi-product firms plays a key role in the dynamics of this industry, both during transition and in steady state.[2]

Our assumptions capture salient features of the data. According to Cao *et al.* (2019), 95% of firms in the U.S. economy are single-establishment firms, but their share in employment is only 45%. Furthermore, Kehrig and Vincent (2019) report that during 1972-2007 an average of 72% of the plants in manufacturing belonged to single-plant firms and 28% belonged to multi-plant firms (see their Table A1). At the same time, single-plant firms manufactured 22% of value added compared to 78% of the value added manufactured by multi-plant firms. Both studies point out that firm growth took place mostly through the extensive margin, by opening new plants that often produced new products. In addition, Hsieh and Rossi-Hansberg (2020) provide evidence that firm growth through the acquisition of new product lines played an important role in the business strategies of U.S. corporations in three major sectors: services, retail and wholesale. Growth of firms in these sectors was driven by expansions to new locations, i.e., new product lines. Finally, studying the size distribution of firms between 1995 and 2014, Cao *et al.* (2019) conclude that the largest contributors to the increase in the number of establishments per firm were declining costs of external innovation and declining exit rates.

The theory developed in this paper predicts that large multi-product firms grow through innovation that expands their product lines, eventually reaching a steady state. In the processes, these firms raise their markups and reduce their labor share. As a result, the average markup—measured

[1]Nevertheless, we developed in the Appendix a model with a continuum of sectors that is suitable for growth analysis. In this model every sector is similar to the sector analyzed in the main text of the paper.

[2]Interactions between a monopolistically competitive fringe of single-product firms and oligopolistic large firms have been studied in a static framework by Shimomura and Thisse (2012) in a closed economy and by Parenti (2018) in an open economy. Cavenaile *et al.* (2019) develop a model of endogenous growth with quality-ladders, in which there is a fringe of competitive small firms that produce a homogenous good. In addition, there are single-product large firms that produce different varieties of a product. Their model is mostly quantitative, used to study the relationship between innovation and competition.

with either cost-share or sales-share weights—rises and the aggregate labor share declines. Both the cost-weighted and the sales-weighted markups rise due to rising individual markups of multi-product firms and the reallocation of market shares from single- to multi-product firms.

The steady state size distribution of firms is driven by heterogeneity of labor productivity, with more productive firms having larger market shares. Nevertheless, this monotonic relationship does not translate into a monotonic relationship between productivity and product span. The reason is that the marginal profitability of investment in innovation is larger when manufacturing costs are lower, but larger market shares reduce the incentives to invent new product lines. This tension can produce an inverted-U relationship between labor productivity and product span in the cross-section of firms. Using the Compustat data for 2018, we provide evidence in support this prediction. We also show that the inverted-U relationship between labor productivity and product span stems from distortions in the market equilibrium. In the optimal dynamic allocation product span is an increasing function of labor productivity in the steady state.

Our model predicts that improvements in the technology of single-product firms, which raise the competitive pressure on the multi-product oligopolies, lead to a decline in the market share of every large firm on impact. Still, the resulting transition dynamics to a new steady state vary across the large firms according to size. In particular, multi-product firms with large market shares compensate for the initial loss of competitiveness (reflected in the loss of market share) by gradually expanding their product span and raising their market share over time, while firms with small market shares further reduce their product span and market shares over time. As a result, the size distribution of multi-product firms becomes more unequal in the new steady state.

We describe some basic elements of the model in the next section. In Section 3 we detail the entry decisions of single-product firms and their impact on the pricing strategy of large firms. These results are then used in Section 4 to study the innovation decisions of large firms and the resulting transition dynamics. We show that whenever multi-product firms widen their product span in the transition, they grow in size and so do their markups, while the labor share declines. In the following section, Section 5 we study comparative dynamics, some of which were described above. In Section 6 we characterize the optimal dynamic allocation and describe a set of policy measures that implement the optimal allocation. Section 7 concludes.

2 Preliminaries

We consider an economy with a continuum of individuals of mass 1. The labor market is competitive and every individual earns the same wage rate.

There are two sectors. One sector produces a homogeneous good with one unit of labor per unit output and there is always positive demand for its product. We normalize the price of this good to equal one. Therefore the competitive wage also equals one. The other sector produces varieties of a differentiated product.[3]

[3]It is straightforward to generalize the analysis to multiple sectors with differentiated products.

Every individual supplies a fixed amount of labor, l, and has a utility function:[4]

$$u = x_0 + \frac{\varepsilon}{\varepsilon - 1}\left[\int_0^N x(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right]^{\frac{(\varepsilon-1)\sigma}{\varepsilon(\sigma-1)}}, \sigma > \varepsilon > 1, \tag{1}$$

where x_0 is consumption of the homogeneous good, $x(\omega)$ is consumption of variety ω of the differentiated product, σ is the elasticity of substitution between varieties of the differentiated product and ε gauges the degree of substitutability between varieties of the differentiated product and the homogeneous good. The assumption $\sigma > \varepsilon$ asserts that brands of the differentiated product are better substitutes for each other than for the homogeneous good. The assumption $\varepsilon > 1$ ensures that aggregate spending on the differentiated product declines when its price rises (see below).

Real consumption of the differentiated product is:

$$X = \left[\int_0^N x(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right]^{\frac{\sigma}{\sigma-1}}. \tag{2}$$

Using this definition, the price index of X is:

$$P = \left[\int_0^N p(\omega)^{1-\sigma} d\omega\right]^{\frac{1}{\sigma-1}},$$

where $p(\omega)$ is the price of variety ω. In this setup an individual chooses consumption to maximize utility subject to the budget constraint $x_0 + PX = l + y$, where y is non-wage income. This yields $X = P^{-\varepsilon}$ as long as consumers purchase the homogenous good and varieties of the differentiated product, which we assume always to be the case (this requires l to be large enough). Clearly, in this case the demand for variety ω is:

$$x(\omega) = P^{\delta} p(\omega)^{-\sigma}, \delta = \sigma - \varepsilon > 0. \tag{3}$$

Aggregate spending on the differentiated product equals $PX = P^{1-\varepsilon}$, which declines in P, because $\varepsilon > 1$. An individual's consumption choice yields the indirect utility function

$$V = l + y + \frac{1}{\varepsilon - 1} P^{1-\varepsilon}, \tag{4}$$

where the second term on the right-hand side represents consumer surplus.

Two types of firms operate in sector X: atomless single-product firms and m large multi-product firms. Every large firm has a positive measure of product lines (recall the discussion in the introduction of evidence in support of this specification). Single-product firms produce a total of $\overline{r} > 0$ varieties, each one specializing in a single brand. Large firm i has $r_i > 0$ product lines, $i = 1, 2, ..., m$. All the brands supplied to the market are distinct from each other.

[4] We can add to this utility function a disutility of effort $\psi \frac{l^{1+\nu}}{1+\nu}$, $\nu, \psi > 0$, as is common in some of the macro literature. Due to the quasi-linearity of the utility function, this would lead every individual to optimally choose a constant supply of labor, $l = \psi^{-1/\nu}$. For this reason we simplify by assuming that l is constant.

All single-product firms share the same technology, which requires $\bar{a}$ unit of labor per unit output.[5] Facing the demand function (3), a single-product firm maximizes profits $P^{\delta}p(\omega)^{-\sigma}\left[p(\omega)-\bar{a}\right]$, taking as given the price index P. Therefore, a single-product firm prices its brand ω according to $p(\omega)=\overline{p}$, where:

$$\overline{p}=\frac{\sigma}{\sigma-1}\bar{a}. \tag{5}$$

This yields the standard markup $\overline{\mu}=\sigma/\left(\sigma-1\right)$ for a monopolistically competitive firm.

A large firm i has a technology that requires a_i units of labor per unit output, and it faces the demand function (3) for each one of its brands. As a result, it prices every brand equally. We denote this price by p_i. The firm chooses p_i to maximize profits $r_iP^{\delta}p_i^{-\sigma}\left(p_i-a_i\right)$. However, unlike a single-product firm, a large firm does not view P as given, because it recognizes that

$$P=\left(\overline{r}\,\overline{p}^{1-\sigma}+\sum_{j=1}^{m}r_jp_j^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, \tag{6}$$

and therefore that its pricing policy has a measurable impact on the price index of the differentiated product. Accounting for this dependence of P on the firm's price, the profit maximizing price is:

$$p_i=\frac{\sigma-\delta s_i}{\sigma-\delta s_i-1}a_i, \tag{7}$$

where s_i is the market share of firm i and:[6]

$$s_i=\frac{r_ip_i^{1-\sigma}}{P^{1-\sigma}}=\frac{r_ip_i^{1-\sigma}}{\overline{r}\,\overline{p}^{1-\sigma}+\sum_{j=1}^{m}r_jp_j^{1-\sigma}}. \tag{8}$$

Equations (7) and (8) jointly determine prices and market shares of large firms. The markup factor of firm i is $\mu_i=(\sigma-\delta s_i)/(\sigma-\delta s_i-1)$, which is increasing in its market share. When the market share equals zero the markup is $\sigma/(\sigma-1)$, the same as the markup of a single product firm. The markup factor varies across firms as a result of differences in either the product span, r_i, or the marginal production cost, a_i. We analyze the dependence of prices, market shares and markups on marginal costs and product spans in the next section.

3 Entry of Single-Product Firms

The number of large firms is given and we analyze in the next section the evolution of their product spans, r_i. Unlike large firms, single-product firms enter the industry until their profits equal zero. Firms in this sector play a two-stage game: in the first stage single-product firms enter; in the second stage all firms play a Bertrand game as described in the previous section. Under these

[5]It is straightforward to allow for heterogeneity of the single-product firms, by assuming that each one of them draws a unit labor requirement from a known distribution. Since this type of heterogeneity plays no essential role in our analysis, we have chosen to work with the simpler formulation.

[6]Note that $\sigma-\delta s_i-1=\sigma\left(1-s_i\right)+\varepsilon s_i-1>0$ and $\sigma-\delta s_i=\sigma\left(1-s_i\right)+\varepsilon s_i>0$.

circumstances, (5) and (7) portray the equilibrium prices, except that the number of single product firms, $\bar{r}$, is endogenous. We seek to characterize a subgame perfect equilibrium of this stage game.

To determine the equilibrium number of single-product firms, assume that they face an entry cost f and they enter until profits equal zero. In a subgame perfect equilibrium every entrant correctly forecasts the number of entrants, and the price that will be charged for every variety in the second stage of the game. Therefore, every single-product firm correctly forecasts the price index P. Using the optimal price (5) and the profit function $P^{\delta}\overline{p}^{-\sigma}\left(\overline{p}-\bar{a}\right)$, this free entry condition can be expressed as:

$$\frac{1}{\sigma}P^{\delta}\left(\frac{\sigma}{\sigma-1}\bar{a}\right)^{1-\sigma}=f. \tag{9}$$

The left-hand side of this equation describes the operating profits, which equal a fraction $1/\sigma$ of revenue, while the right-hand side represents the entry cost. In these circumstances the price index P is determined by f and $\bar{a}$, and it is rising in both f and $\bar{a}$. Importantly, it does not depend on the number of large firms nor on their product spans.

We now use (7) and (8) to calculate the response of prices and market shares to changes in the number of product lines, changes in marginal costs and changes in the price index P. Denoting by a hat the proportional rate of change of a variable, i.e., $\hat{x}=dx/x$, differentiating these two equations yields the solutions:

$$\hat{p}_i=\frac{\beta_i}{1+(\sigma-1)\beta_i}\hat{r}_i+\frac{1}{1+(\sigma-1)\beta_i}\hat{a}_i+\frac{(\sigma-1)\beta_i}{1+(\sigma-1)\beta_i}\hat{P}, \tag{10}$$

$$\hat{s}_i=\frac{1}{1+(\sigma-1)\beta_i}\hat{r}_i-\frac{\sigma-1}{1+(\sigma-1)\beta_i}\hat{a}_i+\frac{\sigma-1}{1+(\sigma-1)\beta_i}\hat{P}, \tag{11}$$

where:

$$\beta_i=\frac{\delta s_i}{(\sigma-\delta s_i-1)(\sigma-\delta s_i)}>0. \tag{12}$$

Due to the fact that the price index P responds neither to changes in r_i nor changes in a_i, an increase in r_i raises p_i and s_i, but it has no impact on prices and market shares of the other large firms. For the same reason, an increase in a_i raises p_i and reduces s_i, but has no impact on prices and market shares of the other large firms. Moreover, an increase in a_i raises the price of firm i less than proportionately, and therefore there is only partial pass-through of marginal costs to prices. The extent of the pass-through is smaller for a firm with a larger β_i, which is a firm with a larger market share. Finally, an increase in the price index P, which represents a decline in the competitive pressure in the industry, raises the price and the market share of *every* large firm. However, the price rises proportionately more and the market share rises proportionately less in firms with larger β_is, which are firms with larger market shares. Finally, the market share of a firm is larger the larger is its product span or the lower is its marginal cost of production. Noting again that the markup of every firm i is larger the larger its market share, we summarize these findings in

Proposition 1. *Suppose that the number of large firms and their product spans are given, but there is free entry of single-product firms. Then: (i) an increase in r_i raises the price, markup and market share of firm i, but has no impact on prices, markups and market shares of the other large firms; (ii) a decline in a_i reduces the price and raises the markup and market share of firm i, but has no impact on prices, markups and market shares of other large firms; (iii) a decline in the price index P, either due to a decline in $\bar{a}$ or a decline in f, reduces the price, markup and market share of every large firm, with prices changing proportionately more and market shares changing proportionately less for firms with initially larger market shares.*

It is clear from this proposition that free entry of single-product firms leads large firms to compete for market share with single-product firms rather than with each other.[7] An increase in r_i or a decline in a_i, each of which raises the market share of firm i, does not impact the market share of other large firms, but do reduce the market share of single-product firms. Since the price index P does not change in response to changes in the number of product lines r_i or the marginal cost a_i, the market share of single-product firms must decline, which materializes through a decline in their joint product span. We, therefore, have

Proposition 2. *Suppose that the number of large firms and their product spans are given, but there is free entry of single-product firms. Then a decline in a_i or an increase in r_i reduces the number of single-product firms and their market share.*

4 Transition Dynamics

We next study the dynamics that arise when large firms can expand their product lines. Time is continuous and the economy starts at time $t = 0$. The range of products of firm i at time t is $r_i(t)$ for $t \geq 0$ and $r_i(0) = r_i^0$ is given.

