Firm Growth and Selection in a Finite Economy

Staley, Mark

University of Ontario Institute of Technology

12 October 2015

Firm Growth and Selection in a Finite Economy

Mark Staley*

October 12, 2015

Preliminary version. Comments welcomed.

Abstract

A model of firm dynamics is presented in which the growth rate of knowledge capital is linked to productivity, and productivity fluctuates randomly. The distribution of productivity forms a stable traveling wave, representing a growing economy. Granularity is maintained by way of spinoffs, resulting in a firm size distribution that rapidly approaches the Zipf distribution. An unexpected consequence of the model is that the growth rate is proportional to the log of the number of firms. The model also implies that specialization is positive for growth.

Keywords: Growth, Ideas, Selection, Scale effect.

JEL codes: O40, O41

*Faculty of Science, University of Ontario Institute of Technology, 2000 Simcoe Street North, Oshawa, Ontario, L1H 7K4, Canada. Email: mark.staley@uoit.ca.

1 Introduction

One of the most striking observations about firm productivity is that it is very heterogeneous, even within a single industry. The aggregate productivity of the economy could be increased substantially if resources were reallocated from less-productive firms to more-productive firms (Restuccia & Rogerson, 2008, Hsieh & Klenow, 2009). Assuming that resources are in finite supply, it is possible to imagine forces of selection working to realize such a reallocation. For example, high-productivity firms may hire workers more rapidly than low-productivity firms because the former can more vigorously pursue "organic growth" strategies than the latter. However in order for this growth mechanism to be sustainable there must also be firm-level sources of noise maintaining the dispersion of productivity, else the fuel of selection will be exhausted.

In this paper, a model of growth is presented in which firms experience idiosyncratic shocks to productivity in the form of new ideas. Firms are selected based on how quickly those ideas can be translated into the skills and knowledge of new employees, or in other words how rapidly *knowledge capital* can be accumulated. The growth rate of a firm is an increasing function of its productivity. Occasionally, workers leave to start their own spinoff firms, taking knowledge with them. Ideas are nonrival, but it takes time for them to be copied, and they are confined within firms except when a spinoff occurs.

The core assumptions of the present model (and similar models) are that productivity is linked to people's knowledge, knowledge is subject to random shocks, and knowledge is copied at a rate that is increasing in its usefulness. The first two assumptions are straightforward but the third requires elaboration because it can be modeled in different ways. One approach is to assume passive imitation. In Staley (2011) and Luttmer (2012), people meet randomly and during each meeting the less-productive person copies the ideas of the more-productive person. This simple rule leads to a selection effect that favors the most productive ideas.[1] Another approach is to assume that search and imitation activities are costly or risky and that people or firms rationally decide what portion of their resources to allocate to them (König

[1] This model is based on models of idea flows developed by Alvarez et al (2008) and Lucas (2009).

et al., 2012, Lucas & Moll, 2014, Perla & Tonetti, 2014). In these approaches there are incentives to imitate and no incentives to be imitated, but there are also no costs to being imitated. The focus of this paper is on a different type of knowledge transfer: that between managers (teachers) and laborers (students). This type of learning requires that managers work alongside laborers and actively teach them. Managers also hire new laborers, which leads to firm growth. What is the incentive for such activity? One answer to that question, suggested below, leads us in the direction of a new growth model.

Since ancient times the primary mechanism for knowledge transfer was the master-apprentice relationship. The master was willing to teach the apprentice because the apprentice agreed to work for the master for a low wage for an extended period of time and there were other anti-competitive policies in place to protect the masters.[2] One can think of the present-day manager-worker relationship as the modern equivalent of the master-apprentice relationship. Even though the labour market is much freer than it was in past centuries, knowledge transfer still occurs. According to Lucas (1978) there are diminishing returns to labor and physical capital due to management's finite "span of control". In a competitive economy this implies that labor and physical capital do not capture all the output of the firm; the remainder is paid to owners and managers. Presumably this remainder can be the incentive to set up firms and train workers.

In the present model, span of control is captured as follows. There are two factors of production, labor L and knowledge capital K (a form of intangible capital), sometimes simply referred to as "capital" or "knowledge" in this paper. There is no physical capital. Firms combine the two factors according to $Y = F(AK, L)$, where A stands for productivity, and F is first-order homogeneous and concave in AK and L separately. The economy is perfectly competitive. The symbol A can be interpreted as the productivity of a set of ideas that are shared by the employees of the firm. The symbol K can be interpreted as a count of blueprints that contain the ideas, and blueprints are "read" by the employees. It is best to think of knowledge AK as separate from labor. Firm owners capture the rents from knowledge even though that knowledge resides in the heads of (some) employees.[3]

[2]Smith, 1776, chapter X part II. In the medieval era, associations of masters were called universities, for example the "University of Taylors".

[3]Knowledge capital is similar to organization capital (Prescott & Visscher, 1980). Atkeson & Kehoe (2005) distinguish organization capital from management expertise. Here I

Knowledge capital accumulates according to $\mathrm{d}K/\mathrm{d}t = \beta F$, where β is a constant. When labor is held fixed this dynamic process leads to increases in output that are subject to diminishing returns as in the familiar "learning curves" described by Arrow (1962). A depreciation term can also be added. In the present model labor is not held fixed. Firm owners maximize their return on capital by adjusting the amount of labor. As K becomes larger, L tends to follow suit. Finally, productivity is subject to geometric Brownian shocks: $\Delta A/A \sim N(\mu\,\Delta t, \sigma^2\,\Delta t)$, which are coordinated across the firm.[4] Together these components add up to a model as described in the first paragraph. The growth rate of each firm depends on its productivity, which fluctuates randomly.

The case of $\mu = 0$ is of particular interest, because in that case growth is dependent on random shocks only. If one sets up a simulation of the above model with some fixed number of firms and sets μ to zero, one typically observes that the economy grows for a while but then grinds to a halt. What happens is that one of the firms draws a very large productivity shock and grows much faster than the others. That firm proceeds to capture most of the labor force and becomes almost a monopoly. It continues to experience fluctuating productivity, but the fuel of selection is gone. What is missing from this model is a mechanism for generating new firms (entrants) and hence maintaining diversity of innovation shocks.[5]

There is an economic incentive for workers to leave their places of employment and set up their own spinoff firms because then they can earn rent on their previously un-compensated knowledge in addition to the market wage.[6] Of course if everyone decided to become an entrepreneur there would be no incentive for firms to grow. But most people do not become entrepreneurs.

abstract from this separation and treat knowledge capital as a single entity. One can think of the rents on knowledge as being split between managers and firm owners (the split is 3:1 in favor of managers according to Atkeson & Kehoe).

[4] Why would a person or firm allow productivity to fall? I don't know, but according to Harberger (1998) it is a frequent occurrence. It is assumed that if productivity falls, a firm will lay off workers but keep 'AK'. One interpretation is that worker's knowledge is very inhomogeneous and the least knowledgeable are the first to be sacked.

[5] In Staley (2011) and Luttmer (2012) diversity is maintained because each person is an independent generator of productivity shocks.

[6] This is consistent with Klepper & Sleeper (2005), who have studied spinoffs in the laser industry. They say "...nearly all the spinoffs initially produced lasers that their parents had previously produced."

In designing a specific model for spinoffs we can be guided by empirical observation. Of primary relevance is that both entering firms and existing firms are small. The present model assumes that all new firms are individual proprietorships ($L = 1$) and that $L = 1$ is the minimum amount of labour required to run a firm. When the optimal employment size of a firm falls below 1, the firm exits. Finally, the spinoff rate is determined via a simple rule for the formation of spinoff capital: $\mathrm{d}K_{\text{spinoff}}/\mathrm{d}t = \varepsilon\beta F$. That is, some fixed percentage ε of newly produced blueprints gets spun out to new firms. The rate of spinoffs is determined by the rate of growth of the parent firm, which seems consistent with Klepper & Sleeper (2005) and Franco & Filson (2006), who find that the spinoff rate is primarily a function of parental success rather than parental size.

Simulation and analysis shows that with the addition of the spinoff mechanism described above, the growth rate of GDP per capita converges to a steady value and the distribution of firm size rapidly converges to Zipf's law, consistent with observation (Axtell, 2001). Zipf's law for firms states that the proportion of firms having size greater than s is proportional to $1/s$. The reason that the model exhibits this behavior is that the assumptions of firm growth are similar to those used by Gabaix (1999) in his explanation for why *city* size follows Zipf's law: the distribution of relative shocks is independent of size (Gibrat's law) and there is a lower-bound on size.[7]

One characteristic of the present model ensures rapid convergence to Zipf's law and so is worth mentioning. Stochastic productivity shocks translate into shocks to size and also into shocks to the *growth rate* of size due to more rapid accumulation of knowledge. So if a firm gets a large positive productivity shock it grows faster than other firms for a period of time. Given that wages are increasing over time the firm does not stay on the rapid growth path forever (that would require a never-ending series of positive productivity shocks), but this high-growth period helps to generate the extreme size skewness implied by Zipf's law.[8]

[7] The empirical evidence for Gibrat's law is mixed. Stanley et al (1996) report that firm growth volatility declines with size. Audretsch et al (2004) report that when services are included in the sample Gibrat's law attains a higher standing.

