Competition as an Engine of Economic Growth with Producer Heterogeneity

Christian Jensen*
University of South Carolina

February, 2014

Abstract

When producers are heterogeneous, the degree of competition between them does not only affect aggregate output via mark-ups and dead-weight losses, but also through aggregate productivity due to specialization. As competition tightens, high productivity producers gain market shares at the cost of low productivity ones, generating economic growth through increased aggregate productivity and capital accumulation, in line with what is observed empirically. Consequently, competition is not limited to reducing dead-weight losses, and can play a greater role in economic growth and development than traditionally considered. Economic growth spurs profits, which leads to entry and increased competition that generates growth, so competition provides a channel through which the economy generates growth internally. When strong enough, this channel can make the returns to scale in the inputs that the economy accumulates endogenously go from being decreasing to nondecreasing at the aggregate level, thus enabling endogenous growth. In fact, these returns to scale are determined endogenously in our model, and vary with the scale of production, the degree of producer heterogeneity and the barriers to entry.

JEL Codes: O40; O43
Keywords: Competition; Specialization; Economic growth; Producer heterogeneity; Returns to scale

*Department of Economics; 1705 College Street; Columbia, SC 29208. cjensen@alumni.cmu.edu. 803-777-2786.

1 Introduction

Competition is usually not considered a source of long-term economic growth, but merely a factor that can affect how much growth is generated by other sources. For example, Romer (1987, 1990), Grossman and Helpman (1991), and Aghion and Howitt (1992) stress the role market power and profits play in providing incentives for the innovations that drive the technological improvements typically considered to be the main engine of growth. On the other hand, Nickell (1996), Blundell, Griffith and Van Reenen (1999) and Aghion et al (2005) provide evidence that greater competition may encourage technological progress. The present study asks if competition can play a greater role in explaining economic growth, in particular whether it can contribute toward generating the rise in aggregate productivity that empirically tends to accompany economic development, even when it does not affect the rate of technological innovation. We find it can, when productivity is heterogeneous across producers, since increased competition then makes high-productivity producers gain market shares from those with low productivity, thereby raising aggregate productivity through increased specialization. Moreover, this effect can be so strong as to make the returns to scale in the inputs the economy accumulates endogenously go from being decreasing to increasing, at the aggregate level, thus enabling endogenous growth.

One of the fundamental propositions of economic theory is that market power limits production by imposing mark-ups that push prices above marginal costs. As competition intensifies, the demand for each particular good becomes more sensitive to its price, and profit-maximizing producers find it optimal to lower mark-ups, thus reducing dead-weight losses and raising production. Once mark-ups are zero, though, production growth from reducing dead-weight losses comes to a halt, so competition is typically deemed able to produce growth-spurts, but not sustained growth. Furthermore, reducing dead-weight losses raises production by increasing the quantity of inputs used, leaving their productivity unchanged, so competition increasing over time cannot explain why growth in aggregate productivity and income go hand-in-hand empirically, as described by Solow (1957). However, with heterogeneous producers, aggregate productivity does rise with competition, due to increased specialization, so its effects are not limited to reducing dead-weight losses.

Competition is driven by rent-seeking. Profits attract new producers, and their entry leads to more brands to choose among. Consequently, brands become less distinct, and consumers become more willing to substitute between these, lowering the market power of each producer. Hence, with free entry, one should expect competition to intensify as long as profits are positive. As the economy grows, demand increases, making markets expand and profits rise, which in turn leads to increased competition and economic growth.

As a result, competition can generate growth from within the economy, just as physical capital does in the Solow (1956) model, thereby magnifying the impact of all engines of growth. This channel is particularly important with producer heterogeneity, which amplifies its impact by raising aggregate productivity when competition increases. When strong enough, it can make the returns to scale in the inputs the economy accumulates endogenously (capital) be nondecreasing in terms of aggregate output, even if they are decreasing for individual producers.

The returns to scale in endogenously accumulated inputs are key for whether or not an economy can keep growing endogenously.[1] However, despite their crucial role, endogenous-growth models impose the necessary nondecreasing returns by assumption. Instead, the returns to scale are determined by the model itself in the present framework, moreover, they vary with the scale of production. As capital is accumulated, the aggregate returns to capital, taking into account the intensification of competition and boost to aggregate productivity, go from being decreasing to increasing. This switch happens sooner the higher the degree of competition, that is, the lower the barriers to entry. Consequently, our model can explain not only why different countries experience different growth trajectories, but also how the growth pattern of each of these can change over time. We find that a country with high barriers to entry will not only be poorer and experience a lower growth rate than an identical country with lower barriers, but that it could even stagnate, while that with low barriers keeps on growing endogenously.

Empirically, we have no evidence that economic growth has been accompanied by lower mark-ups historically, nor that these are lower in developed economies than in underdeveloped ones. However, there is little doubt that market consolidation and specialization in production have accompanied development, and is a significant difference between rich and poor countries. Production in poor countries tends to be decentralized with many small independent producers with varying degree of productivity, while in rich countries it is highly concentrated and specialized. For example, Comin and Hobijin (2004) and Banerjee and Duflo (2005) suggest that differences in total factor productivity across countries arise because the share of unproductive producers is greater in poor countries. Dollar (1992), Sachs and Warner (1995), and Frankel and Romer (1999) provide empirical evidence for openness to international trade, and hence international competition and specialization among producers, being associated with higher growth rates, (see Rodriguez and Rodrik (2000) for a critique of this evidence). One should expect the same to apply for competition and specialization within a country. Empirical evidence for barriers to entry being associated with lower levels and growth rates of

[1]Returns to scale also have important implications for the possible causes of business cycles, the amplification of the shocks generating these, and for the measurement of Solow residuals, see Benhabib and Farmer (1994), Cole and Ohanian (1999) and Deveraux, Head and Lapham (1996).