Similarly to Klette and Kortum (2004), at every point in time firm i can invest to increase the number of its product lines. An investment flow of ι_i per unit time expands r_i by $\phi(\iota_i)$ units per unit time, where $\phi(\cdot)$ is the innovation function. We assume that $\phi(\iota)$ is increasing, concave, $\phi(0) = 0$ and it satisfies the Inada conditions $lim_{\iota \to 0} \phi'(\iota) = +\infty$ and $lim_{\iota \to +\infty} \phi'(\iota) = 0$. Furthermore, r_i depreciates at the rate θ per unit time, which randomly hits the continuum of available brands. It follows that the product span r_i satisfies the differential equation:

$$\dot{r}_i = \phi(\iota_i) - \theta r_i, \text{ for all } t \geq 0, \tag{13}$$

where we have suppressed the time index t in $\dot{r}_i$, r_i and ι_i.

The endogenous expansion of product lines plays an important role in our theory. Hsieh and Rossi-Hansberg (2020) provide evidence that it also played an important role in the business strategies of U.S. corporations in three key sectors—services, retail and wholesale—where firm growth

[7] This is different, for example, from Atkeson and Burstein (2008), who have a continuum of industries, each one populated by a finite number of large single-product firms, and no small firms. Nevertheless, our and their pricing formulas have common elements.

was dominated by expansion to new locations, i.e., new product lines. Cao *et al.* (2019) make a similar argument more broadly; firms grew predominantly on the extensive margin, through new establishments that often represented new product lines.[8]

At every point in time firms play a two stage game. In the first stage single-product firms enter and large firms invest in innovation. Single-product firms live only one instant of time. For this reason they make profits only in this single instant. This assumption represents an extreme form of the empirical property that the turnover of small firm establishments is much larger than the turnover of establishments of large firms. In the second stage all firms choose prices, in the manner described in the previous section. Under the circumstances the price index P is determined by the free entry condition (9), and it remains constant as long as the cost of entry and the cost of production of the single-product firms do no change.

In this economy the state vector is $\mathbf{r} = (r_1, r_2, ..., r_m)$, a function of time t, and the price p_i is a function of $\mathbf{r}$. Note, however, from (10) and (11) that p_i and s_i depend only on element r_i of $\mathbf{r}$. It follows that the profit flow of large firm i is:

$$\pi_i(\iota_i, r_i) = r_i P^{\delta} p_i(r_i)^{-\sigma} [p_i(r_i) - a_i] - \iota_i, \text{ for all } t \geq 0, \tag{14}$$

where P is the same at every t and $p_i(r_i)$ is the price of firm i's brands as a function of r_i, given by (7). Evidently, if the firm's product span changes over time so do π_i, r_i, p_i and ι_i. Moreover, the firm's market share is also a function of r_i, $s_i(r_i)$. From (10) and (11) we obtain the elasticities of the function $p_i(r_i)$:

$$\frac{\partial p_i}{\partial r_i}\frac{r_i}{p_i} = \frac{\beta_i}{1 + (\sigma - 1)\beta_i}, \tag{15}$$

where β_i is defined in (12). Note that β_i is increasing in s_i and that, due to (11), s_i is increasing in r_i. Therefore β_i is increasing in r_i. As a result, the elasticity of the price function is larger the larger is s_i.

Next assume that the interest rate is constant and equal to ρ. This interest rate can be derived from the assumption that individuals discount future utility flows (1) with a constant rate ρ, so that they maximize the discounted present value of utility $\int_0^\infty e^{-\rho t} u(t) dt$. Under these circumstances firm i maximizes the discounted present value of its profits net of investment costs, π_i. It therefore solves the following optimal control problem:

[8] Aghion *et al.* (2019) develop a model of economic growth in which the total number (measure) of product lines is constant, but a single firm can operate multiple product lines. They focus on explaining the decline in the long-run growth rate. The key trigger of their dynamics is a decline in a static cost function $c(n)$ that describes a firm's overhead cost of operating n product lines. They argue that a fall in these costs was caused by the IT revolution. Their firms have constant unit costs and the quality of products can be improved by investing in innovation, as in standard models of endogenous growth with quality ladders (see Grossman and Helpman (1991) and Aghion and Howitt (1992)). A firm acquires new product lines by gaining leadership positions through quality competition. They characterize a steady state of an economy with two types of firms—high- and low-productivity (unit labor requirements)—and study the impact of a decline in $c(n)$ on concentration, labor shares, the reallocation of market shares and the long-run growth rate.

$$\max_{\{\iota_i(t), r_i(t)\}_{t\geq 0}} \int_0^\infty e^{-\rho t} \pi_i \left[\iota_i(t), r_i(t)\right] dt$$

subject to (13), (14), $r_i(0) = r_i^0$, and a transversality condition to be described below, where $\pi_i(\iota_i, r_i)$ is defined in (14). In this problem ι_i is a control variable while r_i is a state variable. The current-value Hamiltonian is:

$$\mathcal{H}(\iota_i, r_i, \lambda_i) = \left\{ r_i P^\delta p_i(r_i)^{-\sigma} \left[p_i(r_i) - a_i\right] - \iota_i \right\} + \lambda_i \left[\phi(\iota_i) - \theta r_i\right],$$

where λ_i is the co-state variable of constraint (13). The first-order conditions of this optimal control problem are:

$$\frac{\partial \mathcal{H}}{\partial \iota_i} = -1 + \lambda_i \phi'(\iota_i) = 0,$$

$$-\frac{\partial \mathcal{H}}{\partial r_i} = -\frac{\partial \left\{ r_i P^\delta p_i(r_i)^{-\sigma} \left[p_i(r_i) - a_i\right] \right\}}{\partial r_i} + \theta \lambda_i = \dot{\lambda}_i - \rho \lambda_i,$$

and the transversality condition is:

$$\lim_{t\to\infty} e^{-\rho t} \lambda_i(t) r_i(t) = 0.$$

In addition, the optimal path of (ι_i, r_i) has to satisfy the differential equation (13).

The above first-order conditions can be expressed as:

$$\lambda_i \phi'(\iota_i) = 1, \tag{16}$$

$$\dot{\lambda}_i = (\rho + \theta) \lambda_i - P^\delta p_i(r_i)^{-\sigma} \left\{ p_i(r_i) - a_i - r_i \left(\sigma p_i(r_i)^{-1} \left[p_i(r_i) - a_i\right] - 1 \right) p_i'(r_i) \right\}. \tag{17}$$

From (16) we obtain the investment level ι_i as an increasing function of λ_i, which we represent as $\iota_i(\lambda_i)$. Substituting this function into (13) yields the autonomous differential equation:

$$\dot{r}_i = \phi\left[\iota_i(\lambda_i)\right] - \theta r_i. \tag{18}$$

Next we substitute (7), (12) and (15) into (17) to obtain a second autonomous differential equation:

$$\dot{\lambda}_i = (\rho + \theta) \lambda_i - \Gamma_i(r_i), \tag{19}$$

where

$$\Gamma_i(r_i) = a_i^{1-\sigma} P^\delta \sigma \left[\frac{\sigma - \delta s_i(r_i)}{\sigma - \delta s_i(r_i) - 1}\right]^{-\sigma} \frac{1}{\left[\sigma - \delta s_i(r_i) - 1\right] \sigma + s_i(r_i)^2 \delta^2} \tag{20}$$

represents the profitability of a new product line, given the firm's product span r_i; that is, it represents the *marginal* profitability of r_i. We show in the Appendix that this marginal profitability

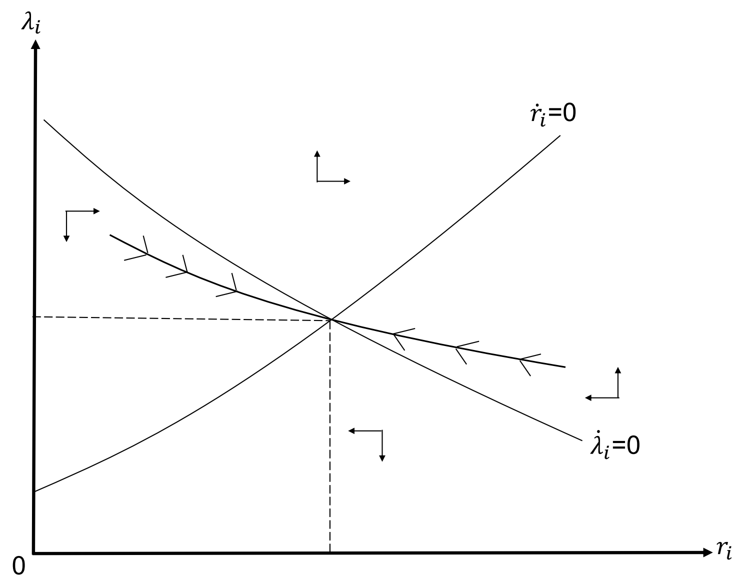

Figure 1: Transition Dynamics

declines in r_i, i.e., $\Gamma_i'(r_i) < 0$.[9]

A solution to the autonomous system of differential equations (18) and (19) that satisfies the transversality condition is also a solution to the firm's optimal control problem, because $\mathcal{H}(\iota_i, r_i, \lambda_i)$ is concave in the first two arguments. This can be seen by observing that the Hamiltonian is additively separable in ι_i and r_i, and it is strictly concave in ι_i and in r_i. The steady state of these differential equations is characterized by:

$$\phi\left[\iota_i\left(\lambda_i\right)\right] = \theta r_i, \tag{21}$$

$$\left(\rho + \theta\right)\lambda_i = \Gamma_i\left(r_i\right). \tag{22}$$

The left-hand side of (21) is an increasing function of λ_i. Therefore, the curve in (r_i, λ_i) space along which r_i is constant is upward sloping. The right-hand side of (22) is declining in r_i, because $\Gamma_i'(r_i) < 0$. Therefore, the curve in (r_i, λ_i) space along which λ_i is constant is downward sloping. These curves are depicted in Figure 1. Based on the differential equations (18)-(19), the figure also depicts the resulting dynamics. There is a single stable saddle-path along which (r_i, λ_i) converge to the steady state and the transversality condition is satisfied in this steady state. On this saddle path either r_i rises and λ_i declines or r_i declines and λ_i rises, depending on whether r_i^0 is below or above its steady-state value.

[9]By definition:

$$\Gamma_i\left(r_i\right) \equiv \frac{\partial\left\{r_i P^{\delta} p_i\left(r_i\right)^{-\sigma}\left[p_i\left(r_i\right) - a_i\right]\right\}}{\partial r_i},$$

where the right-hand side represents marginal profits of r_i. We show in the Appendix that $\Gamma_i\left(r_i\right)$ can be expressed as $\tilde{\Gamma}_i\left[s_i\left(r_i\right)\right]$, where $\tilde{\Gamma}_i\left(\mathrm{s}_i\right)$ is a declining function (see (20) above and equation (40) in the Appendix). Since (11) implies that $s_i\left(r_i\right)$ is an increasing function, it follows that $\Gamma_i\left(r_i\right)$ is a declining function.

Now suppose that all the r_i^0s are below their steady state values (this case arises, for example, when the economy is in steady state and innovation costs decline; see below). Then every large firm expands its range of products over time. As a result, the number of single-product firms shrinks. This process continues until the economy reaches a steady state.

If at some point in time the number of single-product firms drops to zero, the dynamics change.[10] We focus, however, on the case in which $\bar{r} > 0$ for all $t \geq 0$. In this case the price index P remains constant as long as f and $\bar{a}$ do not change.

What can be said about the dynamics of profits net of investment costs? Changes of these profits over time can be expressed as (see (14)):

$$\begin{aligned}\frac{\partial \pi_i(\iota_i, r_i)}{\partial t} &= -\frac{\partial \iota_i}{\partial t} + \frac{\partial \left\{ r_i P^\delta p_i(r_i)^{-\sigma} \left[p_i(r_i) - a_i \right] \right\}}{\partial r_i} \frac{\partial r_i}{\partial t} \\ &= -\frac{\partial \iota_i}{\partial t} + \Gamma_i(r_i) \frac{\partial r_i}{\partial t}.\end{aligned}$$

From (16) we see that ι_i is an increasing function of λ_i and λ_i declines in a firm that expands its product range. As a result, the firm's investment level ι_i declines over time, raising profits net of investment costs through a decline in investment spending. Moreover, $\Gamma_i(r_i) > 0$, and therefore an increase in r_i raises operating profits, thereby raising profits net of investment costs.[11] It follows that every firm that adds new product lines enjoys rising profits net of investment costs. Since wages are constant, this implies that the share of labor in national income declines when all large firms grow.[12]

Next consider the average markup in the differentiated product sector, defined as aggregate revenue divided by aggregate variable costs:

$$\mu_{av} = \frac{\bar{r}\,\bar{p}^{1-\sigma} + \sum_{j=1}^m r_j p_j^{1-\sigma}}{\bar{r}\,\bar{a}\,\bar{p}^{-\sigma} + \sum_{j=1}^m r_j a_j p_j^{-\sigma}}.$$

[10] From that point on the optimal strategy of large firm i depends on the entire state vector $\mathbf{r}$. As a result, the firms engage in a differential game. Since no firm can commit to the entire path of its investments ι_i, one needs to adopt the closed loop solution to this game, in which the investment level ι_i is a function of the state vector $\mathbf{r}$. There do not exist user-friendly characterizations of solutions to such games. Instead, we provide in the Appendix an analysis of the impact of changes in the state variables r_i on prices, markups and market shares of the large firms.