[8] This mechanism of fast convergence to Zipf is similar to Luttmer (2011), who describes a model of growth based on the accumulation of organization capital where there are two modes, one high-growth and the other low-growth, and the probability per unit time of transitioning from the former to the latter is constant.

The maintenance of dispersion of productivity is manifest as a stable *traveling wave*, similar to that seen in the model of Staley (2011) and Luttmer (2012).[9] This traveling wave can be pictured as a graph showing the density of firms as a function of log-productivity. The wave maintains its shape (the product of a Bessel function and an exponential function) and moves to the right at a constant speed g, which is the growth rate of the economy. The value of g is related to the shape of the traveling wave, which is not completely constrained by the model. So it seems that any growth rate is possible! However there is one value of g that is stable, and that is the value that is selected.

The analysis required to determine the stable growth rate involves investigating the behavior of the right tail of the productivity distribution where stochastic effects are important.[10] There is a transition region where the deterministic traveling wave, representing many firms, gives way to a probability distribution of individual firms experiencing random innovations. Stability requires that this transition be smooth (no kinks or jumps). When there is a large number of firms this stability requirement means that the traveling wave must have a very long low-density right tail, which is consistent with a high growth rate. Conversely, when there is a small number of firms the shape of the traveling wave must be less spread out, which implies a lower rate of growth.

It turns out that the growth rate is proportional to the log of the number of firms, which is in turn proportional to the log of the number of workers. This so-called scale effect is quite modest, but may have bearing on changes in growth rates over time.[11] For example, a back-of-the-envelope calculation suggests that this scale effect can account for just under half of the increase in the growth rate in the United States since the 1800s.

The growth rate is also proportional to the variance of productivity shocks,

[9] In physics and chemistry, these kinds of traveling waves appear in reaction-diffusion systems. A representative text is Grindrod (1996). See also Whitham (1974).

[10] Fisher, a biophysicist, discusses traveling waves of fitness seen in microbial populations and offers the amusing analogy of the random snuffing of an exploring dog's nose: "The balance between the irregular snuffling and the inertial motion of the body determine the overall speed; yet, as the owner of a large, headstrong dog knows, predicting its speed is very hard!" (Fisher, 2011).

[11] Jones (1999) provides a review of scale effects in various models of endogenous growth. The scale effect described by Staley (2011) is bounded from above.

which suggests a role for specialization. If firms do not specialize they perform many tasks, which diversifies the productivity shocks, but leads to firm-level variances that are low relative to an economy containing specialist firms. Specialization allows the economy to capitalize on positive productivity shocks wherever they arise in the production chain. The possible link between specialization and growth has garnered little attention in modern growth theory (Stigler, 1976) with one exception being the attempt of Romer (1987) to link increasing returns to specialization. The framework of random productivity shocks and selection offers one avenue to explore this relationship. Finally, the growth rate is increasing and convex in the share of knowledge capital.

The present model falls under the umbrella of growth models emphasizing heterogeneous productivity and knowledge diffusion. Recent examples are Luttmer (2007, 2011, 2012a,b, 2014), Lucas (2009), Staley (2011), Alvarez et al (2012), König et al (2012), Lucas & Moll (2014), Perla & Tonetti (2014) and Perla et al (2015). The closest to the present effort is Luttmer (2014), who introduces a tuition schedule into a model of stochastic shocks and imitation. In the present model the compensation for teaching comes instead from ownership of intangible capital. This gives rise to a selection mechanism based on "organic growth", which recalls the work of Nelson & Winter (1982). There are also similarities between the present model and Luttmer (2011), who models the distribution of firm sizes also based on the accumulation of intangible capital. Whereas Luttmer's model has two growth modes, fast and slow, the present model has a continuum that reflects random productivity shocks.

In the present model, entrants are important for maintaining diversity but they are *not* individually more innovative than incumbents. However since there are many more small firms than large firms, radical innovations are more likely to come from small firms. In "Schumpeterian" models of growth (for example Aghion & Howitt, 1992), the image is of small firms discovering new products, or discovering new ways to improve productivity, and displacing large incumbents in the marketplace. In the present model the same thing happens only slower: it takes time for the small victors to accumulate sufficient knowledge capital to affect the market. Chatterjee & Rossi-Hansberg (2012) describe a model in which spinoffs are key to generating productivity improvements. In their model a person comes up with a new idea and decides whether to sell it to their current employer or start a new firm. In contrast, the present model assumes that the person launching a spinoff is merely taking a copy of an already existing idea in order to profit from the knowledge capital associated with it.

The remainder of the paper is organized as follows. Section 2 describes the model, including the traveling-wave solution, the firm-size distribution, and the stochastic behavior at the frontier of knowledge. In Section 3, a calibration is undertaken. The focus is on the spinoff rate, the volatility of firm growth rates, and the share of knowledge capital in output. A predicted autocorrelation effect for growth rates is discussed, and the predicted scale effect is explored in some detail. Section 4 concludes. An appendix includes the derivation of the formula for the growth rate.

2 The Model

2.1 The Basic Setup

The output of firm i is

$$Y_i = (A_i K_i)^\alpha \, L_i^{1-\alpha}, \tag{1}$$

where A_i, K_i and L_i are productivity, knowledge capital, and labor respectively.

Given a market wage w, each firm chooses L_i to maximizes profit $\pi_i = Y_i - wL_i$ given A_i and K_i. The optimum labor for firm i is

$$L_i = A_i K_i \left(\frac{1-\alpha}{w}\right)^{\frac{1}{\alpha}} \tag{2}$$

There is a fixed quantity of labor L that is supplied inelastically. The wage that clears the labor market is obtained by summing Equation (2) over all firms and rearranging:

$$w = (1-\alpha)\left(\frac{\sum_i A_i K_i}{L}\right)^\alpha . \tag{3}$$

The dynamics of this economy is described by the following three equations:

$$\frac{\mathrm{d}K_i}{\mathrm{d}t} = \beta Y_i - \delta K_i, \tag{4}$$

$$\mathrm{d}A_i = A_i\left(\mu \mathrm{d}t + \sigma \mathrm{d}z_i\right), \tag{5}$$

$$\frac{\mathrm{d}L}{\mathrm{d}t} = nL, \tag{6}$$

Equation (4) captures learning by doing and includes a depreciation term δK_i. Equation (5) describes random innovation, where $\mathrm{d}z_i \sim N(0, \mathrm{d}t)$ and draws are uncorrelated across firms. In Equation (6), n is the population growth rate. The model can be simplified by setting δ, μ and n to zero, which is useful in keeping the number of parameters manageable when calibrating to data, but for now we retain these parameters for completeness.

The focus of attention is on finding a balanced growth path, in which case the growth rates of the major variables including w are constant. Following standard procedures we can write down the relationships between the various growth rates, which is useful in determining the next step in the analysis. Summing Equation (4) over i and dividing by K (total capital), the growth rate of K is $g_K = \beta \bar{A}^{\alpha}(L/K)^{1-\alpha} - \delta$, where $\bar{A}$ is the capital-weighted average productivity. The quantity g_K is constant on a balanced growth path, so from observation of the right hand side of this expression we must have $0 = \alpha g_{\bar{A}} + (1-\alpha)(n - g_K)$, where $g_{\bar{A}}$ is the growth rate of $\bar{A}$. From Equation (3) we have $g_w = \alpha(g_{\bar{A}} + g_K - n)$. Combining these two expressions, we obtain the following useful relationship between growth rates

$$g_w = \frac{\alpha}{1-\alpha} g_{\bar{A}} = g_K - n. \tag{7}$$

Our primary interest is in determining g_w so Equation (7) tells us that we must find the growth rate of average productivity. An expression for $g_{\bar{A}}$ can be derived from Equations (1) - (5) assuming there is sufficient granularity of firms that the stochastic shocks in (5) wash out:

$$g_{\bar{A}} = \mu + \beta \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} \frac{\mathrm{Var}_A}{\bar{A}}, \tag{8}$$

where Var_A is the variance of productivity. But from (7), $w^{\frac{1-\alpha}{\alpha}}$ must be linear in $\bar{A}$. Hence on a balanced growth path the relative variance of productivity $\mathrm{Var}_A/\bar{A}^2$ must be constant. The image of a traveling wave comes to mind.