output per worker and productivity is provided by Barseghyan (2008), Nickell (1996) and Nicoletti and Scarpetta (2003). Hopenhayn (1992) provides a theoretical explanation of this based on entry costs protecting incumbent producers, leading to lower firm-level productivity. Productivity differences across producers, even withing the same industry or sector, are widely documented, see Bartelsman and Doms (2000) for a survey. That increased specialization can make aggregate productivity rise in the presence of producer heterogeneity has been exploited in trade and industry theory, see for example Hopenhayn (1992), Bernard et. al. (2003) and Melitz (2003). This strand of the literature, and the model in the present paper, differ from that on growth and specialization, such as Romer (1987) and Grossman and Helpman (1991), in that it is based on producer heterogeneity, instead of a preference for variety (Ethier (1982)). That is, aggregate productivity does not rise as a result of increased specialization that results in a larger variety of goods, but rather specialization in the sense that the most efficient producers become more dominant in production. Hence, our results rely on producer heterogeneity and increased competition between goods, not on a preference for variety. Romer (1986) provides empirical evidence for economies growing with nondecreasing, and even increasing, returns to scale. Caballero and Lyons (1992) argue that the statistical evidence is stronger for increasing returns to scale in aggregate production than at the disaggregate level. While market power has received attention in the literature, the main focus has been on its effects on innovation dynamics, or growth generated by other sources, not on competition as an engine of growth in its own right.

We show that endogenous economic growth is possible in a closed economy without technological progress, knowledge accumulation or product innovation, even when returns to scale are originally decreasing in the factors of production that the economy accumulates endogenously, through increased competition. We do so not because we doubt the importance of other sources of growth, or to suggest that these are less significant. Instead, we seek to illustrate that competition can also play a role, a more important role than typically considered, when one takes into account that producers are heterogeneous. In particular, it can be a relatively easy and immediate route to growth, as it does not require resources to import or develop innovations and human capital. This claim is supported by the fact that struggling economies are typically advised to undertake market reforms to enhance competition in order to generate swift economic growth. Such reforms do, however, involve removing barriers to entry, raising market consolidation and striking down monopolies, which can be difficult enough. Struggling countries' typical reluctance to undertake the recommended market reforms is evidence of this.

Building on the work of Dixit and Stiglitz (1977), Romer (1987) and Grossman and Helpman (1991), our model has of an infinite number of differentiated intermediate goods that households acquire to compose final goods. The producers of intermediate goods, which compete monopolistically, rent capital, labor and land from households in competitive factor markets. Producers are subject to idiosyncratic productivity shocks that generate the heterogeneity that is key in order for competition to have an impact on aggregate total factor productivity. The production side of the economy is presented in the next section, followed by a section describing households, which solve a standard intertemporal consumption problem. The subsequent two sections describe the equilibrium conditions, and producer heterogeneity and aggregation, respectively. We then show how producer heterogeneity makes aggregate productivity rise with competition, and how this channel can make aggregate returns to capital become nondecreasing, despite being decreasing for each individual producer. We conclude that there is greater scope for competition to generate economic growth than traditionally considered when one takes into account that productivity differs across producers.

2 Production

Imagine a continuum of measure one of identical households, indexed by $j \in [0,1]$, each producing y_j units of final good by combining a continuum of measure $I > 1$ of differentiated intermediate goods in quantities x_{ij}, where $i \in [0, I]$. Households use the technology

$$y_j = \left(I^{-\frac{1}{\varepsilon}} \int_0^I \left(e^{\gamma_i} x_{ij}\right)^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}} \tag{1}$$

where $\varepsilon > 1$ is the elasticity of substitution between any two intermediate goods, while γ_i are idiosyncratic shocks to preferences, or technology, that affect the relative weight of each intermediate good in the production of final good. The elasticity ε determines the degree of competition between differentiated goods, and thus the market power of each intermediate-good producer, and the overall degree of competition in the economy. We imagine that the larger the measure of producers I is, the less distinct each type of intermediate good is, and the easier it is to substitute between these. As a result, competition becomes fiercer the larger the measure of producers I, so that ε is increasing in I. Apart from the idiosyncratic shocks, the production function (1) is a standard Dixit-Stiglitz (1977) aggregator. The term $I^{-1/\varepsilon}$ is required so that aggregate productivity is not increasing in the measure of producers I, unless accompanied by increased competition ε, that is, it is required so that the model does not feature a preference for variety.

Assuming intermediate goods are the only inputs required to produce final goods, at any point in time each household j chooses the optimal mix of these so as to minimize the cost of provisioning the final good by solving

$$\min_{[x_{ij}]_{i=0}^{I}} \int_0^I P_i x_{ij} di \tag{2}$$

subject to the production function (1), where P_i is the price of intermediate good i. The resulting demand for intermediate good i from household j is

$$x_{ij} = \left(\frac{P_i}{P}\right)^{-\varepsilon} e^{(\varepsilon-1)\gamma_i} I^{-1} y_j \tag{3}$$

where P is the marginal cost of producing the final good. Because all households are identical, they compose identical final goods at identical cost, and since the market for final good is perfectly competitive, its market price must equal its marginal cost of production. Combined with the fact that the production technology (1) satisfies constant returns to scale, this implies that profits in final good production $Py_j - \int_0^I P_i x_{ij} di$ must equal zero (Shaw, Chang and Lai (2006)), so that its price and marginal cost are given by

$$P = \frac{\int_0^I P_i x_{ij} di}{y_j} = \left(I^{-1} \int_0^I \left(e^{-\gamma_i} P_i\right)^{1-\varepsilon} di\right)^{\frac{1}{1-\varepsilon}}. \tag{4}$$

Integrating intermediate-good demands (3) across all the identical households yields the aggregate demand for intermediate good i,

$$X_i \equiv \int_0^1 x_{ij} dj = \left(\frac{P_i}{P}\right)^{-\varepsilon} e^{(\varepsilon-1)\gamma_i} I^{-1} Y \tag{5}$$

where $Y \equiv \int_0^1 y_j dj$ is the aggregate demand for final goods.