[11] Intuitively, an increase in a product line raises directly a firm's profits due to the fact that the price exceeds marginal costs. But it also reduces profits indirectly as a result of an increase in the price in response to the expansion of product span. The former effect is represented by $p_i - a_i > 0$ in the tilted brackets of (17) while the latter effect is represented by $-r_i\left[\sigma p_i^{-1}(p_i - a_i) - 1\right] p_i' < 0$ in these brackets, noting that $p_i' > 0$ and $-\left[\sigma p_i^{-1}(p_i - a_i) - 1\right] = 1 - \sigma\left(1 - \mu_i^{-1}\right) < 0$. The last inequality results from the fact that the markup, $\mu_i = p_i/a_i$, is larger than $\sigma/(\sigma - 1)$ (see (7)). Nevertheless, the joint impact is positive, as shown in (20).

[12] In this economy aggregate income equals wages plus profits net of investment costs, i.e.,

$$y_{ag} = l + \sum_{j=1}^m \pi_i(\iota_i, r_i),$$

and the labor share is l/y_{ag}. When all large multi-product firms have r_i^0s below their steady state values they raise their product span over time and aggregate profits increase. As a result, the aggregate labor share declines.

This statistic can be expressed as

$$\mu_{av} = \left(1 - \sum_{i=1}^{m} \varrho_i\right) \bar{\mu} + \sum_{i=1}^{m} \varrho_i \mu_i,$$

where

$$\varrho_i = \frac{r_i a_i p_i^{-\sigma}}{\overline{r}\,\overline{a}\,\overline{p}^{-\sigma} + \sum_{j=1}^{m} r_j a_j p_j^{-\sigma}} = \mu_{av} s_i \mu_i^{-1} \tag{23}$$

is the variable cost share of large firm i, $\bar{\mu} = \bar{p}/\bar{a}$ is the markup of a small firm and $\mu_i = p_i/a_i$ is the markup of large firm i. In other words, the average markup μ_{av} is a cost-weighted average of the markups of single- and multi-product firms. Next note from (23) that $\sum_{i=1}^{m} \varrho_i \mu_i = \mu_{av} \sum_{i=1}^{m} s_i$, and therefore:

$$\mu_{av} = \left(1 - \mu_{av} \sum_{i=1}^{m} s_i \mu_i^{-1}\right) \bar{\mu} + \mu_{av} \sum_{i=1}^{m} s_i,$$

which implies:

$$\mu_{av} = \frac{1}{\left(1 - \sum_{i=1}^{m} s_i\right) \bar{\mu}^{-1} + \sum_{i=1}^{m} s_i \mu_i^{-1}}. \tag{24}$$

When all large firms grow on the dynamic path, the market share of every one of them rises and so does its markup. The hike in each firm's market share and markup consequently contribute to a rise in the average markup, μ_{av}, because the large firms have larger markups than the single-product firms.

While the cost-weighted average markup represents the ratio of aggregate revenue to aggregate variable cost, an alternative measure of average markups is a sales-share weighted average of the markups of all single- and multi-product firms (see, for example, Edmond *et al.* (2019)). In our model this average is:

$$\mu_{av}^{s} = \left(1 - \sum_{i=1}^{m} s_i\right) \bar{\mu} + \sum_{i=1}^{m} s_i \mu_i.$$

Since the markup of every large firm is higher than the markup of every single-product firm and the market share of every multi-product firm rises over time, this average markup increases over time. The growth in this average markup is driven by the same two forces that drive the rise in the cost-weighted average markup μ_{av}: rising markups of the large firms and market share reallocation from low-markup (single-product) to high-markup (multi-product) firms. We conclude that both measures of the average markup, μ_{av} and μ_{av}^{s}, are rising over time when large firms expand their product span.

Next compare the size of these markup statistics. Their ratio is given by:

$$\frac{\mu_{av}^{s}}{\mu_{av}} = \left[\left(1 - \sum_{i=1}^{m} s_i\right) \bar{\mu} + \sum_{i=1}^{m} s_i \mu_i\right] \left[\left(1 - \sum_{i=1}^{m} s_i\right) \bar{\mu}^{-1} + \sum_{i=1}^{m} s_i \mu_i^{-1}\right].$$

Since $1/\mu$ is a convex function of μ, Jensen's inequality implies that the right-hand side of this equation is larger than one, and therefore that $\mu_{av}^{s} > \mu_{av}$, which is what Edmond *et al.* (2019)

found in the Compustat data.[13] We summarize these findings in

Proposition 3. *Consider an economy in which the initial range of products* r_i^0 *is smaller than its steady state value for every* i*, and in which* $\bar{r} > 0$ *at all times. Then over time: (i) every large firm* i *widens its product span, raises its markup, and experiences rising profits net of investment costs; (ii) the cost-weighted average markup and the sales-weighted average markup rise over time; (iii) the sales-average markup exceeds the cost-average markup at every point in time; and (iv) the share of labor in national income declines over time.*

Since wages are constant, so is wage income. Nonetheless, in view of Proposition 3(i), aggregate income—which consists of labor income plus aggregate profits net of investment costs—rises during the transition to a steady state. In view of the indirect utility function (4), this implies that aggregate utility rises over time (recall that P remains constant). Moreover, if this economy is populated by some individuals who own shares in large firms and other individuals who do not, the growth of large multi-product firms widens the disparity of well-being between these two groups.

5 Comparative Dynamics

For an economy that is initially in steady state, we study in this section the dynamics that arise in response to changes in the cost of inventing new product lines, the marginal costs of production and the cost of entry of single-product firms.

First, consider a change in the cost of innovation, as reflected in a shift of the function $\phi(\iota_i)$. We take $\kappa > 0$ to be a productivity measure of innovation and express the modified innovation function as $\kappa\phi(\iota_i)$. Initially $\kappa = 1$. An upward shift in κ represents a rise in the productivity of investment in innovation or a decline in innovation costs, while a decline in κ represents a decline in the productivity of investment in innovation or a rise in innovation costs. The latter may arise when it becomes harder to invent new product lines. With the new innovation function the dynamics of product span, (13), become:

$$\dot{r}_i = \kappa\phi(\iota_i) - \theta r_i, \text{ for all } t \geq 0. \tag{25}$$

In this case the first-order condition of the optimal control problem (16) becomes:

$$\lambda_i \kappa \phi'(\iota_i) = 1, \tag{26}$$

while the differential equation (19) does not change. From (26) we obtain the investment level ι_i as an increasing function of $\kappa\lambda_i$, which we express as $\iota_i(\kappa\lambda_i)$. This is the same $\iota_i(\cdot)$ function that we

[13]Using a dynamic model of monopolistic competition with a Kimball aggregator, Edmond *et al.* (2019) decompose the welfare cost of markups into three sources of influence: (i) aggregate markup; (ii) misallocation of inputs; and (iii) inefficiently low entry of firms. Their quantitative model implies that (i) accounts for 3/4 of the welfare cost while (ii) accounts for 1/4. The impact of entry is negligible. They also show that in the Compustat data the sales-weighted aggregate markup is higher than the cost-weighted aggregate markup, in line with our theoretical prediction, and that the gap between them has widened over time (see their Figure 8). Moreover, in their model the cost-weighted aggregate markup turns out to be the relevant measure for (i).

had before. Substituting this function into (25) yields the autonomous differential equation:

$$\dot{r}_i = \kappa\phi\left[\iota_i\left(\kappa\lambda_i\right)\right] - \theta r_i.$$

The steady state of this differential equations is characterized by:

$$\kappa\phi\left[\iota_i\left(\kappa\lambda_i\right)\right] = \theta r_i,$$

while the second steady state equation, (22), does not change, because the differential equation (19) remains the same. For $\kappa = 1$, the steady state and the dynamics depicted in Figure 1 remain the same.

Now consider an increase in κ, representing a decline in the costs of inventing new product lines. Since $\kappa\phi\left[\iota_i\left(\kappa\lambda_i\right)\right]$ is increasing in κ, this leads to a downward shift of the $\dot{r}_i = 0$ curve without changing the $\dot{\lambda}_i = 0$ curve. As a result, λ_i declines on impact to a new saddle path, starting transition dynamics with declining values of λ_i and rising values of r_i. This process takes place in every large firm, leading to a new steady state in which every large firm has a larger product span, a larger market share and a higher markup. The average markups μ_{av} and μ^s_{av} rise during the transition and they are higher in the new steady state. The flow of aggregate utility also rises during this transition and is higher in the new steady state. The flow utility rises because profits net of investment costs rise while the price index P remains the same. We therefore have

Proposition 4. *Suppose that every large firm i is in steady state and $\bar{r} > 0$ at all times. Then a decline in the cost of innovation, i.e., an increase in κ, leads all large firms to expand product ranges, raise their market shares and raise their markups. Contemporaneously, the average markups μ_{av} and μ^s_{av} increase and so does the aggregate flow of utility.*

We next turn to changes in the marginal costs of production and the cost of entry of single-product firms. As is evident from (21) and (22), such changes impact the new steady state through the function $\Gamma_i\left(r_i\right)$ only. A change that raises $\Gamma_i\left(r_i\right)$ shifts upward the $\dot{\lambda}_i = 0$ curve in Figure 1. After the impact effect, which results from the upward jump in λ_i, the dynamic process leads to a gradual widening of the span of products and increases in the markup and profits net of investment costs. In contrast, a change that reduces $\Gamma_i\left(r_i\right)$ shifts downward the $\dot{\lambda}_i = 0$ curve. After the downward jump of λ_i on impact, the dynamic process then leads to a gradual narrowing of the span of products and declines in markups and profits net of investment costs.

First, consider a decline in a_i, resulting from a technical improvement in the firm's technology. We show in the Appendix that the impact of a_i on $\mathit{\Gamma}_i$ can be expressed as:

$$\begin{aligned} \hat{\mathit{\Gamma}}_i &= -\left(\sigma - 1\right)\hat{a}_i + \left(\frac{\partial \mathit{\Gamma}_i}{\partial s_i}\frac{s_i}{\mathit{\Gamma}_i}\right)\left(\frac{\partial s_i}{\partial a_i}\frac{a_i}{s_i}\right)\hat{a}_i \\ &= \frac{\left(\sigma - 1\right)s_i^2\delta^2 - \left(\sigma - \delta s_i - 1\right)^2\left(\sigma^2 - \delta^2 s_i^2\right)}{\left[\left(\sigma - \delta s_i - 1\right)\sigma + s_i^2\delta^2\right]^2}\left(\sigma - 1\right)\hat{a}_i. \end{aligned} \tag{27}$$

The relationship between a_i and $\mathit{\Gamma}_i$ portrayed by this equation does not depend on the cost structure of other firms. Moreover, it implies that a decline in a_i shifts upward the $\dot{\lambda}_i = 0$ curve if and only if:

$$(\sigma - \delta s_i - 1)^2 \left(\sigma^2 - \delta^2 s_i^2\right) > (\sigma - 1)\, s_i^2 \delta^2. \tag{28}$$

The potential ambiguity of the response of $\mathit{\Gamma}_i$ to changes in a_i results from the existence of two channels through which the marginal cost impacts the profitability of a new variety (the marginal profitability of r_i), as can be seen from (20). A decline in a_i raises the marginal profitability of a new variety, $\mathit{\Gamma}_i$, for a given market share, s_i, due to cost savings in production. But, a decline in a_i makes firm i more competitive, thereby raising its market share, as shown in (11). A rise in the firm's market share reduces in turn the profitability of a new variety, as we show formally in the Appendix (see (40) in the Appendix). It follows that the shift of the $\dot{\lambda}_i = 0$ curve depends on the strength of these two effects: if the response of the market share dominates, the curve shifts down; and if the response of the market share does not dominate, the curve shifts up. The strength of the market share effect depends in turn on the firm's initial size. For low values of s_i the impact through the market share channel is small, and (28) is satisfied. But (28) is less likely to be satisfied the larger s_i is, because the left-hand side of this inequality is declining in s_i while the right-hand side is increasing. This leads to the following

Lemma 1. *If $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) > (\sigma - 1)\,\delta^2$, then (28) is satisfied for all market shares $s_i \epsilon\, [0, 1]$. And if $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) < (\sigma - 1)\,\delta^2$, then there exists a market share $s^o \epsilon\, (0, 1)$, defined by:*

$$(\sigma - \delta s^o - 1)^2 \left[\sigma^2 - \delta^2 (s^o)^2\right] = (\sigma - 1)\,(s^o)^2\,\delta^2,$$

such that (28) is satisfied for $s_i < s^o$ and violated for $s_i > s^o$.