It will prove useful to express the equations of the model in terms of the rent on capital, which is firm-dependent:

$$r_i \equiv \frac{\partial Y_i}{\partial K_i} = \alpha A_i \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}. \tag{9}$$

From the above equation and (7) and (8) it can be seen that if a firm does not innovate, the rent will decay exponentially at a rate that is greater than μ due to the increasing market wage. So if a firm wants to maintain a given level of profitability it must generate a constant stream of positive productivity shocks in excess of drift μ.[12]

We can now rewrite the capital accumulation equation (4) in terms of r_i as

$$\frac{\mathrm{d}K_i}{\mathrm{d}t} = \left(\frac{\beta}{\alpha} r_i - \delta\right) K_i, \tag{10}$$

which implies that the growth rate of total capital is

$$g_K = \frac{\beta}{\alpha}\bar{r} - \delta. \tag{11}$$

The capital-weighted average rent $\bar{r}$ can be expressed as

$$\bar{r} = \alpha \bar{A} \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}. \tag{12}$$

From Equations (7) and (12), we can see that $\bar{r}$ is constant on the balanced growth path. So $\{r_i\}$ must be the set of *de-trended productivities*. From Equation (5) the dynamic equation for r_i is

$$\frac{\mathrm{d}r_i}{r_i} = \left(\mu - \frac{1-\alpha}{\alpha} g_w\right) \mathrm{d}t + \sigma \mathrm{d}z_i, \tag{13}$$

where $g_w = \frac{\beta}{\alpha}\bar{r} - \delta - n$, which is obtained from substituting Equation (11) into (7). Finally, defining $\tilde{K}_i$ to be the proportion of total knowledge capital allocated to firm i, ie $\tilde{K}_i \equiv K_i/K$, we have[13]

$$\frac{\mathrm{d}\tilde{K}_i}{\mathrm{d}t} = \frac{\beta}{\alpha}\left(r_i - \bar{r}\right)\tilde{K}_i, \quad \text{where} \quad \bar{r} = \sum_i \tilde{K}_i r_i, \tag{14}$$

[12] The Red Queen said to Alice "Now, *here*, you see, it takes all the running you can do, to keep in the same place." (Carroll, 1871).

[13] Equation (14) is the replicator equation used in evolutionary game theory (Taylor & Jonker, 1978, Schuster & Sigmund, 1983).

As we shall see, Equations (13) and (14) describe the dynamics of a traveling wave.

2.2 Spinoffs

As mentioned in the Introduction, sustained growth is only possible in the current model if there is a diversity of productivity shocks. If a single firm takes over, capturing all the knowledge rents available in the economy, growth will stop. In the real world the distribution of firm size is very skewed, with a huge number of small firms, a lot of medium size firms, and a few large firms. Remarkably, the observed size distribution almost exactly matches *Zipf's law*: the proportion of firms having an employee count greater than s is approximately a/s where a is a constant (Axtell, 2001). The corresponding density function is a/s^2. The goal of this section is to augment the model described in the previous section with a model of firm entry based on *spinoffs* such that Zipf's law is satisfied.

The following model is motivated partially by empirical observation, and partially by the theoretical framework of Gabaix (1999), who describes how a Zipf distribution arises from random growth and a lower bound on size. No attempt is made to link entry and exit to the rational behavior of firms.

Entry. According to Klepper & Sleeper (2005) and Franco & Filson (2006), the rate of spinoffs of new firms is a function of parental success, which presumably correlates with growth. Accordingly, let us assume that some portion ε of newly-formed knowledge capital (gross of depreciation) is spun out to new firms. Further assume that new firms are individual proprietorships ($L = 1$). This assumption is partially consistent with Klaesson & Karlsson's study of Swedish firms (Klaesson & Karlsson, 2014, figure 5.3), which shows that the size distribution of entrants is very skewed towards single proprietorships. Finally, let us assume that the amount of capital in a startup is inferred from Equation (2) using $L_i = 1$. The implicit assumption is that the number of blueprints taken by the entrepreneur from the parent firm is equal to the average number of blueprints per person in that firm.

Exit. A realistic model of firm dynamics should include an exit mechanism. Let us assume that when the optimum amount of labor falls below 1 the firm exits and all knowledge capital is lost. The owner then enters into the labor market. This simple-minded model of exit implies a lot of churn in the small-firm segment of the economy, consistent with observation. This mechanism also enforces a lower bound, as in Gabaix (1999).

It is useful to write down an expression for the rate of entry, since that is an observable quantity, hence useful for calibration purposes. Starting from Equation (10) and converting to labor flows using (2) and (9), the number of single-employee spinoff firms produced by firm i per unit time is

$$\frac{\mathrm{d}N_{i\,\mathrm{spin}}}{\mathrm{d}t} = \varepsilon\,\frac{1-\alpha}{\alpha}\frac{\beta}{\alpha}\frac{r_i^2 K_i}{w}.$$

To get the total spinoff rate for the economy, we must sum the above expression over i, which leaves us with an expression containing the second moment of $\{r_i\}$, which we will label M_2. Recall that on the balanced growth path $\bar{r} = \sum_i r_i \tilde{K}_i$ is stationary. Using (13) and (14) we can write down an expression for the change in $\bar{r}$ over time that contains the quantity M_2. Setting the change in $\bar{r}$ equal to zero we obtain an expression for M_2 that can be combined with other equations in Section 2.1 to obtain

$$\frac{1}{L}\frac{\mathrm{d}N_{\mathrm{spin}}}{\mathrm{d}t} = \varepsilon\left\{\frac{g_w}{\alpha} + \delta + n - \mu\right\}. \tag{15}$$

The rate of spinoffs can be expressed as a percentage of the number of firms N_F by multiplying both sides of (15) by L/N_F.

2.3 Simulation of the Model

A simulation exercise shows that with the addition of the above mechanisms of entry and exit something close to a Zipf distribution for employment size is observed for the entire population of firms (see Figure 1). Convergence to Zipf is rapid, which means that the largest firms spend time on very fast growth tracks.[14] In Figure 1 the largest firm (extreme bottom-right point) is 39 years old and has 1.49 million employees. The average growth rate since birth of the largest twenty-five firms is 17.6% per annum and their average age is 56 years. These large firms manage to stay on high-growth tracks for a long time before getting bumped to a different (usually slower) growth track.

[14] This is consistent with observation. See Luttmer (2011) for a discussion of what is takes to produce a Zipf distribution.

The simulation exercise also shows that productivity forms a traveling wave, so the distribution over r is stationary. The shape can be seen in Figure 2. Note that the Zipf distribution for firm size is observed not just for the whole population of firms but for each small range of r. Imagine if you will a group of firms satisfying Zipf's law residing within the thin column shown in Figure 2. The firm-size distribution is Zipf within each such column, suggesting that the density of firms can be factored as $\rho(r)/s^2$ where $\rho(r)$ is the density over rent and s is employment size.

The above factorization implies Gibrat's law, which is often taken as a starting point in devising mechanisms to explain the observed Zipf distribution. Gibrat's law states that the distribution of growth rates is independent of size. To see the connection between factorization and Gibrat's law, recall from Equation (14) that any firm in a given column will grow (both labor and capital) at the steady rate $\beta/\alpha(r - \bar{r})$. We can think of a column as a *growth track*. So the factorization implies that we have the same firm-size distribution on each growth track. This in turn implies that the distribution of growth rates is independent of size.

Finally, the simulation shows that the market wage w grows geometrically at a constant rate. The spinoff mechanism described in the previous section prevents any single firm from taking over the economy, hence the fuel of selection is never exhausted.

2.4 Analytic Solution

Let $P(r, \tilde{K}, t)$ be the density of firms as a function of rent r, capital $\tilde{K}$ and time t. That is, the proportion of firms earning rent between r and $r + \mathrm{d}r$ and having capital between $\tilde{K}$ and $\tilde{K} + \mathrm{d}\tilde{K}$ at time t is $P(r, \tilde{K}, t)\,\mathrm{d}r\,\mathrm{d}\tilde{K}$. Ignoring the spinoff mechanism for now, Equations (13) and (14) imply the following partial differential equation for P:[15]

$$\begin{aligned}\frac{\partial P}{\partial t} &= -\frac{\partial}{\partial \tilde{k}}\left[\tilde{k}\frac{\beta}{\alpha}\left(r-\bar{r}\right)P\right] \\ &\quad -\left[\mu - \frac{1-\alpha}{\alpha}\left(\frac{\beta}{\alpha}\bar{r} - \delta - n\right)\right]\frac{\partial}{\partial r}\left(rP\right) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial r^2}\left(r^2 P\right),\end{aligned} \tag{16}$$

where

$$\bar{r}(t) = \int_0^\infty \int_0^\infty P(r, \tilde{K}, t) r \,\mathrm{d}r\,\mathrm{d}\tilde{K}.$$

If we focus our attention on the part of the size distribution where $s > 1$, we can incorporate the spinoff mechanism by replacing β with $\beta(1-\varepsilon)$ in Equation (4). Then we can simply absorb the factor $(1-\varepsilon)$ into β, and so no real change to the model description in Section 2.1 is required, and Equation (16) remains valid.

To begin our analysis, we first show that a Pareto distribution for size is consistent with Equation (16). It is readily verified that Equation (16) is satisfied by $P(r, \tilde{k}, t) = C\rho(r,t)/\tilde{k}^\theta$, where C and θ are constants and $\rho(r,t)$ satisfies[16]

$$\frac{\partial \rho}{\partial t} = D_I + D_{II}, \tag{17}$$

where

$$\begin{aligned} D_I &= (\theta - 1)\frac{\beta}{\alpha}\left[r - \bar{r}\right]\rho, \\ D_{II} &= -\left[\mu - \frac{1-\alpha}{\alpha}g_w\right]\frac{\partial}{\partial r}\left(r\rho\right) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial r^2}\left(r^2\rho\right).\end{aligned} \tag{18}$$

The differential equation has been divided into two pieces D_I and D_{II} for the purposes of discussion below. We will be most interested in the stationary solution $D_I + D_{II} = 0$ because that solution represents a de-trended traveling wave.

[15] The multidimensional version of the Fokker-Planck equation can be used to derive this differential equation. See the Wikipedia page on "Fokker-Planck equation", and scroll down to the section entitled "Many dimensions". If one defines the vector $\boldsymbol{X}$ to be $(\tilde{k}, r)^\top$, the two-by-two diffusion matrix D contains zeros everywhere except for the bottom-right corner, which contains $r^2\sigma^2$.