Intermediate-good producer i finds the optimal mix of inputs, capital k_i, labor n_i and land l_i, by minimizing their total cost, solving

$$\min_{k_i, n_i, l_i} Rk_i + Wn_i + Fl_i \tag{6}$$

subject to its Cobb-Douglas production function

$$X_i = e^{z_i} k_i^{\alpha} n_i^{1-\alpha-\nu} l_i^{\nu} \tag{7}$$

where W is the wage, R is the rental rate of capital, F is the rental rate of land, $\alpha \in (0,1)$, $\nu \in (0,1)$, $1-\alpha-\nu \in (0,1)$, and z_i represents the technology used by producer i. The resulting first-order conditions yield the factor demands

$$k_i = \alpha \frac{\lambda_i X_i}{R}, \tag{8}$$

$$n_i = (1-\alpha-\nu)\frac{\lambda_i X_i}{W}, \tag{9}$$

$$l_i = \nu \frac{\lambda_i X_i}{F}, \tag{10}$$

where

$$\lambda_i \equiv e^{-z_i}\left(\frac{R}{\alpha}\right)^{\alpha}\left(\frac{W}{1-\alpha-\nu}\right)^{1-\alpha-\nu}\left(\frac{F}{\nu}\right)^{\nu} \tag{11}$$

is the marginal cost of producing intermediate good i.

Producer i must also price its good, and does so by choosing the price P_i that maximizes its profits given the demand (5) it faces, and thus solves

$$\max_{P_i} \Pi_i = (P_i - \lambda_i)\left(\frac{P_i}{P}\right)^{-\varepsilon} e^{(\varepsilon-1)\gamma_i} I^{-1} Y. \tag{12}$$

Profit maximization yields

$$P_i = \frac{\varepsilon}{\varepsilon-1}\lambda_i \tag{13}$$

the usual gross mark-up $\varepsilon/(\varepsilon-1) \in (1,\infty)$. Substituting for the marginal cost (11) and inserting into the price aggregator (4) yields

$$P = \left(\frac{R}{\alpha}\right)^{\alpha}\left(\frac{W}{1-\alpha-\nu}\right)^{1-\alpha-\nu}\left(\frac{F}{\nu}\right)^{\nu}\frac{\varepsilon}{\varepsilon-1}\left(I^{-1}\int_0^I e^{(\varepsilon-1)(\gamma_i+z_i)}di\right)^{\frac{1}{1-\varepsilon}} \tag{14}$$

and thus the relative price

$$\frac{P_i}{P} = \left(I^{-1}\int_0^I e^{(\varepsilon-1)(\gamma_i+z_i)}di\right)^{\frac{1}{\varepsilon-1}} e^{-z_i}. \tag{15}$$

Inserting this relative price (15) into the demand function (5) for intermediate good i, then inserting the resulting equation and the marginal cost of production (11) into the factor demands (8), (9) and (10), and integrating across producers, yields the aggregate demands for capital, labor and land,

$$K = \left(\frac{R}{\alpha}\right)^{\alpha-1}\left(\frac{W}{1-\alpha-\nu}\right)^{1-\alpha-\nu}\left(\frac{F}{\nu}\right)^{\nu} Y\left(I^{-1}\int_0^I e^{(\varepsilon-1)(\gamma_i+z_i)}di\right)^{\frac{1}{1-\varepsilon}}, \tag{16}$$

$$N=\left(\frac{R}{\alpha}\right)^{\alpha}\left(\frac{W}{1-\alpha-\nu}\right)^{-\alpha-\nu}\left(\frac{F}{\nu}\right)^{\nu}Y\left(I^{-1}\int_{0}^{I}e^{(\varepsilon-1)(\gamma_i+z_i)}di\right)^{\frac{1}{1-\varepsilon}}, \tag{17}$$

$$L=\left(\frac{R}{\alpha}\right)^{\alpha}\left(\frac{W}{1-\alpha-\nu}\right)^{1-\alpha-\nu}\left(\frac{F}{\nu}\right)^{\nu-1}Y\left(I^{-1}\int_{0}^{I}e^{(\varepsilon-1)(\gamma_i+z_i)}di\right)^{\frac{1}{1-\varepsilon}}, \tag{18}$$

respectively, where $K \equiv \int_0^I k_i di$, $N \equiv \int_0^I n_i di$ and $L \equiv \int_0^I l_i di$. Without loss of generality, we let the final good be numeraire, so that $P \equiv 1$.

3 Households

In addition to effortlessly composing final goods, households rent labor N, capital K and land L to intermediate-good producers in order to provide for consumption C and the accumulation of capital. Because households are assumed to be identical, aggregation is trivial, so we focus on aggregates directly. In order to simplify, labor and land are assumed to be supplied inelastically, with their supplies normalized to N and one, respectively.[2] Given these assumptions, households seek to maximize the discounted lifetime-utility of consumption

$$\int_0^\infty e^{-\rho t}\frac{C(t)^{1-\theta}-1}{1-\theta}dt \tag{19}$$

subject to the budget constraint

$$\dot{K}+C=WN+RK+F-\delta K+\Pi \tag{20}$$

with respect to the control C and the state K, given a constant relative risk-aversion parameter $\theta>0$, discount rate $\rho\in(0,1)$, capital depreciation rate $\delta\in(0,1)$, and initial condition $K(0)>0$. Here, $\Pi=\int_0^I \Pi_i di$ denotes the profits generated in the production of intermediate goods. The first-order conditions yield

$$\frac{\dot{C}}{C}=\frac{R-\delta-\rho}{\theta} \tag{21}$$

the usual requirement for the optimal consumption path.[3]

[2] Including land as an inelastically supplied input makes it easier to derive the aggregate production function, since otherwise the production side only pins down the factor mix, not the levels.

[3] In addition, we have the transversality condition $\lim_{t\to\infty} e^{-\rho t}C(t)^{-\theta}K(t)$, which is also standard.