Given the assumption $\sigma > \varepsilon > 1$, the inequality $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) > (\sigma - 1)\,\delta^2$ is satisfied when ε is close to σ and violated when ε is close to one (recall that $\delta = \sigma - \varepsilon$). We therefore have

Proposition 5. *Suppose that firm i is in steady state and $\bar{r} > 0$ at all times. Then a decline in a_i triggers an adjustment process that gradually raises r_i as well as i's markup and profits net of investment costs if either $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) > (\sigma - 1)\,\delta^2$ or $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) < (\sigma - 1)\,\delta^2$ and $s_i < s^o$, where s^o is defined in Lemma 1. Otherwise, this technical improvement triggers an adjustment process that gradually reduces r_i while i's markup and profits net of investment costs decline gradually after increasing on impact.*

Using these results, we can examine the dynamics of firm i's market share. Since on impact the span of products does not change (r_i is a state variable), (11) implies that the decline in the marginal cost raises firm i's market share on impact. Moreover, if the adjustment process leads to a gradual expansion of its product span, i's market share rises over time until it reaches a new steady state. In this case the firm has a larger market share in the new steady state. If, however, the adjustment process leads to a narrowing of the firm's product span, then (11) implies that the

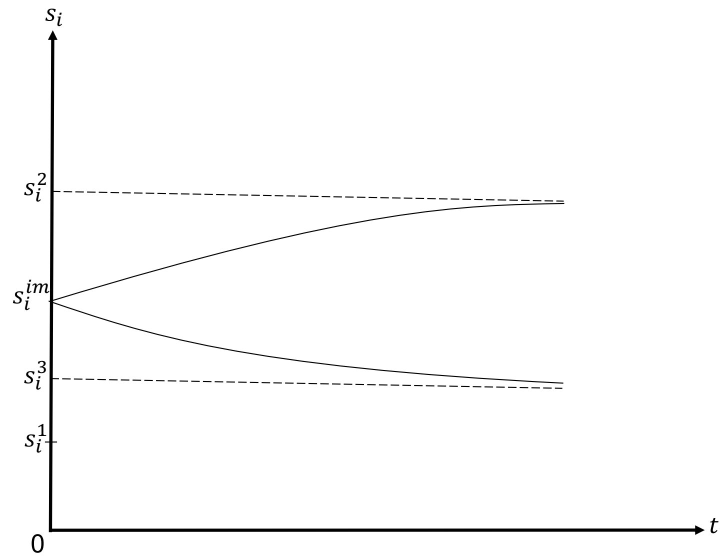

Figure 2: Dynamics of the market share in response to a decline in the marginal cost a_i

initial upward jump in firm i's market share is followed by a gradual decline in its market share. A question then arises whether this firm's market share is larger or smaller in the new steady state. We prove the following

Proposition 6. *Suppose that firm i is in steady state and $\bar{r} > 0$ at all times. Then a decline in a_i triggers an adjustment process that raises s_i in the new steady state.*

Proof. We have shown that the market share is larger in the new steady state when the adjustment process involves expansion of the firm's product span. It therefore remains to show that this is also true when the adjustment process involves contraction of the product span. To this end note that a decline in r_i on the transition path is triggered by a decline in the marginal profitability of r_i in response to a decline in a_i, which leads in turn to a downward shift in the $\dot{\lambda}_i = 0$ curve in Figure 1. In this case the new steady state has a lower r_i as well as a lower λ_i. Next note from the steady state condition (22) that a lower λ_i implies a lower $\mathit{\Gamma}_i$. Recall, however, that for a constant s_i a fall in a_i raises $\mathit{\Gamma}_i$, and therefore $\mathit{\Gamma}_i$ can be lower in the new steady state only if s_i is higher. In sum, independently of whether a decline in a_i shifts upward or downward the $\dot{\lambda}_i = 0$ curve, the market share s_i is larger in the new steady state. □

This result yields the following

Corollary 1. *Consider an economy in steady state with active single-product firms. Then large firms with lower marginal costs have larger market shares.*

The dynamic patterns of the market share that have been unveiled by this analysis are depicted in Figure 2, where s_i^1 is the market share in the initial steady state. First, the market share jumps up to s_i^{im} on impact when a_i declines. Afterward, the market share rises continuously until it reaches

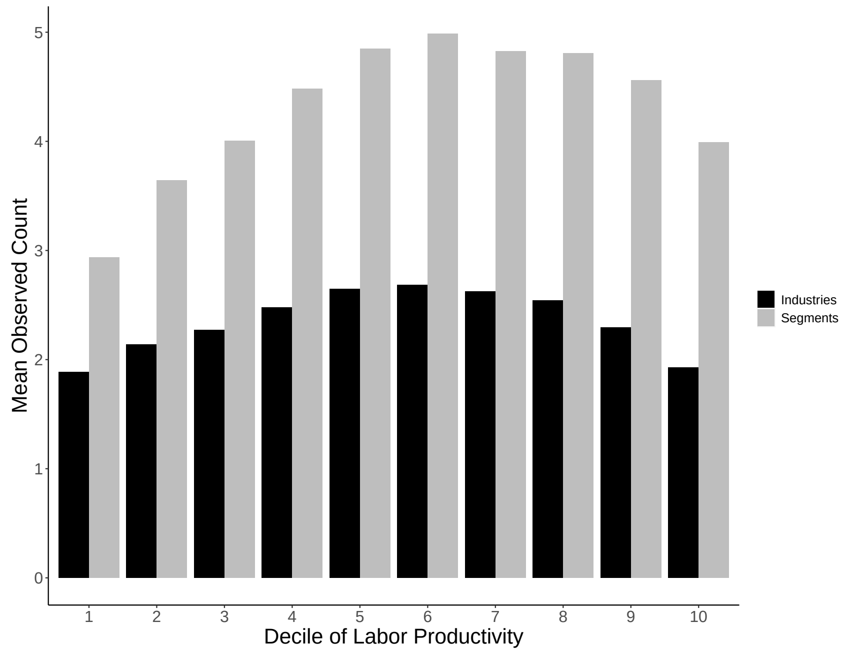

Figure 3: Average Number of Product Lines vs. Labor Productivity Deciles

s_i^2, as portrayed by the upper curve, or it declines continuously until it reaches s_i^3, as portrayed by the lower curve. In both cases the new steady state market share exceeds s_i^1. The former case applies when $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) > (\sigma - 1)\,\delta^2$ or $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) < (\sigma - 1)\,\delta^2$ and $s_i^1 < s^o$, and the latter case applies otherwise.

These results suggest three possible steady state patterns for the relationship between a_i and r_i in the cross section of multi-product firms: lower-cost firms have larger product spans, lower-cost firms have smaller product spans, or the relationship between marginal costs and product spans has an inverted U shape. The first pattern holds for all marginal cost structures when $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) > (\sigma - 1)\,\delta^2$. In the opposite case, when $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) < (\sigma - 1)\,\delta^2$, there exist high values of a_i at which $s_i < s^o$, and among firms with marginal costs in this range those with lower marginal costs have larger product spans. Moreover, there exist low values of a_i at which $s_i > s^o$, and among firms with such low marginal costs lower-cost firms have smaller product spans. Combining these results we have

Proposition 7. *Consider an economy in steady state with active single-product firms. Then, in the cross section of multi-product firms r_i is declining in s_i, rising in s_i, or rising in s_i among firms with low market shares and declining in s_i among firms with high market shares.*

Combining this Proposition with Corollary 1, we note that our model raises the possibility of an inverted-U relationship between labor productivity, as measured by $1/a_i$, and the number of product lines, r_i. We now show that this prediction is not only a theoretical possibility, but that there is suggestive evidence for such a relationship in the Compustat data set. To this end we collected data on revenue, employment, the number of sectors in which a firm operated and the number of segments in which a firm operated, all for 2018. We computed labor productivity as revenue per worker and we treat the number of segments as a proxy for the number of product lines. As a robustness check, we also consider the number of industries in which a firm operated as a proxy for the number of its product lines.[14] Figure 3 depicts the relationships between our two proxies for r_i and our proxy for $1/a_i$. On the horizontal axis the firms are divided into deciles, based on their labor productivity. On the vertical axis we report the mean number of segments and the mean number of industries in each decile. As is evident, these relationships exhibit an inverted-U.

To further examine these relationships, we regressed the number of segments or the number of sectors in which a firm operates on a second-order polynomial of the log of labor productivity. We report in the Appendix the resulting OLS estimates. The coefficient on log labor productivity is positive and the coefficient on the log of labor productivity squared is negative in both case. Moreover, all four coefficients are significantly different from zero. Figure 4 plots the data points that we have used (more than 4,000 observations) as well as the fitted quadratic curve. The first thing to note is that there are many firms with similar numbers of segments and different labor productivity levels, especially when the number of segments is low. Nevertheless, the estimated curve has the shape of an inverted-U. We report in the Appendix a similar graph for the number of industries in which a company operates. In conclusion, while we view this paper as a theoretical contribution, we have also provided suggestive evidence for the inverted-U curve predicted by our model.

5.1 Costs of Single-Product Firms

We next examine the impact of the cost structure of single-product firms. As is evident from (9), a decline in either the marginal cost or the entry cost of single-product firms reduces the price index P, thereby raising the competitive pressure in the economy. How do the large firms respond to this rise in competition? To answer the question, suppose that all firms are in steady state. Equation (20) implies:

$$\hat{\Gamma}_i = \delta \hat{P} + \left(\frac{\partial \Gamma_i}{\partial s_i} \frac{s_i}{\Gamma_i} \right) \left(\frac{\partial s_i}{\partial P} \frac{P}{s_i} \right) \hat{P}. \tag{29}$$

A decline in the price index P elevates the competitive pressure on every large firm and reduces the marginal value of its product span, r_i. Accordingly, the first term on the right-hand side of this equation is negative when $\hat{P} < 0$. In response, firm i reduces its price and market share (see (10) and (11)) and the fall in market share raises the marginal value of r_i. For this reason the second term on the right-hand side is positive when the price index declines. It follows that a decline in P shifts the $\dot{\lambda}_i = 0$ curve downward in Figure 1 if the competition effect dominates and upward if

[14]About 70% of the firms in the Compustat database breakdown the company into segments through Compustat Segments Data. Firms are able to distinguish between business segments, geographic segments, operating segments, state segments. This data is self-reported and thus is not standardized, but is still widely used. We focus on the number of business segments a company lists as a proxy for the number of product lines. Within each business segment the firm can list up to two SIC codes in which the business segment operates. The total number of unique SIC codes listed across business segments is what we define as the number of industries in which a firm operates. This is our second proxy for the number of product lines.

Figure 4: Number of Segments vs. Labor Productivity

the market share effect dominates. Using (11), it is evident that for $\varepsilon \to 1$ (29) is similar to (27), except for the opposite sign on their right-hand sides. Therefore, in this case a decline in P shifts down the $\dot{\lambda}_i = 0$ curve if and only if a decline in a_i shifts it up. Under these conditions a lower P may lead to a lower or higher value of r_i in steady state, and moreover, its impact may vary across firms with different marginal costs and therefore different market shares s_i. For $\varepsilon \to 1$ the inequality $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) > (\sigma - 1)\,\delta^2$ is violated, implying that there exists an s_P^o, such that the decline in P shifts the $\dot{\lambda}_i = 0$ curve down for $s_i < s_P^o$ and up for $s_i > s_P^o$. In this case a rise in the competitive pressure shrinks the product span of multi-product firms with $s_i < s_P^o$ and expands the product span of multi-product firms with $s_i > s_P^o$. As a result, the gaps in market shares between large and small multi-product firms widens, thereby increasing the inequality in the size distribution of firms.[15] Alternatively, for $\varepsilon \to \sigma > 1$ the inequality $(\sigma - \delta - 1)^2 \left(\sigma^2 - \delta^2\right) > (\sigma - 1)\,\delta^2$ is always satisfied, implying that for ε close to σ the competition effect dominates the market share effect. Consequently, the $\dot{\lambda}_i = 0$ curve shifts down for all multi-product firms, decreasing their product span, when the price index decreases.

Finally, note that a decline in P reduces the steady state market share of every large firm. This is clearly the case when every firm's product span declines, because in this case declines in both P and r_i diminish the market share (see (11)). Alternatively, for a firm that expands its steady state r_i, the value of λ_i is higher in the new steady state (see (21)). Therefore, this firm's Γ_i is also larger in the new steady state (see (22)). But the direct impact of the decline in P on Γ_i is negative, and therefore s_i has to be smaller for Γ_i to be larger. We, therefore, have

[15] From (11), $\hat{s}_i - \hat{s}_j = (\hat{r}_i - \hat{r}_j) / [1 + (\sigma - 1)\,\beta_i]$. Therefore $\hat{s}_i > \hat{s}_j$ if and only if $\hat{r}_i > \hat{r}_j$.

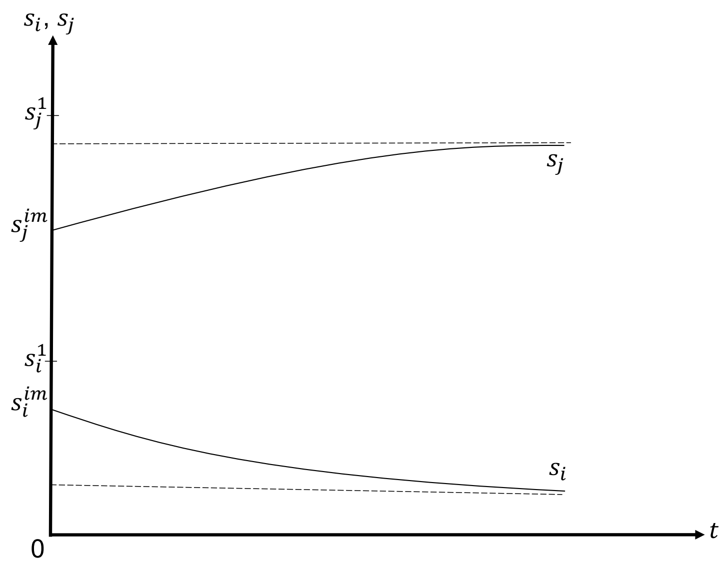

Figure 5: Dynamics of market shares in response to a decline in P

Proposition 8. *Consider an economy in steady state with $\bar{r} > 0$ at all times. Then, a technical improvement that reduces either f or $\bar{a}$ may raise r_i in the new steady state for all i, reduce r_i for all i, or reduce r_i of the small multi-product firms and raise r_i of the large multi-product firms. Nevertheless, s_i is smaller in the new steady state for all large firms i.*

Figure 5 depicts the dynamics of two firms, i and j, for the case in which $s_i < s_P^o$ and $s_j > s_P^o$, where s_P^o is the cutoff market share for the opposite firm dynamics. Firm i starts with $s_i = s_i^1$ while firm j starts with $s_j = s_j^1$. In both firms the market share jumps down on impact as a result of the decline in P, to s_i^{im} and s_j^{im}, respectively. After that, the market share of the smaller firm declines while the market share of the larger firm rises. Yet in both cases, the market share is lower in the new steady state.