[16] If θ is dependent on r the solution does not factorize, which violates Gibrat's law, so we assume that θ is a constant.

Before going further into the derivation of the solution, it will be useful to gain some intuition around the origin of the Pareto distribution for size.[17] Here we will refer to employment size rather than capital, but recall that employment is linear in capital for given r, so if employment size follows a Pareto distribution with exponent θ, so must capital.

All firms start with size $s = 1$. In steady state the density of firms at $s = 1$ must be constant, and this is the result of offsetting flows of entry, exit and growth (firms leaving the starting blocks after they are born). Let us label the last quantity η, so at any given time the number of firms starting to grow from size $s = 1$ is equal to η. Size grows as

$$s = e^{gt} \tag{19}$$

where $g = \beta/\alpha\,(r - \bar{r})$ as in the above discussions. Now consider the small group of firms having size between s and $s + \mathrm{d}s$. Differentiating (19) we have $\mathrm{d}t = \mathrm{d}s/(sg)$. So all the firms with size between s and $s + \mathrm{d}s$ must have originated from a pool of new firms of count $\eta\,\mathrm{d}s/(sg)$. Let us also assume that firms get bumped off the growth track at some hazard rate h due to stochastic productivity shocks. Call the density of firms $f(s)$. Then the number of firms between s and $s + \mathrm{d}s$ must be

$$f(s)\mathrm{d}s = \eta \frac{\mathrm{d}s}{sg} e^{-ht},$$

where t can be obtained from (19). Substituting for t we have

$$f(s) = \frac{\eta}{g} s^{-(1+h/g)}. \tag{20}$$

Referring back to the assumed form of the solution $P(r, \tilde{k}, t) = C\rho(r, t)/\tilde{k}^{\theta}$, we must have $h/g = \theta - 1$. From (18) it is apparent that $D_I = (\theta - 1)g$ and $D_{II} = -h$, the latter because D_{II} captures the changes in density due to stochastic deviations in r (changes in the growth path). The stationary solution $D_I + D_{II} = 0$ is indeed consistent with $h/g = \theta - 1$.

The above reasoning tells us that the size distribution is Pareto, but it doesn't tell us anything about the value of the exponent θ. To get θ we must look at the problem in a different way. Consider the case where the size of the labor

[17] This explanation is inspired by similar arguments in Simon (1955), Krugman (1996), Gabaix (1999) and Luttmer (2011).

force is fixed (the case of a growing labor force will be discussed below). Then the distribution of the growth rate must have a mean of zero. According to Gabaix (1999), the assumption of a zero mean, coupled with a lower bound on size, leads to Zipf's law ($\theta = 2$).[18]

It is often pointed out that a Zipf distribution implies an infinite mean. It is true that if we have a *continuum* of firms the average firm size is

$$\int_1^\infty \frac{1}{s^2}\, s \, \mathrm{d}s = \infty.$$

However in a finite economy the largest firm must have a finite size, so the average must also be finite. Appendix A derives some expressions relating the number of firms N_F, and the size of the largest firm $s_{\max}$, to the size of the labor force L. The conclusion is that the log of the number of firms is linear in the log of total labor and $s_{\max} = N_F$.

In the above, the assumption was that the total size of the labor force was fixed. In reality the population and labor force L grow slowly. Since the log of the number of firms is proportional to the log of L, then for small ranges of L the growth rate of firms will be very close to that of labor. In simulation exercises it is indeed observed that N_F grows almost exactly in parallel with L. So we can assume for the purposes of the size distribution that population growth is like having a sequence of static labor forces. Therefore Gabaix's proof of Zipf's law still holds for a growing population.

The remaining task is to determine the distribution of rent $\rho(r)$. Since we are interested in the stationary solution we can set the left hand side of (17) to zero, set $\theta = 2$, and rewrite the partial derivatives as ordinary derivatives (ρ is now a function of r only):

$$0 = r^2 \frac{\mathrm{d}^2\rho}{\mathrm{d}r^2} + ar\frac{\mathrm{d}\rho}{\mathrm{d}r} + (br + c)\rho, \tag{21}$$

[18] Gabaix's basic proof of the Zipf law assumes random shocks (see page 744), but the proof also works if there is a cross-sectional dispersion of growth rates, as in our model. On page 744, Equation (4), the assumption that the distribution of shocks is independent of size is necessary to show that Zipf's law is satisfied for zero-mean shocks. This is consistent with our approach of factoring the solution such that the size distribution is independent of the growth track.

where

$$a = 4 + \frac{2}{\sigma^2}\left[\frac{1-\alpha}{\alpha} g_w - \mu\right], \tag{22}$$

$$b = \frac{2}{\sigma^2}\frac{\beta}{\alpha}, \tag{23}$$

$$c = 2 + \frac{2}{\sigma^2}\left[\frac{1-\alpha}{\alpha} g_w - \mu\right] - \frac{2}{\sigma^2}\left(g_w + \delta + n\right). \tag{24}$$

The solution is[19]

$$\rho(r) = B\, r^{\frac{1-a}{2}} J_\gamma\left(2\sqrt{br}\right), \tag{25}$$

where J_γ is an order-γ Bessel function of the first kind, B is a normalizing constant, and

$$\gamma = \sqrt{(1-a)^2 - 4c}. \tag{26}$$

The shape of ρ can be seen in Figure 2.

Finally, it is worth mentioning that if one abstracts from knowledge capital, the distribution of total factor productivity is Pareto. This can be seen from Equations (1) and (2). If labor is Zipf distributed, so must be $A_i K_i$. If we absorb $(A_i K_i)^\alpha$ into B_i and write $Y_i = B_i L_i^{1-\alpha}$, $\{B_i\}$ must follow a Pareto distribution with exponent $1/\alpha$.

2.5 Growth Rate

The focus of this section is on determining the unique stable growth rate g_w consistent with stochastic behavior at the the frontier of knowledge.[20] It will be necessary to make some approximations in order to get a tractable

[19] See Polyanin & Zaitsev (2003), p. 228, Equation 132. There is a second solution in Polyanin & Zaitsev that contains a Bessel function of the second kind, Y_γ, but this solution blows up at the origin so has been excluded (see Abramowitz & Stegun, 1964, p. 438, Figure 10.2).

[20] The mathematical techniques used in this section are described in Tsimring et al (1996), Brunet & Derrida (1997) and Staley (2011). Hallatschek (2011) describes a promising new technique that has not been applied to the present growth model. These methods are inspired by Fisher (1937), who analyzed the stochastic properties of fitness distributions to determine the stable speed of advance of a gene over a geographical area.

formula for g_w. The following formula turns out to be quite dependable for $\delta = n = \mu = 0$, and $0 < \alpha < 0.25$:

$$g_w = c \frac{\alpha}{1-\alpha} \sigma^2 \ln L, \tag{27}$$

where $c = 0.455$, which is obtained from a linear regression of simulation results (see Figure 3). The scale effect is evident from the $\ln L$ factor. This is highlighted in Figure 4 (recall from Equation (A.3) that $\ln L$ is proportional to $\ln N_F$). The remainder of this section lays out the analytical justification for Equation (27). Details are contained in Appendix B.

The growth rate formula assumes that $\delta = n = \mu = 0$. By setting $\mu = 0$ we are describing a model of random innovation and selection, with no exogenous growth. The parameter δ is set to zero because it is unclear how to calibrate it. The population growth rate n is set to zero because a simulation exercise shows that the growth rate is quite insensitive to that parameter and the growth rate formula is vastly simplified if $n = 0$.

We will work with the de-trended productivity distribution $\rho(r)$, which is the density of firms with respect to knowledge rent (Figure 2). We will ignore the firm-size distribution because it has no bearing on the growth rate. It is assumed that the distribution can be divided into two portions: a deterministic traveling-wave portion called $\rho_I(r)$ that describes the body of the distribution, and a portion called $\rho_{II}(r)$, which describes stochastic behavior at the sparsely populated frontier (far right) of the distribution. It will useful to work with the variable $z = 2\sqrt{br}$ in order to simplify the algebra. Using the density-transformation formula $\rho(z) = \rho(r)\,\mathrm{d}r/\mathrm{d}z$ we can rewrite Equation (25) as

$$\rho_I(z) = \frac{B}{2b} z^{2-a} J_\gamma(z), \tag{28}$$

where B is a normalization constant and a, b and γ are given by Equations (22), (23) and (26) respectively.

In the stochastic far-right tail, our focus is on the region of the probability distribution where the portion of firms is close to $1/N_F$ where N_F is the number of firms. The density in this region represents potential fluctuations of r beyond the frontier, where there are no firms yet, hence no firm growth. The differential equation governing behavior in this region can be obtained by dropping the first term in (17) (here we are reverting back to r-space momentarily):

$$\frac{\partial \rho_{II}}{\partial t} = -\left(\mu - \frac{1-\alpha}{\alpha} g_w\right) \frac{\partial}{\partial r}\left(r\rho_{II}\right) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial r^2}\left(r^2 \rho_{II}\right). \tag{29}$$

The above formula describes simple geometric Brownian motion with a lower bound (the interface with region I), similar to the dynamics discussed in the previous section. We are interested in the stationary solution so we can guess that it takes the form r^b. Substituting this form into (29) and converting to the variable z we have

$$\rho_{II}(z) = Az^{5-2a}, \tag{30}$$

where A is a constant.