4 Equilibrium

In addition to effortlessly composing final goods, households rent labor, capital and land to intermediate-good producers to provide for consumption C and the accumulation of capital. Setting the aggregate demand for land (18) equal to its inelastic unitary supply yields

$$Y = \left(I^{-1} \int_0^I e^{(\varepsilon-1)(\gamma_i+z_i)} di \right)^{\frac{1}{\varepsilon-1}} K^\alpha N^{1-\alpha-\nu} \tag{22}$$

after exploiting that the aggregate demands for factors of production (16)-(18) imply $R/F = \alpha/(\nu K)$ and $W/F = (1-\alpha-\nu)/(\nu N)$, which guarantee an optimal mix of capital, labor and land in the production of intermediate goods. Combining these two conditions with the one for the price level (14), and exploiting that the final good is numeraire, yields

$$R = \alpha \left(I^{-1} \int_0^I e^{(\varepsilon-1)(\gamma_i+z_i)} di \right)^{\frac{1}{\varepsilon-1}} K^{\alpha-1} N^{1-\alpha-\nu} \frac{\varepsilon-1}{\varepsilon}, \tag{23}$$

$$W = (1-\alpha-\nu) \left(I^{-1} \int_0^I e^{(\varepsilon-1)(\gamma_i+z_i)} di \right)^{\frac{1}{\varepsilon-1}} K^{\alpha} N^{-\alpha-\nu} \frac{\varepsilon-1}{\varepsilon}, \tag{24}$$

$$F = \nu \left(I^{-1} \int_0^I e^{(\varepsilon-1)(\gamma_i+z_i)} di \right)^{\frac{1}{\varepsilon-1}} K^{\alpha} N^{1-\alpha-\nu} \frac{\varepsilon-1}{\varepsilon}, \tag{25}$$

the equilibrium wage and rental rates.

From aggregate production (22) we have aggregate total factor productivity

$$A \equiv \left(I^{-1} \int_0^I e^{(\varepsilon-1)(\gamma_i+z_i)} di \right)^{\frac{1}{\varepsilon-1}} \tag{26}$$

which determines how efficiently labor and capital are converted into final goods. There are two stages in this process, inputs producing intermediate goods, and intermediate goods producing final goods, so total factor productivity depends on the efficiency with which each of these two stages is carried out, which is a function of the two shocks z_i and γ_i, respectively. From above, it is clear that these two shocks are perfect substitutes in terms of aggregate variables, since total factor productivity depends on their sum.

5 Heterogeneity

It follows from the generalized mean inequality that total factor productivity A, defined above (26), is increasing in ε, as long as producers are heterogeneous ($\gamma_i + z_i$ varies across producers).[4] Intuitively, increased competition ε leads to greater substitution between low and high productivity producers, thus raising aggregate productivity. While any non-degenerate distribution of heterogeneity makes total factor productivity increase with ε, assume the particular example where $\gamma_i + z_i$ for all $i \in [0, I]$ is a collection of independent and identically distributed random variables such that conditional on ε, the expected value and variance of $\exp((\varepsilon - 1)(\gamma_i + z_i))$ are both finite. In this case, Uhlig (1996) shows that conditional on ε,

$$I^{-1} \int_0^I e^{(\varepsilon-1)(\gamma_i+z_i)} di = E\left(e^{(\varepsilon-1)(\gamma_i+z_i)}\right) \tag{27}$$

holds due to the law of large numbers, where the right-hand-side is the moment-generating function of the random variable $\gamma_i + z_i$. Assuming furthermore that $\gamma_i + z_i$ is Normally distributed with mean $-\sigma^2/2$ and variance σ^2, so that $E(\exp(\gamma_i + z_i)) = 1$ is not increasing in heterogeneity σ, we have

$$E\left(e^{(\varepsilon-1)(\gamma_i+z_i)}\right) = e^{\frac{\sigma^2}{2}\left((\varepsilon-1)^2-(\varepsilon-1)\right)} \tag{28}$$

and aggregate total factor productivity

$$A = e^{\frac{\sigma^2}{2}(\varepsilon-2)} \tag{29}$$

for all $\varepsilon > 1$.[5] As a result, we have aggregate production

$$Y = e^{\frac{\sigma^2}{2}(\varepsilon-2)} K^\alpha N^{1-\alpha-\nu} \tag{30}$$

and factor prices

$$R = \alpha e^{\frac{\sigma^2}{2}(\varepsilon-2)} K^{\alpha-1} N^{1-\alpha-\nu} \frac{\varepsilon-1}{\varepsilon}, \tag{31}$$

$$W = (1-\alpha-\nu)\, e^{\frac{\sigma^2}{2}(\varepsilon-2)} K^{\alpha} N^{-\alpha-\nu} \frac{\varepsilon-1}{\varepsilon}, \tag{32}$$

$$F = \nu e^{\frac{\sigma^2}{2}(\varepsilon-2)} K^{\alpha} N^{1-\alpha-\nu} \frac{\varepsilon-1}{\varepsilon}, \tag{33}$$

for capital, labor and land, respectively.

---

[4] The generalized mean inequality states that if $q < m$, then $(\int_0^I r_i^q di)^{1/q} \leq (\int_0^I r_i^m di)^{1/m}$, and the two are equal if and only if $r_i = r$ for all i, for any positive real numbers r_i and real $q \neq 0$, with $I > 0$. See for example Hardy, Littlewood and Polya (1952).

[5] If technology improves over time so that the mean of $\gamma_i + z_i$ increases, aggregate total factor productivity will grow even if ε remains constant. Since our focus is on the effects of changes in competition, we imagine $\gamma_i + z_i$ has a constant mean. We assume a mean of $-\sigma^2/2$ so that the average productivity of individual firms does not increase with heterogeneity σ.