Gutierrez and Philippon (2019) find that the elasticity of the number of firms with respect to Tobin's Q declined during 1995-2010. They argue that this resulted from increased entry costs due to regulation rather than due to technological developments or financial frictions. In our model an increase in f generates the above described dynamics independently of the source of variation in the fixed cost of entry. According to Proposition 8, an increase in f raises the long-run market share of all large multi-product firms and reduces the joint market share of the small single product firms. Yet, it may have an uneven impact on the span of products of the large firms. That is, it may increase the number of product lines of the smaller multi-product firms and reduce the number of product lines of the large ones, thereby flattening the relationship between labor productivity (i.e., $1/a_i$) and product span.

6 Optimal Allocation

We study in this section the optimal allocation and discuss policies that support it in a decentralized equilibrium. Recall that the interest rate is constant and equal to ρ. Therefore the optimal allocation is obtained by maximizing the present value of the utility flows $\int_0^\infty e^{-\rho t}u(t)dt$, where u is given by (1). We characterize the optimal allocation in two stages. First, we solve a static optimal allocation for every point in time. Then, in stage two, we solve a dynamic optimal allocation problem that uses the static solution at each point in time. This two-stage procedure provides useful intuition and makes it easier to characterize optimal policies in a market economy.

At a point in time t consumption of the homogeneous good is equal to

$$x_0(t) = l - \left\{ \bar{r}(t)\left[f + \bar{a}\bar{x}(t)\right] + \sum_{i=1}^{m} \left[\iota_i(t) + r_i(t)\, a_i x_i(t)\right] \right\}.$$

The term in the curly brackets represents labor used in the production of varieties of the differentiated product by single- and multi-product firms, plus the entry resource cost of single-product firms, plus investment in multi-product firms for the expansion of their product spans, where all variables have the same meaning as in the previous section. Using the definition of the real consumption index X in (2),

$$X(t) = \left[\bar{r}(t)\,\bar{x}(t)^{\frac{\sigma-1}{\sigma}} + \sum_{i=1}^{m} r_i(t)\, x_i(t)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}. \tag{30}$$

Substituting these equations into (1) we obtain the flow of utility at time t

$$u(t) = l + \frac{\varepsilon}{\varepsilon-1} X(t)^{\frac{\varepsilon-1}{\varepsilon}} - \left[\bar{r}(t)\,\bar{a}\bar{x}(t) + \sum_{i=1}^{m} r_i(t)\, a_i x_i(t) \right] - \bar{r}(t) f - \sum_{i=1}^{m} \iota_i(t).$$

The term in the square brackets on the right-hand side of this equation represents variable labor costs of producing the real consumption $X(t)$. To achieve optimality these labor costs have to be minimized subject to (30), yielding the cost function $C\left[\bar{r}(t), \{r_i(t)\}_{i=1}^m\right] X(t)$, where

$$C\left[\bar{r}(t), \{r_i(t)\}_{i=1}^m\right] = \left[\bar{r}(t)\,\bar{a}^{1-\sigma} + \sum_{j=1}^{m} r_j(t)\, a_j^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{31}$$

is the unit labor cost of $X(t)$. Substituting this result into the utility flow yields:

$$u(t) = l + \frac{\varepsilon}{\varepsilon-1} X(t)^{\frac{\varepsilon-1}{\varepsilon}} - C\left[\bar{r}(t), \{r_i(t)\}_{i=1}^m\right] X(t) - \bar{r}(t) f - \sum_{i=1}^{m} \iota_i(t).$$

Now note that optimality requires to choose $X(t)$ so as to maximize this utility flow, which yields

$X(t) = C[\bar{r}(t), \{r_i(t)\}_{i=1}^m]^{-\varepsilon}$ and the flow of utility

$$u(t) = l + \frac{1}{\varepsilon - 1} C[\bar{r}(t), \{r_i(t)\}_{i=1}^m]^{1-\varepsilon} - \bar{r}(t) f - \sum_{i=1}^{m} \iota_i(t). \tag{32}$$

Next observe that $\bar{r}(t)$ has to be chosen so as to maximize (32) subject to (31), yielding the first-order condition

$$\frac{1}{1-\sigma} C[\bar{r}(t), \{r_i(t)\}_{i=1}^m]^{\delta} \bar{a}^{1-\sigma} = f, \;\; \delta = \sigma - \varepsilon > 0.$$

We expressed this as an equality condition, assuming that in every period the left-hand side of this equation is larger than the right-hand side for $\bar{r}(t) = 0$; otherwise, it is not desirable to have active single-product firms. For example, when the entry cost of single-product firms, f, is very high, it maybe optimal to forgo their services. For comparison with our market results, where we assumed that the small firms are viable, we assume that the fixed cost f is small enough so that $\bar{r}(t) > 0$ for all t on the optimal path. In this case the first-order condition for the choice of $\bar{r}(t)$ is satisfied with equality at every point in time, implying that the unit cost $C(\cdot)$ is constant on the dynamic path and equal to C^*, implicitly defined by

$$\frac{1}{1-\sigma} C^{*\delta} \bar{a}^{1-\sigma} = f. \tag{33}$$

In what follows we use asterisks to denote optimal values of endogenous variables. This result implies that the real consumption index X is also constant on the dynamic path and $X(t) = X^* = C^{*-\varepsilon}$ for all t. Finally, from the minimization problem that was used to derive the cost function $C(\cdot)$ we obtain the optimal output quantities:

$$\bar{x}^* = C^{*\delta} \bar{a}^{-\sigma} = (\sigma - 1)\bar{a}^{-1} f, \tag{34}$$

$$x_i^* = C^{*\delta} a_i^{-\sigma} = (\sigma - 1)\bar{a}^{\sigma-1} a_i^{-\sigma} f, \; i = 1, 2, ..., m. \tag{35}$$

Evidently, the optimal output level of every variety is constant and a firm's output level is larger the lower its marginal cost. These findings are summarized in

Proposition 9. *Consider the optimal allocation in an economy that has a low entry cost f that secures $\bar{r}^* > 0$ for all t. Then: (i) the output levels $\bar{x}^*$ and $\{x_i^*\}_{i=1}^m$ are constant on the optimal path; (ii) the unit cost C^* and the real consumption index X^* are constant on the optimal path and $X^* = C^{*-\varepsilon}$.*

Using these results, we can express the flow of utility as a function of product spans of the multi-product firms and their investment levels. From (31) and (32) and the result that $C(\cdot) = C^*$

on the dynamic path, the flow of utility satisfies:

$$u\left[\{r_i(t)\}_{i=1}^m, \{\iota_i(t)\}_{i=1}^m\right] = l + \frac{1}{\varepsilon - 1} C^{*1-\varepsilon} - \left[C^{*1-\sigma} - \sum_{i=1}^m r_i(t)\, a_i^{1-\sigma}\right] \overline{a}^{\sigma-1} f - \sum_{i=1}^m \iota_i(t)\,. \quad (36)$$

The term in the square brackets shows that larger product spans of multi-product firms call for fewer single-product firms in order to ensure a constant value $C(\cdot) = C^*$. This lowers entry costs of the small firms, thereby saving resources and raising welfare. There is, however, a tradeoff: the growth of product span of large firms requires investment in innovation, which reduces consumption and welfare. This tradeoff is optimized in the dynamic problem that we solve next.

The dynamic optimal allocation problem can be formulated as follows:

$$\max_{\left[\{r_i(t)\}_{i=1}^m, \{\iota_i(t)\}_{i=1}^m\right]_{t\geq 0}} \int_0^\infty e^{-\rho t} u\left[\{r_i(t)\}_{i=1}^m, \{\iota_i(t)\}_{i=1}^m\right] dt$$

subject to (13), (36), $r_i(0) = r_i^0$, and transversality conditions to be described below. In this problem the state variables are $\{r_i(t)\}_{i=1}^m$ while the control variables are $\{\iota_i(t)\}_{i=1}^m$. Dropping t in the notation of time-dependent variables, the current-value Hamiltonian is:

$$\mathcal{H} = l + \frac{1}{\varepsilon - 1} C^{*1-\varepsilon} - \left(C^{*1-\sigma} - \sum_{i=1}^m r_i a_i^{1-\sigma}\right) \overline{a}^{\sigma-1} f - \sum_{i=1}^m \iota_i + \sum_{i=1}^m \lambda_i \left[\phi(\iota_i) - \theta r_i\right],$$

where λ_i is the co-state variable of constraint (13). The first-order conditions of this problem are:

$$\frac{\partial \mathcal{H}}{\partial \iota_i} = -1 + \lambda_i \phi'(\iota_i) = 0, \quad (37)$$

$$-\frac{\partial \mathcal{H}}{\partial r_i} = -\left(\frac{\overline{a}}{a_i}\right)^{\sigma-1} f + \theta \lambda_i = \dot{\lambda}_i - \rho \lambda_i, \quad (38)$$

and the transversality conditions are:

$$\lim_{t\to\infty} e^{-\rho t} \lambda_i(t)\, r_i(t) = 0.$$

In addition, the optimal path of $\{\iota_i, r_i\}_{i=1}^m$ has to satisfy the differential equations (13). Equation (38) yields the autonomous differential equation

$$\dot{\lambda}_i = (\rho + \theta)\,\lambda_i - \left(\frac{\overline{a}}{a_i}\right)^{\sigma-1} f.$$

A second differential equation is obtained from (13) and (37):

$$\dot{r}_i = \phi\left[\iota_i(\lambda_i)\right] - \theta r_i.$$

These equations generate transition dynamics similar to Figure 1, except that now the $\dot{\lambda}_i = 0$ curve is horizontal and therefore λ_i is constant during the transition and equal to:

$$\lambda_i^* = \frac{1}{\rho + \theta} \left(\frac{\overline{a}}{a_i} \right)^{\sigma - 1} f. \tag{39}$$

This implies that the dynamic system travels on the $\dot{\lambda}_i = 0$ curve and it satisfies the transversality conditions. Since all multi-product firms share the same $\dot{r}_i = 0$ curve but their $\dot{\lambda}_i = 0$ curves differ according to a_i, it follows that firms with lower marginal costs a_i end up with larger product spans in the steady state. In other words, in the steady state of the optimal allocation there is a *monotonically* decreasing relationship between marginal cost and product span in the cross section of multi-product firms. Evidently, an inverted-U curve relationship between these variables arises only in a distorted economy. Finally, note that due to the fact that the co-state variable λ_i^* is constant on the optimal path, so is investment in innovation ι_i^*, as can be seen from the first-order condition (37). These findings are summarized in the following

Proposition 10. *Consider the optimal allocation in an economy that has a low entry cost f that secures $\bar{r}^* > 0$ for all t. Then: (i) investment in innovation ι_i^* is constant on the optimal path and larger the smaller is a multi-product firm's marginal cost; (ii) in steady state multi-product firms with lower marginal costs have larger product spans; (iii) if $r_i(0) = r_i^0$ is smaller than the optimal steady state value of r_i^* for all i, then the product span of every large multi-product firm rises and the number of small single-product firms declines on the optimal path.*

To decentralize the optimal allocation, it is necessary to subsidize consumer purchases of every firm's varieties so as to ensure consumer prices that equal marginal costs of production. In addition, every large firm's operating profits have to be taxed to ensure that the firm perceives a constant marginal value of product spans on its entire dynamic path, yet no policy is required to modify entry incentives of single-product firms. We provide in the Appendix a full characterization of these policies. One important feature of the optimal policies is that the subsidy to consumers on purchases of products supplied by single-product firms does not vary over time, while subsidies to consumer purchases of products supplied by large multi-product firms have to vary over time on the transition path. Second, operating profits of large firms have to by taxed in order to induce them to engage in optimal investment in innovation. When the initial product span of a firm is below the optimal steady state value, this firm expands its product span over time. In this case the time pattern of the optimal corporate tax depends on the relative size of the elasticity of substitution σ and the sectoral elasticity of demand ε. If $\sigma > 2\varepsilon$ there exists a market share $s_c = \sigma/2\,(\sigma - \varepsilon)$ such that the tax rate is rising if the share of consumer spending on the firm's products is lower than s_c and the tax rate is declining if the share of consumer spending on the firm's products is larger than s_c. In the opposite case, $\sigma < 2\varepsilon$, the corporate tax rate always rises for a firm that grows its product span. Derivation of these results are provided in the Appendix.

7 Conclusion

We have developed a parsimonious model of industry evolution, in which large multi-product firms grow via investment in new product lines. While these firms are oligopolies, they face competitive pressure from small single-product firms that engage in monopolistic competition. Our model generates time patterns of markups, concentration, and labor shares that are consistent with the data. Moreover, it predicts rich patterns for the cross-section of firms. In particular, it predicts an inverted-U relation between labor productivity and product span, for which we provide supportive evidence. It also predicts that rising competitive pressure from small single-product firms flattens the cross-sectional relationship between labor productivity and product span among the large multi-product firms.

We also characterized the optimal allocation and compared it to the market outcome. In the optimal allocation there is a monotonically increasing relationship between labor productivity and product span, which implies that the inverted-U relationship is caused by misallocation.

Although this study consists of a theoretical contribution, we believe that our model delivers valuable insights into industry dynamics that can be empirically studied. There are few data sets containing information on product span of individual firms, and these data are mostly confidential. Nevertheless, we hope that the predictions of our model will eventually be examined with some of the existing rich data sets. Finally, we show in the appendix how to construct an aggregate economy with a continuum of industries of the type studied in this paper. This model economy can be used to study various macroeconomic issues, including economic growth.

References

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J. and Li, H. (2019) A Theory of Falling Growth and Rising Rents.

Aghion, P. and Howitt, P. (1992) A Model of Growth through Creative Destruction, *Econometrica*, **60**, 323–51.

Atkeson, A. and Burstein, A. (2008) Pricing-to-Market, Trade Costs, and International Relative Prices, *American Economic Review*, **98**, 1998–2031.