We assume that region I meets region II at some point z_c. We have 4 unknowns: g_w, z_c, and the constants A and B. We also have the following four equations:

$$\rho_I(z_c) = \rho_{II}(z_c), \tag{31}$$

$$\rho_I'(z_c) = \rho_{II}'(z_c), \tag{32}$$

$$\int_{z_c}^{\infty} \rho_{II}(z)\mathrm{d}z = \frac{1}{N_F}, \tag{33}$$

$$\int_0^{z_c} \rho_I(z)\mathrm{d}z = 1 - \frac{1}{N_F}. \tag{34}$$

The first equation says that there should be no jump in density at the boundary z_c. The second says that there should be no kink. These two conditions are required to ensure that no terms in (29) blow up at the boundary, which would lead to instability. The third condition is just the definition of z_c. The last condition can be used to determine the normalization constant B (which may be a function of other parameters).

It should be noted that the form of the density in (28) is actually quite strange due to the presence of the Bessel function. The body of the Bessel function is bell-shaped, starting with a value of zero at the origin and reaching a value of zero again at some point z_0 (the first *zero* of the Bessel function). But then it oscillates much like a trigonometric function, taking on both positive and negative values as one moves to the right. Density functions are not supposed to have negative values. But as it turns out the connection point z_c lies to the left of z_0, as shown in the Appendix.

From Equation (31) we get

$$\rho_{II}(z) = \rho_I(z_c)\left(\frac{z}{z_c}\right)^{5-2a}. \tag{35}$$

Substituting this expression into (33) we obtain

$$\rho_I(z_c)\frac{z_c}{2a-6} = \frac{1}{N_F}. \tag{36}$$

Substituting from (28) and taking logs we then obtain

$$(a-3)\ln z_c + \ln(2a-6) - \ln\left[\frac{B}{2b}J_\gamma(z_c)\right] = \ln N_F. \tag{37}$$

In order to proceed further we need to resort to approximation, and the details are contained in Appendix B. But even without going through all that algebra we can guess what the growth rate formula should look like. Noting that the logarithmic function has the effect of killing variability we expect the first term of (37), which is linear in $(a-3)$, to dominate the behavior of the left hand side. So we can guess that $a \approx a_1 \ln N_F + a_2$ where a_1 and a_2 are constants. Substituting for a using Equation (22), assuming $\mu = 0$, and taking into account (A.3), we obtain Equation (27) apart from an additional constant term that turns out to be very small.

3 Testing the Model

In this section the growth rate formula (Equation 27) is tested on U.S. data for the twentieth century, and the scale effect is tested on U.S. data going back to 1820. In addition, a predicted autocorrelation effect for firm-level growth rates is discussed.

There are four numbers in the the growth formula: the growth rate g_w, the share of knowledge capital (or span of control parameter) α, the volatility of productivity shocks σ and the scale factor $\ln L$. The quantities g_w and $\ln L$ are easily determined, there are estimates of α available in the literature, and the parameter σ can be estimated using data on firm growth rates. So we have enough material to test the growth formula.

Let us equate the growth rate g_w with the geometric average growth rate of GDP per capita in the United States from 1900-2000. According to Maddison (2009) this value is $g_w = 0.194$ per annum. Similarly, let us set the scale factor $\ln L$ to be the average log labor over the twentieth century. This can be estimated using Maddison's population data assuming a labor participation rate of one half (roughly today's value). This leads to $\ln L = 18.169$, which corresponds to a geometric average labor force of 77.8 million.

The remaining two parameters are α and σ. Figure 5 shows how σ varies with α given the above values of g_w and $\ln L$.

Let us next consider the span of control parameter α. Atkeson & Kehoe (2005) suggest a value of 0.15 but they are working in an economy in which there is physical capital as well as intangible capital, and the share of physical capital is 0.199. We need to adjust Atkeson & Kehoe's span-of-control parameter so that it can be applied to our economy, which has no physical capital. To do so we scale their value by the factor $1/(1 - 0.199)$, which implies $\alpha = 0.187$. This adjustment can be derived assuming a Cobb-Douglas production function and perfectly elastic capital, and working through the algebra to eliminate physical capital from the production function.

As an aside we can now use the value of α to calibrate the spinoff parameter ε. This parameter does not enter into the formula for the growth rate so is not really necessary for this section, but was required in order to run the simulations, such as used to determine the constant in Equation (27). From Equation (15) we have

$$\text{Spinoff Rate} = \varepsilon \frac{g_w}{\alpha} \frac{L}{N_F}.$$

According to Luttmer (2007, page 1132) the entry rate in the U.S. is 0.116 per anum. Let us assume that all entries are spinoffs. The labor force L in 1990 is 143,280,000 and the number of firms N_F is 5,697,759.[21] Using $g_w = 0.194$ and $\alpha = 0.187$ in Equation (15) we get $\varepsilon = 0.045$. If we use instead the value of $L/N_F = 15$ as suggested by Table 1, we get $\varepsilon = 0.075$. The value $\varepsilon = 0.05$ was used in the simulation studies.

According to the growth formula (and see Figure 5) $\alpha = 0.187$ implies $\sigma = 0.1$. To test this prediction, we can look to the data on firm growth rates.

[21] See http://www.dlt.ri.gov/lmi/pdf/usadj.pdf for labor force statistics and Luttmer (2007) footnote on page 1 for the number of firms.

From Equations (2), and (4) - (6), the relative change in labor for firm i is

$$\frac{\mathrm{d}L_i}{L_i} = \left\{ \frac{\beta}{\alpha}(r_i - \bar{r}) - \frac{1-\alpha}{\alpha} g_w \right\} \mathrm{d}t + \sigma \mathrm{d}z_i, \tag{38}$$

so the parameter σ can be directly compared to the volatility of firm growth rates over time.[22] Several recent estimates of firm-growth volatilities are given in Davis et al (2007). They report values for public firms between 0.09 and 0.13 covering the years 1955 to 2000, which are close to our prediction of $\sigma = 0.1$. However when they include private firms the values range between 0.1 and 0.2. These volatilities are computed for firms with at least ten years of consecutive observations. When short-lived firms are included in the sample, the volatilities become even larger. The average is about 0.45 for the whole dataset and about 0.275 if exiting firms are excluded. In conclusion, the volatility prediction of the present model is only consistent with data on long-lived public firms.

It is possible that there is a size effect at work here. Public firms tend to be larger than private firms, and short-lived firms are overly represented by very small firms. Stanley et al (1996) report that small firms are much more volatile than large firms. So the present model may only be valid for large firms, meaning either that the model fails for small firms or that the excess volatility of small-firm growth rates is driven by factors other than productivity changes.

Equation (38) also implies positive autocorrelation for firm growth rates, which is highest when r_i is large. The challenge in testing this prediction is that r_i is unobservable. One might proxy large r_i with large size, since large firms must have been on a high growth track for some time. Pursuing this line of thought, a prediction of the model is that large firms should exhibit positive autocorrelation in growth rates. This pattern is indeed observed by Coad (2007) and Coad & Hölzl (2009). However those authors also report negative autocorrelation for small firms, which is not seen in the present model.

There are several ways that one might test the predicted scale effect. Cross-country studies, cross-industry studies, and historical studies come to mind.

[22] One must be careful not to use *cross-sectional dispersion* measures as shown in Davis et al (2007). These are always higher than volatility measures and are probably reflecting inhomogeneity of firm-level volatility rates (for example by size as reported by Stanley et al, 1996) and possibly autocorrelation of firm growth rates.

The last choice is arguably the cleanest because one does not have to worry about the many factors that can affect growth rates across different jurisdictions. The present approach is to use data on the U.S. economy going back to 1820 and assume that the only thing that has changed over that long period of time has been the size of the labor force. So the model is very simple:

$$g_w = C \ln L,$$

where L is the labor force and C is calibrated by dividing the growth rate during the twentieth century by the average log labor force during the twentieth century. Figure 6 shows the resulting growth rate by decade vs the actual growth rate.[23] The predicted growth rate during the the nineteenth century is 17.3% versus an actual growth rate of 14.7%. So the scale effect accounts for 43% of the increase in growth rate over the last two centuries. The calculation was repeated for the nonagricultural sectors only (assuming that economic growth is driven by non-agricultural industries only), which resulted in a slightly lower prediction for the 19th century: $g_w = 16.8\%$.[24] In that case the scale effect accounts for half of the increase in growth rate over the last two centuries.

Finally, we can use the scale effect to extrapolate from the current U.S. economy. If population continues to grow at the current rate of about 1% per year, the predicted growth rate of GDP per capita in the year 2115 will be 2.1% per annum. The effect of globalization can also be inferred. If the whole world were to become a single economy similar to the U.S. with a population of 7 billion (22 times the population of the U.S. today), the growth rate of per-capita GDP would be about 2.4% per annum. The scale effect is very modest.

[23] This calculation takes into account that the labor participation rate was lower in the 19th century than in the 20th century. For example in 1830 it was 32% compared to 46% today (Broadberry & Erwin, 2006). Population numbers are from Maddison (2009).