6 Competition-induced growth

Competition reduces mark-ups and the dead-weight losses these generate, raising production by employing greater quantities of inputs. This traditional effect does not rely on producer heterogeneity, and can be seen in the factor prices (31)-(33), which rise toward their respective marginal products as ε increases. With producer heterogeneity, however, competition further boosts production by raising aggregate total factor productivity (29). As competition increases, substitution between intermediate goods becomes easier, so more of the low productivity goods are substituted with high productivity ones, thus raising the amount of final goods that can be produced for any quantity of labor and capital. This effect on aggregate productivity is greater the more heterogeneous productivity is among intermediate-good producers, that is, the greater σ is, since that means there is more to be gained from substituting between these.[6] By raising aggregate productivity, competition has a greater impact on the economy, and is not limited to eliminating dead-weight losses.

In an economy as the one described above, with no innovations in products or production techniques that allow a producer to maintain the uniqueness of her product, one would expect profits to attract new producers, which would in turn make competition intensify. Inserting for the profit-maximizing pricing equation (13), relative price (15), the marginal cost of production (11) and the equilibrium factor prices (31)-(33) into producer i's profit function (12) yields

$$\Pi_i = \frac{Y}{\varepsilon}\left(I^{-1}\int_0^I e^{(\varepsilon-1)(\gamma_i+z_i)}di\right)^{-1} I^{-1}e^{(\varepsilon-1)(\gamma_i+z_i)} \tag{34}$$

which is always positive, since producers apply a positive mark-up. With a fixed cost $\phi > 0$ of staying in business (lump-sum transfers to households directly through marketing efforts, or indirectly through the government as licenses), producers enter or exit until their expected profits equal this fixed cost, so that

$$E\left(\Pi_i\right) = \int_0^I \Pi_i di = \frac{Y}{\varepsilon} = \phi \tag{35}$$

assuming they decide whether or not to stay in business prior to learning their idiosyncratic shocks.[7] Inserting for aggregate output (30) in the entry condition (35) and rearranging

[6] The effect of competition on aggregate productivity does not depend on producer heterogeneity σ increasing with entry, so it is here assumed to be independent of the measure of producers I and the degree of competition ε. If heterogeneity increased with entry, there would be an additional increase in aggregate total factor productivity.

[7] If producers knew their idiosyncratic shocks before deciding to produce, entry and exit would affect the distribution of productivity of those that ended up producing.

yields

$$-\frac{\sigma^2}{2\phi}e^{-\sigma^2}K^\alpha N^{1-\alpha-\nu} = -\frac{\sigma^2}{2}\varepsilon e^{-\frac{\sigma^2}{2}\varepsilon} \tag{36}$$

which has the solution

$$\varepsilon = -\frac{2}{\sigma^2}\,\Omega\left(-\frac{\sigma^2}{2\phi}e^{-\sigma^2}K^\alpha N^{1-\alpha-\nu}\right) \tag{37}$$

where Ω is the Omega function.[8] Substituting into the production function (30) yields

$$Y = e^{-\Omega\left(-\frac{\sigma^2}{2\phi}e^{-\sigma^2}K^\alpha N^{1-\alpha-\nu}\right)-\sigma^2}K^\alpha N^{1-\alpha-\nu} = -\frac{2\phi}{\sigma^2}\Omega\left(-\frac{\sigma^2}{2\phi}e^{-\sigma^2}K^\alpha N^{1-\alpha-\nu}\right) \tag{38}$$

where the last equality follows from the entry condition (35).

For $\omega > 0$, $-\Omega(-\omega)$ is strictly increasing and convex in ω, while $\omega = 2^{-1}\phi^{-1}\sigma^2 exp(-\sigma^2)$ $K^\alpha N^{1-\alpha-\nu}$ is strictly increasing and concave in K, so Y is strictly increasing in K, but can be concave or convex. In fact, computing the second-order derivative

$$\frac{\partial^2 Y}{\partial K^2} = -2\alpha\phi\Omega\left(-\eta K^\alpha\right)\frac{\alpha-1-\left(2+\Omega\left(-\eta K^\alpha\right)\right)\Omega\left(-\eta K^\alpha\right)}{\sigma^2\left(1+\Omega\left(-\eta K^\alpha\right)\right)^3 K^2} \tag{39}$$

where $\eta = \frac{\sigma^2}{2\phi}e^{-\sigma^2}N^{1-\alpha-\nu}$, it turns out that Y is strictly concave in K for $0 < K < K^*$ and strictly convex for $K > K^*$, as is illustrated in figure 1, where

$$K^* \equiv e^{\frac{\sqrt{\alpha}-1}{\alpha}}\left(\frac{2\phi\left(1-\sqrt{\alpha}\right)}{\sigma^2 N^{1-\alpha-\nu}}e^{\sigma^2}\right)^{\frac{1}{\alpha}} \tag{40}$$

which can be obtained by solving $\partial^2 Y/\partial K^2 = 0$.[9] The aggregate returns to scale in capital are determined by two opposing effects. The first is the standard diminishing returns to capital that arise with Cobb-Douglas production functions. The second effect, which relies on heterogeneous producers and endogenous entry, makes aggregate total factor productivity rise with capital accumulation, as a result of increased specialization due to entry driven by profits, which are increasing in aggregate output, and thus in capital. Because total factor productivity grows exponentially as ε rises (29), this effect through

[8] The Omega function, also called the Lambert W function and the product logarithm, is the inverse relation of the function $g(\omega) = \omega e^\omega$, so $\Omega(\omega)e^{\Omega(\omega)} = \omega$. It has no representation in terms of elementary functions, but can be approximated numerically, as discussed in Corless et al. (1996). For $\omega < 0$ it is a multivalued relation, and thus not really a function, with an upper (principal) branch denoted Ω and a lower branch denoted Ω_{-1}. We use the upper branch solution since the lower one would make output decreasing in input use. The function's derivative $\Omega'(\omega) = \Omega(\omega)/(\omega(1+\Omega(\omega)))$ for $\omega \neq \{0, e^{-1}\}$.