Autor, D., Dorn, D., Katz, L. F., Patterson, C. and Van Reenen, J. (2020) The Fall of the Labor Share and the Rise of Superstar Firms, *The Quarterly Journal of Economics*.

Cao, D., Hyatt, H., Mukoyama, T. and Sager, E. (2019) Firm Growth Through New Establishments, *SSRN Electronic Journal*.

Cavenaile, L., Celik, M. A. and Tian, X. (2019) Are Markups Too High? Competition, Strategic Innovation, and Industry Dynamics, *SSRN Electronic Journal*.

De Loecker, J., Eeckhout, J. and Unger, G. (2020) The Rise of Market Power and the Macroeconomic Implications, *The Quarterly Journal of Economics*, **135**, 561–644.

Edmond, C., Midrigan, V. and Yi Xu, D. (2019) How Costly Are Markups?, mimeo.

Grossman, G. M. and Helpman, E. (1991) Quality Ladders in the Theory of Growth, *The Review of Economic Studies*, **58**, 43–61.

Gutierrez, G. and Philippon, T. (2019) The Failure of Free Entry, NBER Working Paper Series 26001, National Bureau of Economic Research, Inc.

Hsieh, C.-T. and Rossi-Hansberg, E. (2020) The Industrial Revolution in Services, Tech. rep., Princeton.

Kehrig, M. and Vincent, N. (2019) Good Dispersion, Bad Dispersion, NBER Working Paper Series 25923, National Bureau of Economic Research, Inc.

Klette, T. J. and Kortum, S. (2004) Innovating Firms and Aggregate Innovation, *Journal of Political Economy*, **112**, 986–1018.

Parenti, M. (2018) Large and Small Firms in a Global Market: David vs. Goliath, *The Journal of International Economics*, **110**, 103–118.

Romer, P. M. (1990) Endogenous Technological Change, *Journal of Political Economy*, **98**, S71–S102.

Shimomura, K.-I. and Thisse, J.-F. (2012) Competition Among the Big and the Small, *The RAND Journal of Economics*, **43**, 329–347.

Appendix

Comparative Dynamics

We first derive the slope of the $\dot{\lambda}_i$=0 curve. Differentiation of the right-hand side of (22) yields:

$$\hat{\mathit{\Gamma}}_i = -(\sigma-1)\,\hat{a}_i + \delta\hat{P} - \frac{\sigma\delta s_i}{(\sigma-\delta s_i-1)(\sigma-\delta s_i)}\hat{s}_i + \frac{\delta s_i(\sigma-2\delta s_i)}{(\sigma-\delta s_i-1)\,\sigma+\delta^2 s_i^2}\hat{s}_i.$$

This equation implies that the right-hand side of (22) is declining in r_i because $\mathit{\Gamma}_i$ is declining in s_i and s_i is rising in r_i (see (11)). The former is seen from this equation by observing that $\sigma\delta s_i > \delta s_i(\sigma-2\delta s_i)$ and $(\sigma-\delta s_i-1)(\sigma-\delta s_i) <(\sigma-\delta s_i-1)\,\sigma+s_i^2\delta^2$. Collecting terms we can rewrite this equation as:

$$\hat{\mathit{\Gamma}}_i = -(\sigma-1)\,\hat{a}_i + \delta\hat{P} - \delta^2 s_i^2\frac{2(\sigma-\delta s_i-1)(\sigma-\delta s_i)+\sigma(\sigma-1)}{(\sigma-\delta s_i-1)(\sigma-\delta s_i)\left[(\sigma-\delta s_i-1)\,\sigma+\delta^2 s_i^2\right]}\hat{s}_i. \tag{40}$$

Next consider the total effect of a shift in the marginal cost a_i on $\mathit{\Gamma}_i$. From (11) we have:

$$\hat{s}_i = -\frac{\sigma-1}{1+(\sigma-1)\,\beta_i}\hat{a}_i = -\frac{(\sigma-1)(\sigma-\delta s_i-1)(\sigma-\delta s_i)}{(\sigma-\delta s_i-1)(\sigma-\delta s_i)+(\sigma-1)\,\delta s_i}\hat{a}_i.$$

Substituting this expression into (40) we obtain the total impact of a_i on $\mathit{\Gamma}_i$:

$$\begin{aligned}\frac{\hat{\mathit{\Gamma}}_i}{(\sigma-1)\,\hat{a}_i} &= -1+\delta^2 s_i^2\frac{2(\sigma-\delta s_i-1)(\sigma-\delta s_i)+\sigma(\sigma-1)}{\left[(\sigma-\delta s_i-1)\,\sigma+s_i^2\delta^2\right]^2}\\ &= \frac{(\sigma-1)\,s_i^2\delta^2-(\sigma-\delta s_i-1)^2\left(\sigma^2-\delta^2 s_i^2\right)}{\left[(\sigma-\delta s_i-1)\,\sigma+s_i^2\delta^2\right]^2}.\end{aligned}$$

It follows that a decline in the marginal cost a_i shifts upward the $\dot{\lambda}_i$=0 curve if and only if $(\sigma-1)\,s_i^2\delta^2 < (\sigma-\delta s_i-1)^2\left(\sigma^2-\delta^2 s_i^2\right)$.

Empirical Analysis

We now provide additional information on the empirical analysis. Table 1 presents the data that has has been used to construct Figure 4 while Table 2 presents the regression results. As pointed out in the main text, the coefficient for log productivity is positive and significantly different from zero and the coefficient for the square of log productivity is negative and significantly different from zero in both specifications; i.e., when we use the number of industries or the number of segments to measure a firm's product span. While in the main text we reported in Figure 3 the curvature of this quadratic form for the number of segments as a proxy for the number of product lines, we now report a similar figure, Figure 6, for the case in which the number of industries is used as a proxy for the number of product lines. As is evident, the two figures are quite similar.

Table 1: Average Number of Product Lines vs. Productivity Deciles

Decile	*Log(Prod)*	*MeanInd*	*MeanSegs*
1	10.05	1.89	2.93
2	11.54	2.14	3.65
3	12.04	2.27	4.00
4	12.31	2.48	4.47
5	12.54	2.64	4.84
6	12.77	2.67	4.98
7	13.06	2.63	4.83
8	13.42	2.53	4.79
9	13.91	2.29	4.57
10	15.31	1.92	3.99

Note: This table shows the deciles of average log labor productivity for firms in the Compustat database for the year 2018, available through WRDS. Labor productivity is defined as the ratio of total sales to employment. It also shows the mean number of industries and business segments that are reported in the Compustat Segments Data. The data was accessed on June 2, 2020.

Table 2: Quadratic Relationship of Productivity on Product Span

	Industries	Segments
log(Prod)	2.85**	5.50**
	(1.33)	(2.54)
log(Prod)^2	-0.11*	-0.21**
	(0.06)	(.11)
Primary Ind. FE	YES	YES
Obs	4126	4126
R^2	0.7334	0.4603

Robust standard errors clustered at the primary industry in parentheses.

$^*p < 0.10, ^{**}p < 0.05$.

Note: This table shows the results of an OLS quadratic regression of the number of industries or segments on the log of labor productivity. The data includes all firms with positive sales and employment in the Compustat database for the year 2018. Labor productivity is defined as the ratio of total sales to employment. Segments here refers to the total number of business segments listed in the Compustat Segments Data by firm. The number of industries is the number of primary and secondary SIC codes listed across all business segments. We also include fixed effects for 4 digit primary SIC code listed on Compustat. Data was accessed on June 2, 2020.

Optimal Allocation

In Section 6 we characterize the optimal allocation, showing that it differs from the market outcome. In this part of the appendix we propose policies that implement the optimal allocation in a market economy with taxes and subsidies. In particular, we show that there exist consumer subsidies for the purchase of varieties of the differentiated product and corporate taxes on operating profits that lead to a market allocation that coincides with the optimal allocation. These taxes and subsidies are firm specific and they vary over time. Moreover, implementation of the optimal allocation requires the policy maker to commit to the entire time path of these taxes and subsidies, which vary across firms and across time.

Let $\overline{\gamma}$ be the factor that converts a producer price $\overline{p}$ into a consumer price $\overline{\gamma}\,\overline{p}$ and by γ_i the factor that converts a producer price p_i into a consumer price $\gamma_i p_i$. We allow these conversion factors to vary over time, although we will find that the optimal value of $\overline{\gamma}$ is constant. Importantly, both consumers and producers treat these factors as exogenous variables. A γ smaller than one

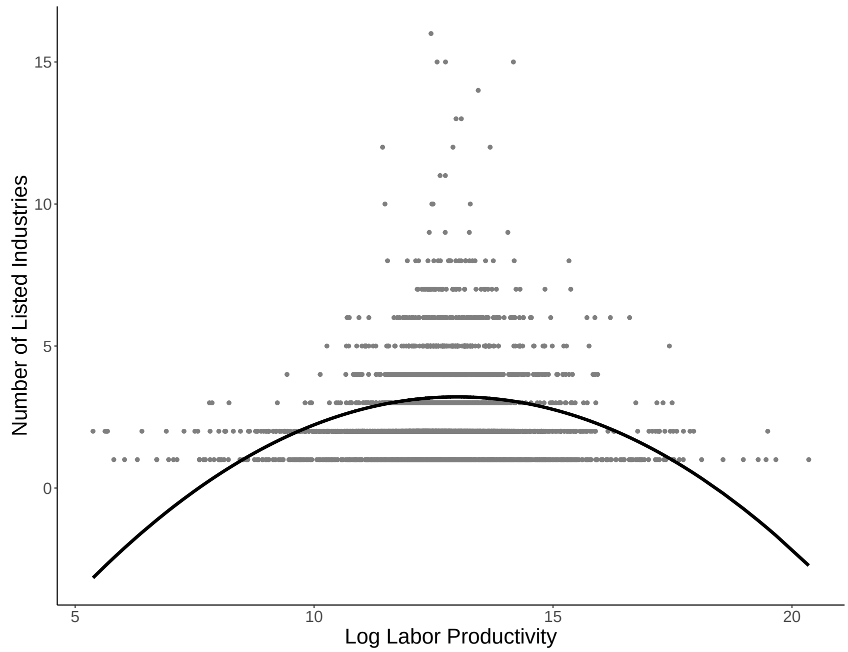

Figure 6: Number of Industries vs. Labor Productivity

represents a subsidy to consumers while a γ larger than one represents a tax. Finally, we denote by τ_i the factor that converts gross operating profits of firm i, $r_i P^{\delta} (\gamma_i p_i)^{-\sigma} (p_i - a_i)$, into net operating profits $\tau_i r_i P^{\delta} (\gamma_i p_i)^{-\sigma} (p_i - a_i)$. The factors τ_i may also vary over time, but the firms treat them as exogenous variables. A τ_i smaller than one represents a corporate tax on operating profits while a τ_i larger than one represents a corporate subsidy to operating profits.

With these policies in place, the demand for varies of the differentiated product (3) can be expressed as:

$$\overline{x} = P^{*\delta} (\overline{\gamma}\,\overline{p})^{-\sigma},$$

$$x_i = P^{*\delta} (\gamma_i p_i)^{-\sigma},$$

where

$$P^* = \left[\overline{r} (\overline{\gamma}\,\overline{p})^{1-\sigma} + \sum_{j=1}^{m} r_j (\gamma_i p_j)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}.$$

In this exposition we use asterisks to denote equilibrium values of endogenous variables in the economy with taxes and subsidies. Large firms now maximize net operating profits $\tau_i r_i P^{*\delta} (\gamma_i p_i)^{-\sigma} (p_i - a_i)$ while small firms maximize operating profits $P^{*\delta} (\overline{\gamma}\,\overline{p})^{-\sigma} (\overline{p} - \bar{a})$. This yields the optimal pricing equations:

$$\overline{p}^* = \frac{\sigma}{\sigma - 1} \bar{a},$$

$$p_i^* = \frac{\sigma - \delta s_i^*}{\sigma - \delta s_i^* - 1} a_i, \tag{41}$$

where s_i^* is the share of consumer spending on goods of firm i, equal to

$$s_i^* = \frac{r_i \left(\gamma_i p_i^*\right)^{1-\sigma}}{P^{*1-\sigma}}. \tag{42}$$

We now propose the following numerical values of these policies:

$$\overline{\gamma} = \frac{\sigma - 1}{\sigma} \quad \text{and} \quad \gamma_i = \frac{\sigma - \delta s_i^* - 1}{\sigma - \delta s_i^*}, \tag{43}$$

which yields $\overline{\gamma}\,\overline{p} = \bar{a}$ and $\gamma_i p_j = a_j$. In other words, these policies lead to consumer prices that equal marginal costs of production. Note that every γ is smaller than one. Therefore consumers enjoy subsidies on all varieties of the differentiated product and the subsidies are larger on products with larger market shares.

With these subsidies a small firm's operating profits are $P^{*\delta} \left(\overline{\gamma}\,\overline{p}^*\right)^{-\sigma} \left(\overline{p}^* - \bar{a}\right)$, and free entry ensures that these profits equal the entry cost f. Using the firm's optimal pricing equation (41) and the subsidy policy (43), this free entry condition yields

$$\frac{1}{\sigma - 1} P^{*\delta} \bar{a}^{1-\sigma} = f.$$

Comparing this to (33), we conclude that $P^* = C^*$, i.e., the price index equals the optimal resource cost of producing a unit of real consumption X. As a result, real consumption X is also at the optimal level, equal to $X^* = C^{*-\varepsilon}$, and the consumption levels of individual varies are at the optimal levels (see (34) and (35)):

$$\overline{x}^* = C^{*\delta} \overline{a}^{-\sigma} = (\sigma - 1)\overline{a}^{-1} f,$$

$$x_i^* = C^{*\delta} a_i^{-\sigma} = (\sigma - 1)\overline{a}^{\sigma-1} a_i^{-\sigma} f.$$

It remains to examine the investment policies of large firms.