[24] The non-agricultural sector in 1830 was only 35% compared to 98% today (Carter, 2005).

4 Conclusion

A model of firm dynamics has been presented based on the accumulation of *knowledge capital*, which is a source of profits and a driver of firm growth. It is assumed that knowledge is disseminated from the "top" and is subject to random shocks that affect the productivity of the entire firm. This dynamic process leads to aggregate growth based on random innovation and selection.

To maintain granularity in this economy a model of spinoffs has been included, which leads to a Zipf distribution of firm size consistent with observation. A salient feature of the model is that convergence to Zipf is rapid, reflecting the existence of rapid "growth tracks".

The distribution of productivity forms a traveling wave, representing a growing economy. The speed of this wave is determined by a combination of stochastic behavior at the frontier of knowledge and inertial drag in the body of the distribution caused by a finite speed of selection. Using tools developed in the biophysics community it has been found that the speed of the traveling wave is proportional to the log of the size of the economy. This behavior implies a modest scale effect that may have bearing on the slow acceleration of growth seen in historical data.

The model implies that the volatility of firm growth rates is around 10% per annum, consistent with data for large firms but not for small firms. Likewise, the predicted positive autocorrelation of growth rates is consistent with data for large firms but not for small firms. These two observations imply that the model is breaking down for small firms. One interpretation is that there are factors other than productivity affecting the size of small firms that have not been captured in the present model.

Finally, it was mentioned in passing that the model implies a connection between specialization and growth. This connection should be present in any model based on random innovation and selection, and is worth pursuing due to its potential to link economic growth with markets and trade.

A Zipf's Law in a Finite Economy

Define N_F to be the number of firms, L to be the size of the labor force, and $s_{\max}$ to be the size of the largest firm. Let $G(s) = 1/s$ be the proportion of firms with size greater than or equal to s. G is only defined between 1 and $s_{\max}$. We have the following

$$G(s_{\max}) = 1/N_F,$$

$$\int_1^{s_{\max}} s f(s)\,\mathrm{d}s = L, \quad \text{where } f(s) \equiv -\frac{\mathrm{d}G}{\mathrm{d}s}.$$

These equations imply

$$s_{\max} = N_F, \tag{A.1}$$

$$N_F \ln(N_F) = L, \tag{A.2}$$

Equation (A.1) says that there are as many employees in the largest firm as there are firms. Equation (A.2) can be used to infer the number of firms given L using an iterative procedure. Table 1 lists the values of N_F for a range of values of L. The third column lists the corresponding average size. Apparently the function $N_F(L)$ is slightly convex, and the elasticity is nearly constant. We can use linear regression to derive the following handy formula ($R^2 = 0.9987$ based on the data in Table 1):

$$\ln N_F \approx 0.8917 \, \ln L - 0.7594. \tag{A.3}$$

B An Approximation for the Growth Rate

The purpose of this appendix is to use some approximations to justify Equation (27). The stating point for this derivation is Equation (37). We need to use Equations (32) and (34) to eliminate the rest of the unknowns.

We would expect the connection point z_c to be very close the first zero of the Bessel function z_0 because we are dealing with a region where ρ_I is small. Our first approximation then is to assume that $\rho_I(z)$ is linear between z_c and z_0:

$$\rho_I(z_c) \approx \rho_I'(z_0)\,(z_c - z_0). \tag{B.1}$$

L	N_F	Average Firm Size (L/N_F)
10	6	1.75
100	30	3.39
1,000	190	5.25
10,000	1,383	7.23
100,000	10,771	9.28
1,000,000	87,848	11.38
10,000,000	739,955	13.51
100,000,000	6,382,029	15.67
1,000,000,000	56,048,388	17.84
10,000,000,000	499,283,712	20.03

Table 1: *Number of employees L and number of firms N_F according to Equation (A.2). The Excel Solver was used to determine N_F.*

From (B.1) and (32) we then get

$$z_c = z_0 \frac{2a-5}{2a-4}, \tag{B.2}$$

which confirms that z_c lies to the left of z_0.

We now use the following asymptotic form for $J_\gamma(z)$ for large z near the zero (Abramowitz & Stegun, 1964, page 364)

$$J_\gamma(z) \approx \sqrt{\frac{2}{\pi z}} \cos\left(z - \frac{1}{2}\gamma\pi - \frac{1}{4}\pi\right). \tag{B.3}$$

Using (37), (B.2), along with (B.1) and (B.3) we can eliminate z_c to obtain

$$\frac{B}{2b}\sqrt{\frac{2}{\pi}} z_0^{\frac{7}{2}-a} \frac{2a-5}{(2a-4)^2(2a-6)} = \frac{1}{N_F}. \tag{B.4}$$

To get $B/2b$ we must use Equation (34), which cannot be solved analytically. When $\delta + n$ is small (e.g. on the order of a few percent) the shape of ρ_I is very skewed, with a very sharp peak near $z = 0$. Hence we can concentrate on the form of the integrand near $z = 0$. For small z the Bessel function can be approximated as (Abramowitz & Stegun, 1964, page 360)

$$J_\gamma(z) \approx \frac{1}{\Gamma(\gamma+1)} \left(\frac{z}{2}\right)^\gamma \tag{B.5}$$

The integral on the left hand side of (34) is then

$$\begin{aligned} I &= \frac{B}{2b}\frac{1}{\Gamma(\gamma+1)2^{\gamma}}\int_0^{z_0} z^{2-a+\gamma}\,\mathrm{d}z \\ &= \frac{B}{2b}\frac{z_0^{3-a+\gamma}}{\Gamma(\gamma+1)2^{\gamma}(3-a+\gamma)}. \end{aligned} \tag{B.6}$$

The above integral is well defined at $z = 0$, even though the integrand blows up there, because according to the definitions of a and γ we must have $3 - a + \gamma > 0$.

Setting $I = 1$ (ignoring the difference between N_F and $N_F - 1$) and substituting $B/2b$ using (B.6) into (B.4) and then taking logs, we have

$$\begin{aligned} &\gamma\left\{\ln 2 - \ln z_0 + \ln(\gamma+1) - 1\right\} \\ &+ \ln 2 - \frac{1}{2}\ln z_0 + \frac{1}{2}\ln(\gamma+1) - 1 \\ &+ \ln(3-a+\gamma) + \ln(2a-5) - 2\ln(2a-4) - \ln(2a-6) \\ &= -\ln(N_F) \end{aligned} \tag{B.7}$$

For typical values encountered in this project the quantity in the curly brackets in the first row of (B.7) is negative. So again, remembering that the log function kills variability, we can guess that γ is approximately linear in $\ln N_F$. From (26) we have (assume μ is zero):

$$\gamma = \left\{\left(1 + \frac{2}{\sigma^2}\frac{1-\alpha}{\alpha}g_w\right)^2 + \frac{8}{\sigma^2}(g_w + \delta + n)\right\}^{\frac{1}{2}}$$

For values of α less than 0.5, and for small δ and n (a few percent) the first term under the square root dominates over the second, and we have

$$\frac{2}{\sigma^2}\frac{1-\alpha}{\alpha}g_w \approx a_1 \ln N_F + a_2,$$

where a_1 and a_2 are constants (a_1 is positive). Given the relation between $\ln N_F$ and $\ln L$ in Equation (A.3), we finally have

$$g_w \approx b_1 \frac{\alpha}{1-\alpha}\sigma^2 \ln L + b_2,$$

where b_1 and b_2 are constants. The regression results shown in Figure 3 indicate that b_2 must be close to zero. Given all the approximations that have been made, it is remarkable that the formula for the growth rate works as well as it does.

References

[1] Abramowitz, M., Stegun, I. (1964) *Handbook of Mathematical Functions.* National Bureau of Standards, Washington.

[2] Aghion, P., Howitt, P. (1992) A Model of Growth Through Creative Destruction. *Econometrica* **60**: 323-351.

[3] Arrow, K. (1962) The Economic Implications of Learning by Doing. *The Review of Economic Studies* **29**: 155-173.

[4] Alvarez, F., Buera, F., Lucas, R. (2008) Models of Ideas Flows. NBER Working Paper no. 14135.

[5] Alvarez, F., Buera, F., Lucas, R. (2012) Idea Flows, Economic Growth, and Trade. NBER Working Paper No. 19667.

[6] Atkeson, A, Kehoe, P. (2005) Modeling and Measuring Organization Capital. *Journal of Political Economy* **113**: 1026-1053.

[7] Audretsch, D.B., Klomp, L., Santarelli, E., Thurik, A.R. (2004) Gibrat's Law: Are the Services Different? *Review of Industrial Organization* **24**: 301-324.

[8] Axtell, R. (2001) Zipf Distribution of U.S. Firm Sizes. *Science* **293**: 1818-1820.

[9] Broadberry, S., Irwin, D. (2006) Labor Productivity in the United States and the United Kingdom During the Nineteenth Century. *Explorations in Economic History* **43**: 257-279.

[10] Brunet, E., Derrida, B. (1997) Shift in the velocity of a front due to a cutoff. *Physical Review E* **56**: 2597-2604.

[11] Carroll L. (1871) *Through the Looking-Glass and What Alice Found There.*

[12] Carter S. B. (2005) *Labor Force, for Historical Statistics of the United States, Millennial Edition.* Cambridge University Press, New York.