[9] Output Y is not defined for $K > (2\phi/(\sigma^2 N^{1-\alpha-\nu})\exp(\sigma^2-1))^{1/\alpha} > K^*$, since $\Omega(-\omega)$ is not real-valued for $\omega > e^{-1}$, meaning that there is no real-valued ε that solves the entry condition (36).

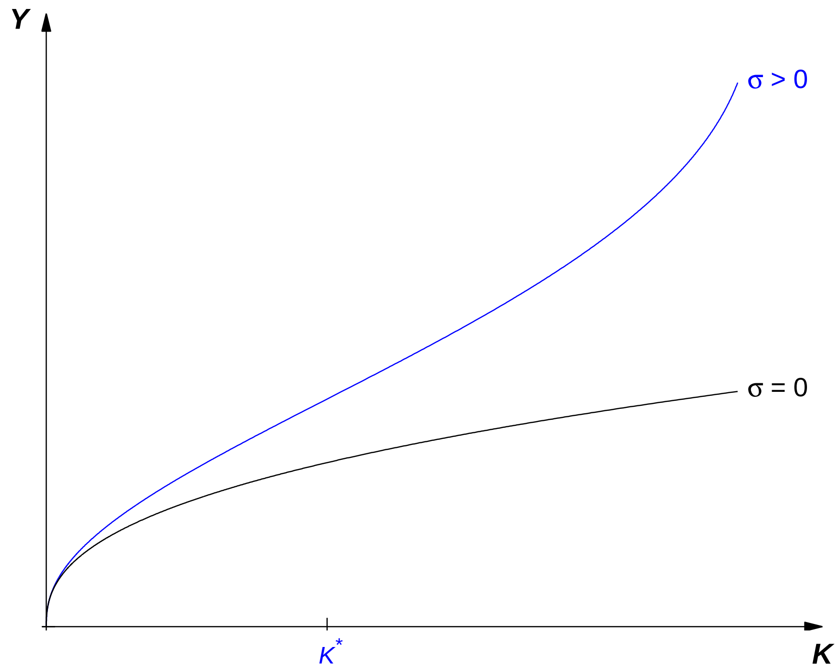

Figure 1: Aggregate production as a function of capital with and without heterogeneity.

competition eventually makes the returns to capital go from being decreasing to increasing, as is illustrated in the figure. This occurs for all $\alpha \in (0,1)$, so returns to capital eventually become increasing in terms of aggregate output no matter how decreasing they are from the perspective of each individual producer. However, the threshold K^* is lower the smaller ϕ is, as this makes the degree of competition ε larger and more responsive to changes in aggregate output.[10] In the limit case where ϕ approaches zero, returns to capital become increasing for all $K > 0$. Hence, barriers to entry and the degree of competition can determine whether there are decreasing or increasing returns to capital at the aggregate level, and thus whether or not the economy can keep on growing endogenously simply by accumulating capital.

As usual, households' intertemporal consumption decision (21) makes the growth rate of consumption be determined by the rental rate of capital R, in addition to the parameters for risk aversion θ, discount rate ρ and depreciation rate δ. Inserting the solution (37) for ε into the equilibrium rental rate (31) yields

$$R = -\alpha\phi\left(\frac{2}{\sigma^2}\Omega\left(-\frac{\sigma^2}{2\phi}e^{-\sigma^2}K^{\alpha}N^{1-\alpha-\nu}\right)+1\right)K^{-1} = \alpha\left(\frac{Y}{K}-\frac{\phi}{K}\right) \tag{41}$$

[10]Raising the labor input N reduces the threshold K^*, as it also makes ε larger and more responsive to changes in the aggregate capital stock. Greater heterogeneity σ can push the threshold either way.

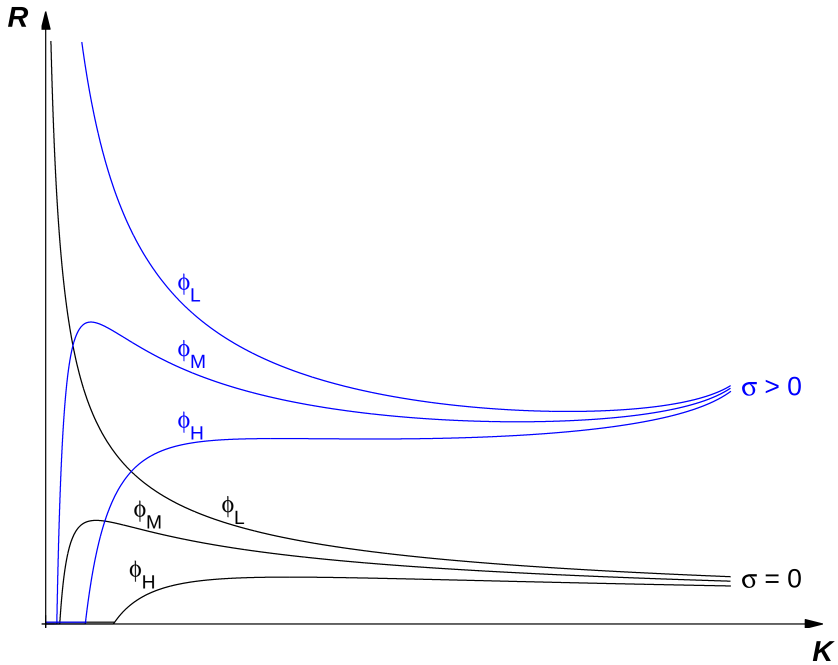

Figure 2: Rental rate as a function of capital with and without heterogeneity.

where the last equality follows from the solution for output (38), and shows how barriers to entry (ϕ) lower the rental rate, and hence the growth rate of consumption.[11] Figure 2 plots several possible paths for the rental rate, for low, medium and high barriers to entry (ϕ_L, ϕ_M and ϕ_H respectively). For any given $\phi > 0$, $Y/K - \phi/K$ is negative for $K < \phi^{1/\alpha} N^{(\alpha+\nu-1)/\alpha} \exp(\sigma^2/(2\alpha))$. However, since the marginal product of capital is always positive, its rental rate (31) can only be negative if $\varepsilon < 1$. This is not allowed in our model, as it would make mark-ups and profits negative, so intermediate-good producers would stop producing. Consequently, figure 2 assumes that the rental rate of capital is zero whenever the capital stock is so low that $Y - \phi$ is negative (though the model is not really defined for $\varepsilon < 1$).