Recognizing that $P^* = C^*$ is constant on the dynamic path, (41) and (42) implicitly define the optimal price of firm i as a function of its product span, $p_i^*(r_i)$, similarly to the analysis of the market economy without government intervention. The only difference is that now there are policy instruments that the firms treat as exogenous. As a result, profits of firm i net of taxes and investment costs are

$$\pi_i(\iota_i, r_i) = \tau_i r_i C^{*\delta} \left[\gamma_i p_i^*(r_i)\right]^{-\sigma} \left[p_i^*(r_i) - a_i\right] - \iota_i$$

and the current value Hamiltonian of the firm's optimal control problem is

$$\mathcal{H}(\iota_i, r_i, \lambda_i) = \tau_i r_i C^{*\delta} \left[\gamma_i p_i^*(r_i)\right]^{-\sigma} \left[p_i^*(r_i) - a_i\right] - \iota_i + \lambda_i \left[\phi(\iota_i) - \theta r_i\right].$$

The first-order conditions for the optimal control problem are therefore:

$$\frac{\partial \mathcal{H}}{\partial \iota_i} = -1 + \lambda_i \phi' (\iota_i) = 0,$$

$$-\frac{\partial \mathcal{H}}{\partial r_i} = -\tau_i \mathit{\Gamma}_i^* (r_i) + \theta \lambda_i = \dot{\lambda}_i - \rho \lambda_i,$$

where

$$\mathit{\Gamma}_i^* (r_i) \equiv \frac{\partial \left\{ r_i C^{*\delta} \left[\gamma_i p_i^* (r_i) \right]^{-\sigma} \left[p_i^* (r_i) - a_i \right] \right\}}{\partial r_i},$$

and the transversality conditions are:

$$\lim_{t \to \infty} e^{-\rho t} \lambda_i (t) r_i(t) = 0.$$

Now recall that the optimal investment in innovation is constant on the dynamic path and satisfies $\lambda_i^* \phi' (\iota_i^*) = 1$, where λ_i^* is is given in (39), i.e.,

$$\lambda_i^* = \frac{1}{\rho + \theta} \left(\frac{\overline{a}}{a_i} \right)^{\sigma - 1} f.$$

The first-order conditions of the firm's optimal control problem imply that this investment pattern is attained if and only if:

$$\tau_i \mathit{\Gamma}_i^* (r_i^*) = \left(\frac{\overline{a}}{a_i} \right)^{\sigma - 1} f \tag{44}$$

at every point in time, where r_i^* is the optimal product span. It follows from this result that operating profits of firm i are taxed ($\tau_i < 1$) if and only if

$$\mathit{\Gamma}_i^* (r_i^*) > \left(\frac{\overline{a}}{a_i} \right)^{\sigma - 1} f.$$

We show next that $\tau_i \epsilon (0, 1)$; that is, the optimal policy consists of taxing operating profits.

First note that

$$\frac{\mathit{\Gamma}_i^* (r_i)}{C^{*\delta}} = \frac{\partial \left\{ r_i \left[\gamma_i p_i^* (r_i) \right]^{-\sigma} \left[p_i^* (r_i) - a_i \right] \right\}}{\partial r_i}$$

$$= \left[\gamma_i p_i^* (r_i) \right]^{-\sigma} \left[p_i^* (r_i) - a_i \right] - \left\{ \sigma r_i \gamma_i^{-\sigma} p_i^* (r_i)^{-\sigma - 1} \left[p_i^* (r_i) - a_i \right] - r_i \left[\gamma_i p_i^* (r_i) \right]^{-\sigma} \right\} \frac{\partial p_i^* (r_i)}{\partial r_i}$$

$$= \gamma_i^{-\sigma} p_i^* (r_i)^{-\sigma} \left(\left[p_i^* (r_i) - a_i \right] - \left\{ \sigma r_i p_i^* (r_i)^{-1} \left[p_i^* (r_i) - a_i \right] - r_i \right\} \frac{\partial p_i^* (r_i)}{\partial r_i} \right)$$

$$= \gamma_i^{-\sigma} p_i^* (r_i)^{-\sigma} \left[\frac{1}{\sigma - \delta s_i^* (r_i) - 1} a_i - r_i \left[\frac{\sigma}{\sigma - \delta s_i^* (r_i)} - 1 \right] \frac{\partial p_i^* (r_i)}{\partial r_i} \right]$$

$$= \gamma_i^{-\sigma} p_i^* (r_i)^{-\sigma} \left[\frac{1}{\sigma - \delta s_i^* (r_i) - 1} a_i - r_i \frac{\delta s_i^* (r_i)}{\sigma - \delta s_i^* (r_i)} \frac{\partial p_i^* (r_i)}{\partial r_i} \right].$$

However,

$$\frac{\partial p_i^* (r_i)}{\partial r_i} \frac{r_i}{p_i^*} = \frac{\beta_i^* (r_i)}{1 + (\sigma - 1) \beta_i^* (r_i)}$$

where

$$\beta_i^* (r_i) = \frac{\delta s_i^* (r_i)}{[\sigma - \delta s_i^* (r_i) - 1] [\sigma - \delta s_i^* (r_i)]}.$$

Therefore

$$\begin{aligned} \varGamma_i^* (r_i) &= C^{*\delta} \gamma_i^{-\sigma} p_i^* (r_i)^{-\sigma} \left[\frac{1}{\sigma - \delta s_i^* (r_i) - 1} a_i - \left[p_i^* (r_i) \frac{\delta s_i^* (r_i)}{\sigma - \delta s_i^* (r_i)} \right] \frac{\beta_i^* (r_i)}{1 + (\sigma - 1) \beta_i^* (r_i)} \right] \\ &= \gamma_i^{-\sigma} a_i^{1-\sigma} C^{*\delta} \left[\frac{\sigma - \delta s_i^* (r_i)}{\sigma - \delta s_i^* (r_i) - 1} \right]^{-\sigma} \frac{\sigma}{[\sigma - \delta s_i^* (r_i) - 1] \sigma + \delta^2 s_i^* (r_i)^2} \\ &= \gamma_i^{-\sigma} \left(\frac{\overline{a}}{a_i} \right)^{\sigma - 1} f \left[\frac{\sigma - \delta s_i^* (r_i)}{\sigma - \delta s_i^* (r_i) - 1} \right]^{-\sigma} \frac{\sigma (\sigma - 1)}{[\sigma - \delta s_i^* (r_i) - 1] \sigma + \delta^2 s_i^* (r_i)^2}, \end{aligned}$$

where we used (33) in deriving the last line. Now compare this formula to (20). Since we showed that the expression on the right-hand side of (20) declines in r_i, it follows that— holding γ_i constant— $\varGamma_i^* (r_i)$ also declines in r_i. This ensures concavity in r_i of the firm's decision problem.

Finally, we show that $\tau_i \epsilon (0, 1)$ in every time period, implying that the optimal policy consists of a tax on operating profits. To this end use the formula for the subsidy factor γ_i together with the optimal tax formula (44) to obtain:

$$\begin{aligned} \tau_i &= \varGamma_i^* (r_i)^{-1} \left(\frac{\overline{a}}{a_i} \right)^{\sigma - 1} f = \frac{\sigma [\sigma - \delta s_i^* (r_i) - 1] + \delta^2 s_i^* (r_i)^2}{\sigma (\sigma - 1)}, \\ &= 1 - \frac{[\sigma - \delta s_i^* (r_i)] \delta s_i^* (r_i)}{\sigma (\sigma - 1)}, \end{aligned}$$

which shows that $\tau_i \epsilon (0, 1)$ at every point in time.

For a firm with rising product span the share of consumer spending on its products rises over time, i.e., $s_i^* (r_i)$ is an increasing function. Therefore the corporate tax rate is rising over time (τ_i is decreasing) if and only $\sigma > 2 \delta s_i^* (r_i)$. Since $\delta = \sigma - \varepsilon > 0$, it follows that for $\sigma > 2\varepsilon$ there exists a market share $s_c = \sigma / 2 (\sigma - \varepsilon)$ such that the tax rate is rising for market shares below s_c and declining for larger market shares. In the opposite case, when $\sigma > 2\varepsilon$, the corporate tax rate always rises for firms that expand their product span.

Comparative Statics: Given Number of Brands

In this section we examine the case in which the number of single-product firms, $\bar{r}$, as well the number of products available to each one of the large firms, r_i, are given. Equations (7) and (8) imply:

$$\hat{p}_i = \hat{a}_i + \frac{\delta s_i}{(\sigma - \delta s_i - 1)(\sigma - \delta s_i)}\hat{s}_i, \tag{45}$$

$$\hat{s}_i = \hat{r}_i - \sum_{j=1}^{m} s_j \hat{r}_j - (\sigma - 1)(\hat{p}_i - \sum_{j=1}^{m} s_j \hat{p}_j).$$

Substituting the last equation into (45) yields:

$$[1 + \beta_i(\sigma - 1)]\hat{p}_i - \beta_i(\sigma - 1)\sum_{j=1}^{m} s_j \hat{p}_j = \hat{a}_i + \beta_i(\hat{r}_i - \sum_{j=1}^{m} s_j \hat{r}_j), \text{ for all } i.$$

These equations can also be expressed as:

$$\mathbf{B}\hat{\mathbf{p}} = \mathbf{R}\hat{\mathbf{r}} + \hat{\mathbf{a}}, \tag{46}$$

where $\mathbf{B}$ is an $m \times m$ matrix with elements:

$$b_{ii} = 1 + \beta_i(\sigma - 1)(1 - s_i),$$

$$b_{ij} = -\beta_i(\sigma - 1)s_j, \text{ for } j \neq i,$$

$\hat{\mathbf{p}}$ is an $m \times 1$ column vector with elements p_i, where a hat represents a proportional rate of change (i.e., $\hat{p}_i = dp_i/p_i$), $\mathbf{R}$ is an $m \times m$ matrix with elements:

$$r_{ii} = \beta_i(1 - s_i),$$

$$r_{ij} = -\beta_i s_j, \text{ for } j \neq i,$$

$\hat{\mathbf{r}}$ is an $m \times 1$ column vector with elements $\hat{r}_i$, where a hat represents a proportional rate of change, and $\hat{\mathbf{a}}$ is an $m \times 1$ column vector with elements $\hat{a}_i$, where a hat represents a proportional rate of change.

Since

$$|b_{ii}| - \sum_{j \neq i} |b_{ij}| = 1 + \beta_i(\sigma - 1)(1 - \sum_{j=1}^{m} s_j) > 1,$$

$\mathbf{B}$ is a diagonally dominant matrix with positive diagonal and negative off-diagonal elements. It therefore is an M-matrix and its inverse has all positive entries. This inverse, denoted by $\tilde{\mathbf{B}} = \mathbf{B}^{-1}$, is therefore an $m \times m$ matrix with elements $\tilde{b}_{ij} > 0$. Next note that $\mathbf{B}$ can be expressed as:

$$\mathbf{B} = \mathbf{I} + (\sigma - 1)\mathbf{R},$$

where $\mathbf{I}$ is the identity matrix. Therefore:

$$\mathbf{B}^{-1}\mathbf{B} = \tilde{\mathbf{B}} + (\sigma - 1)\tilde{\mathbf{B}}\mathbf{R} = \mathbf{I}. \tag{47}$$

It follows from this equation that:

$$\tilde{b}_{ii} + (\sigma - 1)\sum_{j=1}^{m} \tilde{b}_{ij} r_{ji} = 1,$$

$$\tilde{b}_{ik} + (\sigma - 1)\sum_{j=1}^{m} \tilde{b}_{ij} r_{jk} = 0, \text{ for } k \neq i.$$

Summing these up yields:

$$\sum_{k=1}^{m} \tilde{b}_{ik} + (\sigma - 1)\sum_{j=1}^{m} \tilde{b}_{ij} \sum_{k=1}^{m} r_{jk} = 1, \text{ for all } i. \tag{48}$$

Since:

$$\sum_{k=1}^{m} r_{jk} = \beta_j (1 - \sum_{k=1}^{m} s_k) > 0$$

and $\tilde{b}_{ik} > 0$ for all i and k, it follows from (48) that:

$$0 < \tilde{b}_{ik} < 1 \text{ for all } i \text{ and } k.$$

Equation (47) implies:

$$(\sigma - 1)\tilde{\mathbf{B}}\mathbf{R} = \mathbf{I} - \tilde{\mathbf{B}},$$

and therefore $\tilde{\mathbf{B}}\mathbf{R}$ has positive diagonal elements and negative off-diagonal elements.

Going back to the comparative statics equations (46), we have:

$$\hat{\mathbf{p}} = \tilde{\mathbf{B}}\mathbf{R}\hat{\mathbf{r}} + \tilde{\mathbf{B}}\hat{\mathbf{a}}.$$

It follows from the properties of $\tilde{\mathbf{B}}$ that a decline in a_i reduces every price p_j, but less than proportionately. Equation (45) than implies that all market share $s_j, j \neq i$, decline while the market share s_i rises. And it follows from the properties of $\tilde{\mathbf{B}}\mathbf{R}$ and (45) that an increase in r_i raises the price and market share of firm i and reduces the price and market share of every other firm $j \neq i$. Noting that the markup of every firm i is larger the larger its market share, we therefore have:

Proposition 11. *Suppose that the number of firms and their product ranges are given. Then: (i)* *an increase in r_i raises the price, markup and market share of firm i, and reduces the price, markup and market share of every other large firm; (ii) a decline in a_i reduces the price of every large firm less than proportionately, raises the markup and market share of firm i, and reduces the markup and market share of every other large firms.*

Aggregative Economy

In this section we show how to construct an aggregative economy with a continuum of industries, each one of the type analyzed in the main text of this paper.

We consider an economy with a continuum of individuals of mass 1, each one providing one unit of labor. The labor market is competitive and every individual earns the same wage rate.