[13] Chatterjee, S., Rossi-Hansberg, E. (2012) Spinoffs and the Market for Ideas. *International Economic Review* **53**: 53-93.

[14] Coad, A. (2007) A Closer Look at Serial Growth Rate Correlation. *Review of Industrial Organization* **31**: 69-82.

[15] Coad, A., Hölzl, W. (2009) On the Autocorrelation of Growth Rates. *Journal of Industry, Competition and Trade* **9**: 139-166.

[16] Coad, A., Hölzl, W. (2010) Firm Growth: Empirical Analyis. In *Handbook on the Economics and Theory of the Firm*, edited by Dietrich, M., Edward Elgar, Cheltenham.

[17] Cox, D., Miller, H. (1995) *The theory of stochastic processes.* Chapman and Hall, London.

[18] Davis, S., Haltiwanger, J., Jarmin, R., Miranda, J. (2007) Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms. *NBER Macroeconomics Annual 2006* **21**: 107 - 180.

[19] Fisher, D.S. (2011) Leading the Dog of Selection by its Mutational Nose. *Proceedings of the National Academy of Sciences* **108**: 2633-2634.

[20] Fisher, R.A. (1937) The Wave of Advance of Advantageous Genes. *Annals of Eugenics* **7**: 335-369.

[21] Gabaix, X. (1999) Zipf's Law For Cities: An Explanation. *The Quarterly Journal of Economics* **114**: 739-767.

[22] Grindrod, P. (1996) *The Theory and Applications of Reaction Diffusion Equations: Patterns and Waves.* Oxford University Press, Oxford.

[23] Hallatschek, O. (2011) The Noisy Edge of Traveling Waves. *Proceedings of the National Academy of Sciences* **108**: 17831787.

[24] Harberger, A.C. (1998) A Vision of the Growth Process. *The American Economic Review* **88**: 1-32.

[25] Hsieh, C.T., Klenow, P.J. (2006) Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics* **124**: 1403-1448.

[26] Jones, C. (1999) Growth: With or Without Scale Effects? *American Economic Review* **89**: 139-144.

[27] Klaesson, J., Karlsson, C. (2014) Market potential, start-up size and the survival of new firms. CESIS Working Paper No. 384, Royal Institute of Technology, Sweden.

[28] König, M.D., Lorenz, J., Zilibotti, F. (2012) Innovation vs Imitation and the Evolution of Productivity Distributions. CEPR Discussion Paper DP8843.

[29] Krugman, P. (1996) Confronting the Urban Mystery. *Journal of the Japanese and International Economies* **10**: 399-418.

[30] Lucas, R. (1978) On the Size Distribution of Business Firms. *Bell Journal of Economics* **9**: 508-523.

[31] Lucas, R. (2009) Ideas and Growth. *Economica* **76**: 1-19.

[32] Lucas, R., Moll, B. (2014) Knowledge Growth and the Allocation of Time. *Journal of Political Economy* **122**: 1-51.

[33] Luttmer, E. (2007) Selection, Growth, and the Size Distribution of Firms. *The Quarterly Journal of Economics* **122**: 1103-1144.

[34] Luttmer, E. (2011) On the Mechanics of Firm Growth. *Review of Economic Studies* **78**: 1042-1068.

[35] Luttmer, E. (2012a) Eventually, Noise and Imitation Implies Balanced Growth. Working Paper 699, Federal Reserve Bank of Minneapolis.

[36] Luttmer, E. (2012b) Technology Diffusion and Growth. *Journal of Economic Theory* **147**: 602622.

[37] Luttmer, E. (2014) An Assignment Model of Knowledge Diffusion and Income Inequality. Federal Reserve Bank of Minneapolis Research Department Staff Report 715.

[38] Maddison, A. (2009) "Statistics on World Population, GDP and Per Capita GDP, 1-2008 AD", 2009. Available online at http://www.ggdc.net/maddison/.

[39] Nelson, R., and Winter, S. (1982) *An Evolutionary Theory of Economic Change.* Harvard University Press, Harvard.

[40] OECD (2003) *The Sources of Economic Growth in OECD countries.* OECD Publications, Paris.

[41] Perla, J., Tonetti, C., (2014) Equilibrium Imitation and Growth. *The Journal of Political Economy* **122**: 52-76.

[42] Perla, J., Tonetti, C., Waugh, M. (2015) Equilibrium Technology Diffusion, Trade, and Growth. NBER Working Paper 20881.

[43] Polyanin, A., Zaitsev, V. (2003) *Handbook of Exact Solutions for Ordinary Differential Equations, 2nd ed.* CRC Press, Boca Raton.

[44] Prescott, E.C., Visscher, M. (1980) Organization Capital. *Journal of Political Economy* **88**: 446-461.

[45] Restuccia, D., Rogerson, R. (2008) Policy Distortions and Aggregate Productivity with Heterogeneous Plants. *Review of Economic Dynamics* **11**: 707-720.

[46] Romer, P.M. (1987) Growth Based on Increasing Returns due to Specialization. *The American Economic Review* **77**: 56-62.

[47] Schuster, P., Sigmund, K. (1983) Replicator Dynamics. *Journal of Theoretical Biology* **100**: 533-538.

[48] Simon, H. (1955) On a Class of Skew Distribution Functions. *Biometrika* **42**: 277-294.

[49] Smith, A. (1776) *The Wealth of Nations.*

[50] Staley, M. (2011) Growth and the Diffusion of Ideas. *Journal of Mathematical Economics* **47**: 470-478.

[51] Stanley, M., Amaral, L., Buldyrev, S., Havlin, S., Leschhorn, M., Stanley, H. (1996) Scaling Behavior in the Growth of Companies. *Nature* **379**: 804-806.

[52] Stigler, G.J. (1976) The Successes and Failures of Professor Smith. *The Journal of Political Economy* **84**: 1199-1213.

[53] Taylor, P, Jonker, L. (1978) Evolutionarily Stable Strategies and Game Dynamics. *Mathematical Biosciences* **40**: 145-156.

[54] Tsimring, L., Levine, H., Kessler, D. (1996) RNA Virus Evolution via a Fitness-Space Model. *Physical Review Letters* **76**: 4440-4443.

[55] van Saarloos, W. (1987) Dynamical Velocity Selection: Marginal Stability. *Physical Review Letters* **58**: 2571-2574.

[56] Whitham, G. (1974) *Linear and Nonlinear Waves.* Wiley-Interscience, New York.

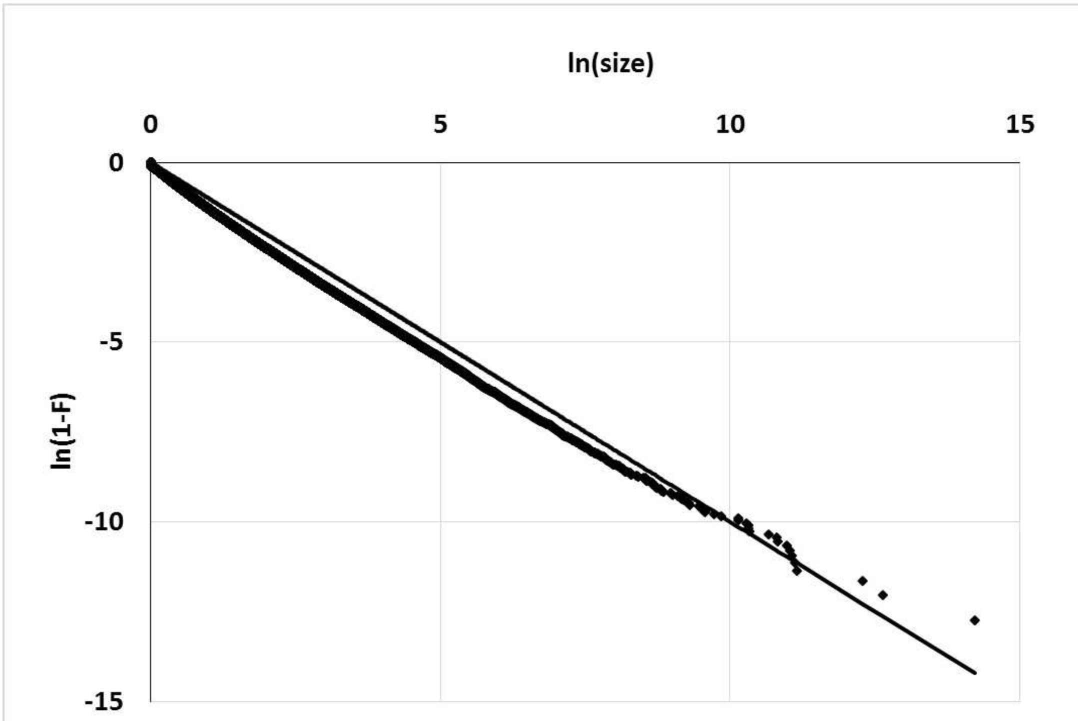

Figure 1: ***Zipf's law for firm size.*** *The x-axis is the natural logarithm of the number of employees s. The y-axis is the natural logarithm of the proportion of firms having size greater than s. The dots represent the distribution of firm size after 100 years of simulation assuming a labor force of* 5 *million (approximately three percent of the U.S. labor force), and* $s = 23$ *initially for all firms. New entrants are born with* $s = 1$. *The solid line is positioned at a 45-degree angle and represents Zipf's law. The mean squared error between the points and the 45-degree line is 0.06. The corresponding value after one thousand years is 0.006. Growth parameters used here: Volatility of Productivity* $\sigma = 0.1$, *Capital Share* $\alpha = 0.2$. *Other parameters used here and in subsequent figures: Capital Spinoff Rate* $\varepsilon = 0.05$, *Investment Rate* $\beta = 0.02$, *Capital Depreciation Rate* $\delta = 0$, *Population Growth Rate* $n = 0$.