As capital is accumulated, its marginal product declines for each individual producer, which puts a downward pressure on its rental rate. At the same time competition increases, lowering mark-ups and raising aggregate total factor productivity, both of which contribute to raise the rental rate of capital. Depending on the relative sizes of these effects, which depends on ϕ, the rental rate can be increasing or decreasing in the capital stock. However, since the only effect that does not diminish as capital is accumulated is the one on aggregate productivity, it eventually comes to dominate, so that the rental rate becomes increasing

[11] This applies also without heterogeneity, since then $R = \alpha K^{\alpha-1} N^{1-\alpha-\nu} - \phi/K = \alpha(Y/K - \phi/K)$.

in capital for a large enough capital stock, thus preventing it from converging to zero. For comparison, figure 2 includes paths for the rental rate when there is no heterogeneity ($\sigma = 0$), which can also be increasing or decreasing in capital, depending on whether the effect of a falling marginal product or falling mark-up dominates, which in turn depends on ϕ. However, when competition does not affect aggregate productivity, eventually the rental rate converges toward zero. Hence, without heterogeneity, consumption growth eventually comes to a stop, or becomes negative, for a high enough level of capital accumulation, but this is not necessarily the case with heterogeneity. In particular, when ϕ is low enough for the rental rate never to fall below $\delta + \rho$, the consumption growth rate (21) is positive for all capital levels, which implies that the economy itself must be growing, which it can only do endogenously by accumulating capital (in the present model). Intuitively, if the rental rate of capital never falls below the threshold at which households stop raising the capital stock, because its accumulation raises competition and hence aggregate productivity, the capital stock, and the economy, can keep on growing endogenously.

Not only do barriers to entry lower the rental rate, and hence the consumption growth rate and incentives to accumulate capital, but in addition they can halt consumption growth long before the decreasing returns to capital become an obstacle. The reason is that as competition approaches its lower limit $\varepsilon \to 1$, the rental rate (31) converges toward zero no matter what the marginal product of capital is. Hence, when the barriers to entry are high enough, the rental rate can be too low to support any positive consumption growth no matter how low the capital stock is, since competition is decreasing in expected profits and therefore also in physical capital. As figure 2 shows, for high enough values of ϕ, the rental rate will be lower with a small capital stock than with a large one, implying that capital accumulation and growth could come to a halt long before they really take off. This applies whether or not there is heterogeneity, that is, whether or not competition affects the degree of competition.

Aggregate total factor productivity

$$A = e^{-\Omega\left(-\frac{\sigma^2}{2\phi}e^{-\sigma^2}K^{\alpha}N^{1-\alpha-\nu}\right)-\sigma^2} \tag{42}$$

is increasing in capital K for $\sigma > 0$, first concave then convex (plotted as a function of capital it looks similar to the plot of aggregate output in figure 1). This switch in the second-order derivative is what makes aggregate output switch from being concave to convex in capital, a property that comes from total factor productivity growing exponentially as a function of the degree of competition (29). Without heterogeneity, $A = 1$ and $Y = K^{\alpha}N^{1-\alpha-\nu}$, since $\Omega(0) = 0$, and aggregate output would always be concave in capital, as in the Solow (1956) model. With heterogeneity, competition works as an externality. Because producers take entry and the degree of competition to be independent of their individual actions, they perceive the marginal product of each of the inputs to be decreasing in the quantity of the input itself, even when it is constant or increasing for the economy as a whole. Likewise, households, who take the real rental rate as given, ignore the effect capital accumulation has on aggregate total factor productivity through entry and competition. Hence, this effect is a pure externality that does not affect individual behavior, nor their optimal decisions (as in the endogenous-growth models of Romer (1986) and Lucas (1988)).

7 Conclusion

Explaining differences in the level and growth rate of income across countries to a great extent boils down to explaining differences in the level and growth rate of aggregate productivity, since the two go hand-in-hand empirically. With the usual assumption of homogeneous producers, competition has no impact on aggregate total factor productivity, and therefore plays little or no role in accounting for differences in income and growth. However, with heterogeneous producers, competition does influence aggregate productivity, by affecting the degree to which high-productivity producers substitute low-productivity ones. As a result, countries with dissimilar degrees of domestic competition can experience important differences in aggregate productivity, despite using identical production technologies. Moreover, we find that barriers to entry and the degree of competition can affect the returns to scale in aggregate production, and thus whether or not an economy can keep on growing simply through the endogenous accumulation of inputs. This illustrates the importance of having well-functioning institutions, markets and incentive mechanisms, in addition to low barriers to entry, so that profits attract competition, which in turn raises productivity and production. The degree of competition can also be affected by legislation, the enforcement of anti-trust measures, and other efforts by government or consumer advocacy groups, as well as actions taken by producers to distinguish their brand through marketing, or other efforts such as collusion, to lower the degree of competition. The availability of cheap transportation, communication and national or international markets can also affect barriers to entry and competition.

Our model assumes there are no innovations in products or production techniques, and free entry, so mark-ups fall as the economy grows. Obviously, the average mark-up would not systematically fall with the continuous development of proprietary innovations protected by patents or as trade secrets. Our stylized model ignores such innovations in order to isolate the effects of competition and the specialization in production that it leads to. In the limit ($\varepsilon \to \infty$), when all production is undertaken by the most efficient producer, growth from increased competition would come to a halt, since it is driven by the substitution from low to high productivity producers. However, with the development of new products and industries, the process would start over again, with the entry of new producers as patents expire, leading to increased competition and the gradual substitution towards the most efficient producers in the industry. Hence, while mark-ups fall as industries mature, the continuous development of new industries would prevent the average mark-up in the economy from declining.