There is a continuum of sectors of measure one, each one producing a differentiated product. Real consumption in sector k is:

$$X^k = \left[\int_0^{N^k} x^k(\omega)^{\frac{\sigma^k - 1}{\sigma^k}} d\omega \right]^{\frac{\sigma^k}{\sigma^k - 1}}, \ \sigma^k > 1,$$

where N^k is the number of varieties available in sector k, $x^k(\omega)$ is consumption of variety ω in sector k, and σ^k is the elasticity of substitution in sector k. Using this definition, the price index of X^k is:

$$P^k = \left[\int_0^{N^k} p^k(\omega)^{1-\sigma^k} d\omega \right]^{\frac{1}{\sigma^k - 1}},$$

where $p^k(\omega)$ is the price of variety ω. The log utility of a representative individual is:

$$log\,(u) = \int_0^1 log\left(X^k\right) dk.$$

In these circumstances every individual spends an equal amount of money in every sector. Therefore, if E denotes aggregate spending per capita, spending per capita in sector k also equals E. In this event, aggregate demand for variety ω in sector k is:

$$x^k(\omega) = A^k p(\omega)^{-\sigma}, \tag{49}$$

$$A^k = E\left(P^k\right)^{\sigma^k - 1}. \tag{50}$$

An individual's inter-temporal utility function is:

$$U = \int_0^{\infty} e^{-\rho t} log\,(u_t)\, dt,$$

where ρ is the subjective discount rate. As a result, the intertemporal allocation of spending satisfies:

$$\frac{\dot{E}_t}{E_t} = \zeta_t - \rho, \tag{51}$$

where ζ_t is the interest rate at time t.

Two types of firms operate in sector k: atomless single-product firms and large multi-product firms, each one with a positive measure of product lines. Single-product firms produce $\overline{r}^k > 0$ varieties, each one specializing in a single brand. Large firm i in sector k has $r_i^k > 0$ product lines, $i = 1, 2, ..., m^k$, where m^k is the number of large firms in this sector. All the brands supplied to the market are distinct from each other.

All single-product firms share the same technology, which requires $\bar{a}^k$ unit of labor per unit output in sector k. Facing the demand function (49), a single-product firm maximizes profits $A^k p(\omega)^{-\sigma} \left[p(\omega) - \bar{a}^k\right]$, taking as given the demand shifter A^k. Therefore, a single-product firm prices its brand ω according to $p(\omega) = \overline{p}^k$, where:

$$\overline{p}^k = \frac{\sigma^k}{\sigma^k - 1} \bar{a}^k. \tag{52}$$

This yields the standard markup $\overline{\mu}^k = \sigma^k / \left(\sigma^k - 1\right)$ for a monopolistically competitive firm.

A large firm i has a technology that requires a_i^k units of labor per unit output, and it faces the demand function (49) for each one of its brands. As a result, it prices every brand equally. We denote this price by p_i^k. The firm chooses p_i^k to maximize profits $r_i^k A^k p_i^{-\sigma} \left(p_i - a_i^k\right)$. However, unlike a single-product firm, a large firm does not view A^k as given, because it recognizes that

$$P^k = \left(\overline{r}^k \left(\overline{p}^k\right)^{1-\sigma} + \sum_{j=1}^{m^k} r_j^k \left(p_j^k\right)^{1-\sigma^k} \right)^{\frac{1}{1-\sigma^k}}, \tag{53}$$

and therefore that its pricing policy has a measurable impact on the price index of the differentiated product. It takes, however, the spending level E as given, because sector k is of measure zero. Accounting for this dependence of P^k on the firm's price, the profit maximizing price is:

$$p_i^k = \frac{\sigma^k - \left(\sigma^k - 1\right) s_i^k}{\left(\sigma^k - 1\right)\left(1 - s_i^k\right)} a_i^k, \tag{54}$$

where s_i^k is the market share of firm i in sector k and:

$$s_i^k = \frac{r_i^k \left(p_i^k\right)^{1-\sigma^k}}{\left(P^k\right)^{1-\sigma^k}} = \frac{r_i^k \left(p_i^k\right)^{1-\sigma^k}}{\overline{r}^k \left(\overline{p}^k\right)^{1-\sigma} + \sum_{j=1}^{m^k} r_j^k \left(p_j^k\right)^{1-\sigma^k}}. \tag{55}$$

Equations (54) and (55) jointly determine prices and market shares of large firms. The markup factor of firm i is $\mu_i^k = \left[\sigma^k - \left(\sigma^k - 1\right) s_i^k\right] / \left[\left(\sigma^k - 1\right)\left(1 - s_i^k\right)\right]$, which is increasing in its market share. When the market share equals zero the markup is $\sigma^k / (\sigma^k - 1)$, the same as the markup of a single product firm. The markup factor varies across firms as a result of differences in either the product span, r_i^k, or the marginal production cost, a_i^k. We analyze the dependence of prices, market shares and markups on marginal costs and product spans in the next section.

Entry of Single-Product Firms

The number of large firms in every sector, m^k, is given. Unlike large firms, however, single-product firms enter the industry until their profits equal zero. In every sector the firms play a two-stage game: in the first stage single-product firms enter; in the second stage all firms play a Bertrand game as described above. Under these circumstances, (52) and (54) portray the equilibrium prices, except that the number of single product firms, $\bar{r}^k$, is endogenous. We seek to characterize a subgame perfect equilibrium of this game.

To determine the equilibrium number of single-product firms, assume that they face an entry cost f^k in sector k and they enter until profits equal zero. In a subgame perfect equilibrium every entrant correctly forecasts aggregate spending on the sector's products, the number of entrants, and the price that will be charged for every variety in the second stage of the game. Therefore, every single-product firm correctly forecasts the price index and A^k. Using the optimal price (52) and the profit function $A^k p(\omega)^{-\sigma}\left[p(\omega)-\bar{a}^k\right]$, this free entry condition can be expressed as:

$$\frac{1}{\sigma^k}A^k\left(\frac{\sigma^k}{\sigma^k-1}\bar{a}^k\right)^{1-\sigma^k}=f^k. \tag{56}$$

The left-hand side of this equation describes the operating profits, which equal a fraction $1/\sigma^k$ of revenue, while the right-hand side represents the entry cost. In these circumstances the demand shifter A^k is determined by f^k and $\bar{a}^k$, and it is rising in both f^k and $\bar{a}^k$. Importantly, it does not depend on the number of large firms nor on their product spans. Moreover, given the spending level E, which is determined at the economy-wide level and is not influenced by product spans in sector k (because the sector is of measure zero), the price index P^k is also independent of product spans in sector k. In particular, changes over time in this price index are driven by changes in aggregate spending. For this reason (50) and (51) imply:

$$\frac{\dot{P}_t^k}{P_t^k}=\frac{1}{\sigma^k-1}\left(\rho-\zeta_t\right). \tag{57}$$

Optimal Control

We can now compute the response of p_i^k and s_i^k to changes in r_i^k as we did in the main text, and use the solution in the firm's optimal control problem. In the optimal control problem large firm i in sector k takes as given the path of the interest rate r_t and the path of spending E_t. After characterizing this solution we can use it to express the market clearing conditions. Spending E_t has to equal wage income and aggregate profits net of investment costs. This will give us the growth model. If we use the formulation from the main text, the steady state will have zero growth. But one could add a long-run growth mechanism, such as declining costs of innovation as a function of the cumulative experience in innovation, as is Romer (1990). The steady state should be easy to analyze in either case.

As in the main text, investment is given by

$$\dot{r}_i^k = \phi(\iota_i^k) - \theta r_i^k, \text{ for all } t \geq 0, \tag{58}$$

At every point in time the firms play a two stage game. In the first stage single-product firms enter and large firms invest in innovation. Single-product firms live only one instant of time. For this reason they make profits only in this single instant. Under the circumstances the demand shifter A^k is determined by the free entry condition, and it remains constant as long as the cost of entry and the cost of production of the single-product firms do no change. It follows that the profit flow of large firm i is:

$$\pi_i^k = r_i^k A^k \left(p_i^k\right)^{-\sigma} (p_i^k - a_i^k) - \iota_i^k, \text{ for all } t \geq 0,$$

where A^k is the same at every t while π_i^k, r_i^k, p_i^k and ι_i^k change over time, and p_i^k is given by $p_i^k = \frac{\sigma^k - (\sigma^k - 1)s_i^k}{(\sigma^k - 1)\left(1 - s_i^k\right)} a_i^k$. We can write the optimal control problem as:

$$\max_{\left\{\iota_i^k(t), r_i^k(t)\right\}_{t \geq 0}} \int_0^\infty e^{-\int_0^t \zeta_\tau d\tau} \pi_i^k \left[\iota_i^k(t), r_i^k(t)\right] dt$$

The main difference between this formulation and the formulation in the main text is that now we no longer have $\zeta_t = \rho$ at each point in time, but rather $\zeta_t = \frac{\dot{E}_t}{E_t} + \rho$. The current-value Hamiltonian of this problem is:

$$\mathcal{H}(\iota_i^k, r_i^k, \lambda_i^k) = \left\{r_i^k A^k p_i^k (r_i)^{-\sigma} \left[p_i^k (r_i) - a_i^k\right] - \iota_i^k\right\} + \lambda_i^k \left[\phi\left(\iota_i^k\right) - \theta r_i^k\right],$$

and the first-order conditions are:

$$\frac{\partial \mathcal{H}}{\partial \iota_i^k} = -1 + \lambda_i^k \phi'\left(\iota_i^k\right) = 0,$$

$$-\frac{\partial \mathcal{H}}{\partial r_i^k} = -\frac{\partial \left[r_i^k A^k \left(p_i^k\right)^{-\sigma} (p_i^k - a_i^k)\right]}{\partial r_i^k} + \theta \lambda_i^k = \dot{\lambda}_i^k - \zeta_t \lambda_i^k.$$

Note that the path of the price index P_t^k is determined by the growth rate of the aggregate economy that each firm takes as exogenous. Therefore, the resulting first-order conditions have a similar form to those we derived in the main text:

$$\lambda_i^k \phi'\left(\iota_i^k\right) = 1, \tag{59}$$

$$\dot{\lambda}_i^k = (\zeta_t + \theta)\lambda_i^k - A^k p_i^k\left(r_i^k\right)^{-\sigma^k}\left\{p_i\left(r_i^k\right) - a_i^k - r_i^k\left(\sigma^k p_i^k\left(r_i^k\right)^{-1}\left[p_i^k\left(r_i^k\right) - a_i^k\right] - 1\right)\frac{dp^k\left(r_i^k\right)}{dr_i^k}\right\}. \tag{60}$$

Substituting (59) into (58) yields:

$$\dot{r}_i = \phi\left[\iota_i(\lambda_i)\right] - \theta r_i. \tag{61}$$

The second differential equation is obtained by substituting the pricing equation into (60):

$$\dot{\lambda}_i^k = (\zeta_t + \theta)\lambda_i^k - \Gamma_i^k\left(r_i^k\right), \tag{62}$$

where:

$$\Gamma_i^k\left(r_i^k\right) \equiv a_i^{1-\sigma^k} A^k \sigma\left[\frac{\sigma^k - (\sigma^k - 1)s_i^k\left(r_i^k\right)}{\left(\sigma^k - 1\right)\left(1 - s_i^k\right)}\right]^{-\sigma^k}\frac{1}{\left(\sigma^k - 1\right)\left(1 - s_i^k\right)\sigma + s_i\left(r_i\right)^2(\sigma - 1)^2}. \tag{63}$$

Thus, our two differential equations are similar to the main text, with the caveat that the interest rate is evolving over time. Specifically, the dynamics are such that aggregate spending must satisfy $\zeta_t = \frac{\dot{E}_t}{E_t} + \rho$.

In steady state:

$$\phi\left[\iota_i^k\left(\lambda_i^k\right)\right] = \theta r_i^k, \tag{64}$$

$$(\rho + \theta)\lambda_i^k = \Gamma_i^k\left(r_i^k\right), \tag{65}$$

where we have used the fact that in steady state $\zeta_t = \rho$. The comparative statics of this system have the same form as in the main text. But note that while the key condition for having an inverted-U relationship between productivity and product span was $(\sigma - \delta - 1)^2\left(\sigma^2 - \delta^2\right) < (\sigma - 1)\delta^2$ in the main text, the formula is the same now with the exception that δ is replaced with $\sigma - 1$. This reduces the condition to $0 < \left(\sigma^k - 1\right)^3$, which is always satisfied. Thus, in this formulation we would expect every sector to have the inverted-U property. Another comparative static to note is the effect of an increase in the steady state expenditure level E. This shifts upward the curve associated with (65) in the phase diagram, resulting in an instantaneous increase in λ_i^k and a trajectory of further expansion of r_i^k and rising profits. Thus, firms growing in other sectors reinforce the market dominance of large firms across industries through a pecuniary externality.

In order to close the model we need to solve for the steady state expenditure level. The market clearing condition is simply that revenue must equal net profits plus the total wage bill. With a unit mass of labor and the wage rate as the numeraire, the resulting condition takes the form:

$$E_t = 1 + \int_{k \in K} \left[\sum_{i=1}^{m^k} r_i^k A^k \left(p_i^k \right)^{-\sigma} (p_i^k - a_i^k) - \iota_i^k \right] dk. \tag{66}$$

We can further simplify this by recalling that $A^k = E_t \left(P^k \right)^{\sigma^k - 1}$. This means that we can use (66) to obtain:

$$E_t = \left[1 - \int_{k \in K} \left[\sum_{i=1}^{m^k} r_i^k \left(P^k \right)^{\sigma^k - 1} \left(p_i^k \right) (p_i^k - a_i^k) - \iota_i^k \right] dk \right]^{-1}. \tag{67}$$

Thus, the steady state expenditure level is increasing in the net profits of large firms across sectors. This equation also holds at every point in time, noting that the optimal investment levels depend on the path of aggregate expenditure through the interest rate. It follows that in order to solve the path of spending we need to ensure that the paths of profits of all firms aggregates to the path that rationalizes the optimal investments at each point in time.