While competition can raise aggregate productivity and contribute toward generating sustained economic growth, measuring its importance as a source of growth can be difficult. The reason is that, according to our model, its effects can be indistinguishable from those of technological improvements in terms of aggregate data. However, it should be possible to distinguish between the two using producer-level data, since competition makes aggregate productivity grow without affecting the productivity of each individual producer. Hence, to the extent that productivity has grown faster at the aggregate level than for individual producers, there is scope for competition to have played a role. Alternatively, the effects of increased competition can be observed through increased specialization. Our model features scale effects, in that large economies would be more prone to experiencing a higher degree of competition, nondecreasing returns to capital, and hence higher output and growth rates, contrary to the empirical evidence (Backus, Kehoe and Kehoe (1992) and Jones (1995)). The reason is that profits, and thus entry and competition, are increasing in both capital and labor. Of course, what matters is the size of the market that a producer can reach, so populous economies with fragmented markets would not necessarily have an advantage over less populated economies. Likewise, access to international markets means that the size of a country, or its population, becomes less relevant for the degree of competition.

8 References

Aghion, P., Bloom, N., Blundell, R., Griffith, R. and Howitt, P. (2005), "Competition and Innovation: An inverted-U Relationship," *Quarterly Journal of Economics* 120: 701-728.

Aghion P. and Howitt, P. (1992), "A Model of Growth through Creative Destruction," *Econometrica* 60: 323-351.

Backus, D., Kehoe, P. J. and Kehoe, T. J. (1992), "In Search of Scale Effects in Trade and Growth," *Journal of Economic Theory* 58: 337-409.

Banerjee, A. V. and Duo, E. (2005), "Growth theory through the lens of development," in P. Aghion and S. Durlauf (Eds.), *Handbook of economic growth,* North-Holland.

Barseghyan, L. (2008), "Entry costs and cross-country differences in productivity and output," *Journal of Economic Growth* 13: 145-167.

Bartelsman, E. J. and Doms, M. (2000), "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature* 38: 569-594.

Benhabib, J. and Farmer, R. E. A. (1994), "Indeterminacy and Increasing Returns," *Journal of Economic Theory* 63: 19-41.

Bernard, A. B., Eaton, J., Jensen, J. B. and Kortum, S. (2003), "Plants and Productivity in International Trade," *American Economic Review* 93: 1268-1290.

Blundell, R. Griffith, R. and Van Reenen, J. (1999), "Marketshare, Market Value, and Innovation in a Panel of British Manufacturing Firms," *Reveiw of Economic Studies* 56: 529-554.

Caballero, R. J. and Lyons, R. K. (1992), "External effects in U.S. procyclical productivity," *Journal of Monetary Economics* 29: 209-225.

Cole, H. and Ohanian, L. (1999), "Aggregate Returns to Scale: Why Measurement is Imprecise," *Federal Reserve Bank of Minneapolis Quarterly Review* 23: 19-28

Comin, D. and Hobijn, B. (2004). "Neoclassical growth and the adoption of technologies," *NBER Working Paper* 10733.

Devereux, M. B., Head, A. C. and Lapham, B. J. (1996), "Aggregate fluctuations with increasing returns to specialization and scale," *Journal of Economic Dynamics and Control* 20:627-656.

Dixit, A. K. and Stiglitz, J. E. (1977), "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67: 297-308.

Dollar, D. (1992) "Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976-1985," *Economic Development and Cultural Change* 40: 523-544.

Ethier, W. (1982), "National and international returns to scale in the modern theory of international trade," *American Economic Review* 72: 389-405

Frankel, J. and Romer, D. (1999), "Does Trade Cause Growth?" *American Economic Review* 89: 379-399.

Grossman G. M and Helpman, E. (1991), "Quality Ladders in the Theory of Growth," *Review of Economic Studies* 68: 43-61.

Hardy, G. H., Littlewood, J. E., and Polya, G. (1952), *Inequalities*, Cambridge University Press.

Hopenhayn, H. (1992), "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica* 60: 1127-1150.

Jones, C. I. (1995), "R&D-based Models of Economic Growth," *Journal of Political Economics* 103: 759-784.

Lucas, R. E. (1988), "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22: 3-42.

Melitz, M. J. (2003), "The Impact of Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica* 71: 1695-1725.

Nickell, S. (1996), "Competition and Corporate Performance," *Journal of Political Economy* 104: 724-746.

Nicoletti, G. and Scarpetta, S. (2003), "Regulation, productivity and growth: OECD evidence, " *Policy Research Working Paper Series, World Bank*.

Rodriguez, F. and Rodrik, D. (2000), "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," *NBER Macroeconomics Annual* 2000: 261-325.

Romer, P. M. (1986), "Increasing Returns and Long-Run Growth," *Journal of Political Economy* 94: 1002-1037.

Romer, P. M. (1987), "Growth Based on Increasing Returns Due to Specialization," *American Economic Review* 77: 56-62.

Romer, P. M. (1990), "Endogenous Technological Change," *Journal of Political Economy* 98: 71-102.

Sachs, J. and Warner, A. (1995), "Economic Reform in the Process of Global Integration," *Brookings Papers on Economic Activity* 1: 1-118.

Shaw, M. F., Chang, J. J. and Lai, C. C. (2006), "(Non)optimality of the Friedman Rule and Optimal Taxation in a Growing Economy with Imperfect Competition," *Economics Letters* 90: 412-420.

Solow, R. M. (1956), "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70: 65-94.

Solow, R. M. (1957), "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics* 39: 312-320.

Uhlig, H. (1996), "A law of large numbers for large economies," *Economic Theory* 8: 41-50.