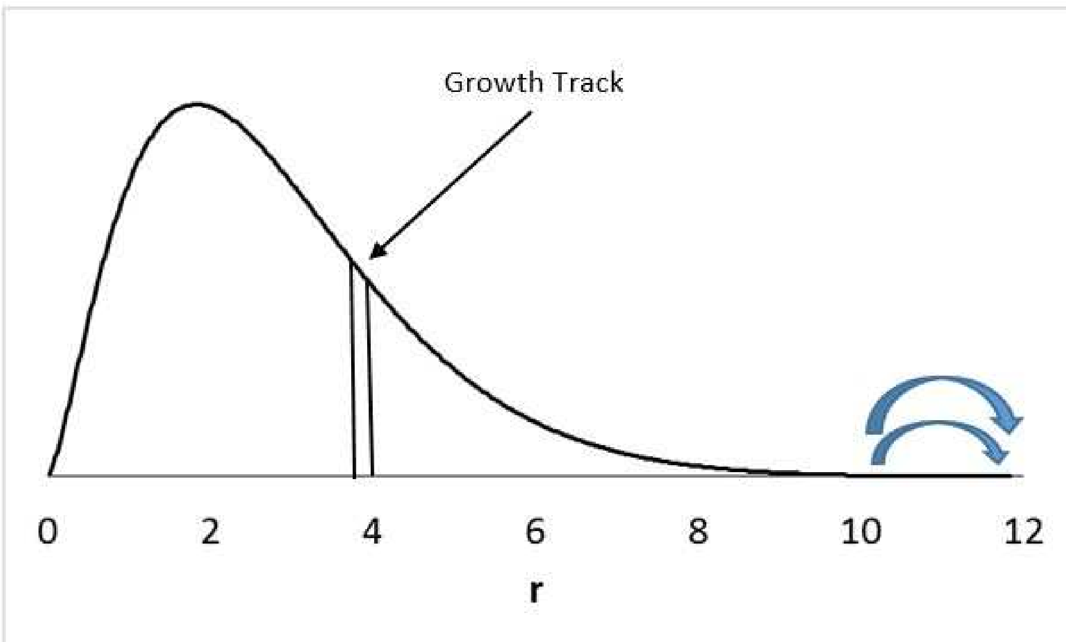

Figure 2: ***Stationary density of firms with respect to knowledge rent.*** *Here $\delta = 0.3$. Smaller values of δ lead to more positive skewness (as $\delta \to 0$ the mode converges toward zero). The thin column represents a growth track, within which the density of firms with respect to employment size s is $\rho(r)/s^2$. The arrows on the right side of the graph represent stochastic productivity shocks near the sparsely-populated frontier.*

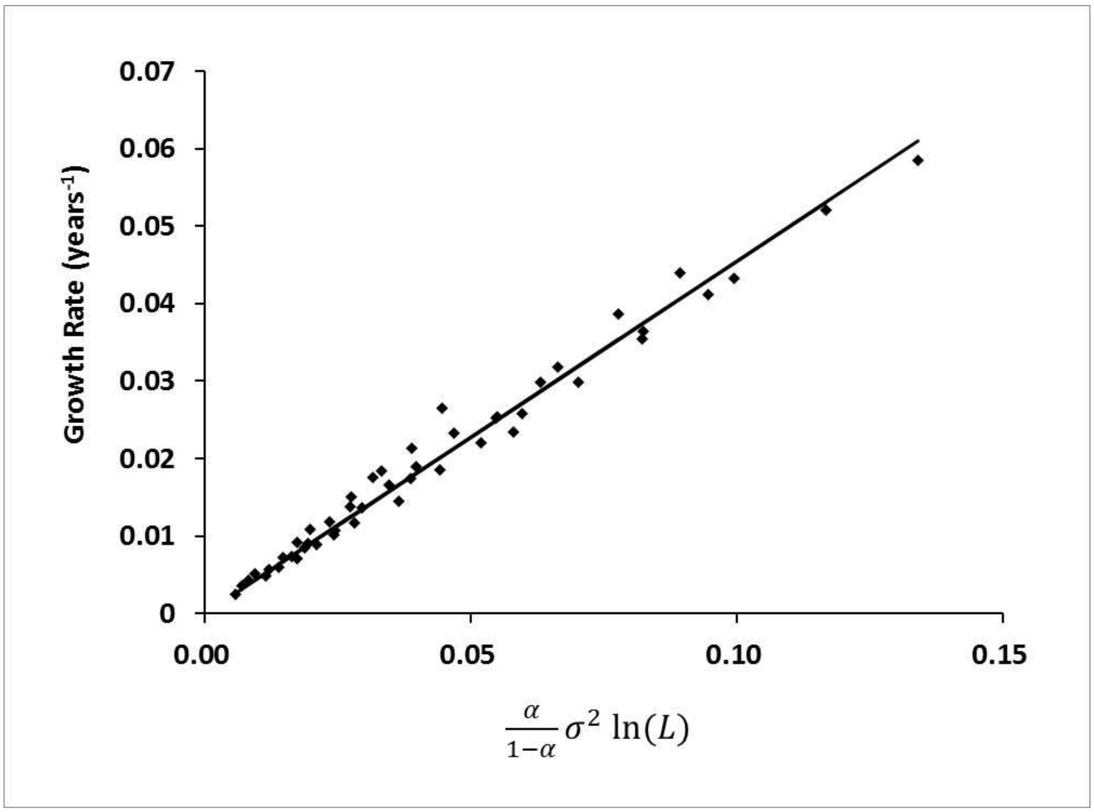

Figure 3: ***Approximation for the growth rate.*** *The growth rate of the market wage is approximately linear in* $\frac{\alpha}{1-\alpha}\sigma^2 \ln L$. *Forty-eight simulation results are shown, with* α *ranging between* 0.05 *and* 0.2, σ^2 *ranging between* 0.01 *and* 0.03, *and* L *ranging between fifty thousand and fifty million. The slope of the best fit line (shown as a solid line) is* 0.455 *with adjusted* $R^2 = 0.973$.

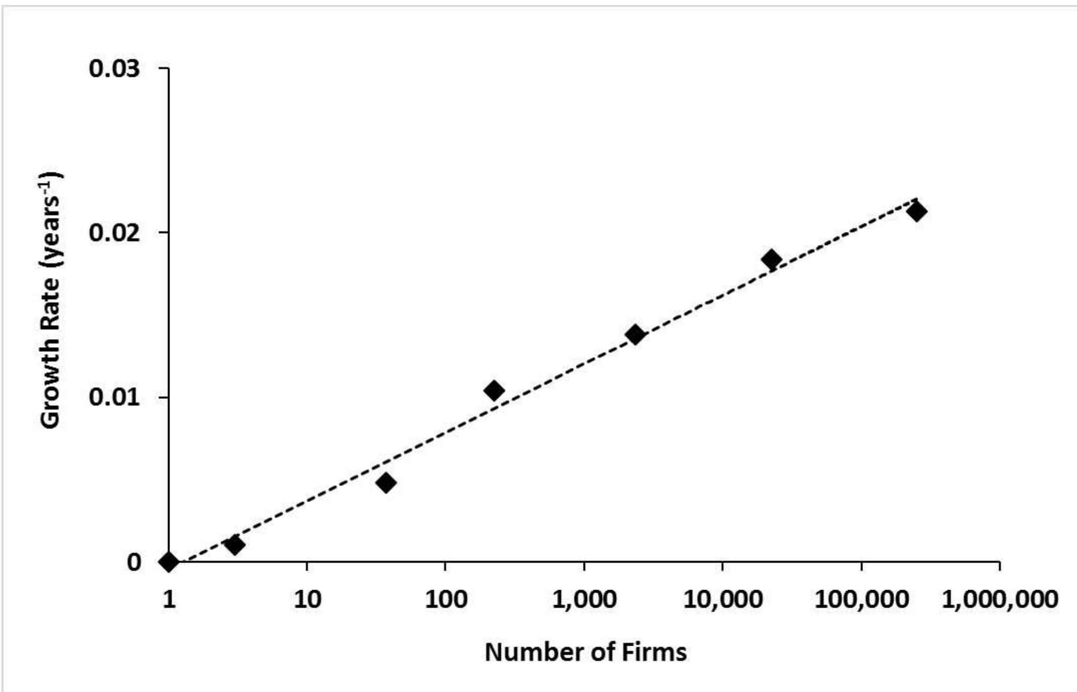

Figure 4: ***Logarithmic scale effect.*** *The growth rate of the market wage as a function of the number of firms, based on a simulation of the model. The dashed line is a best fit. Growth Parameters:* $\sigma = 0.1$, $\alpha = 0.2$.

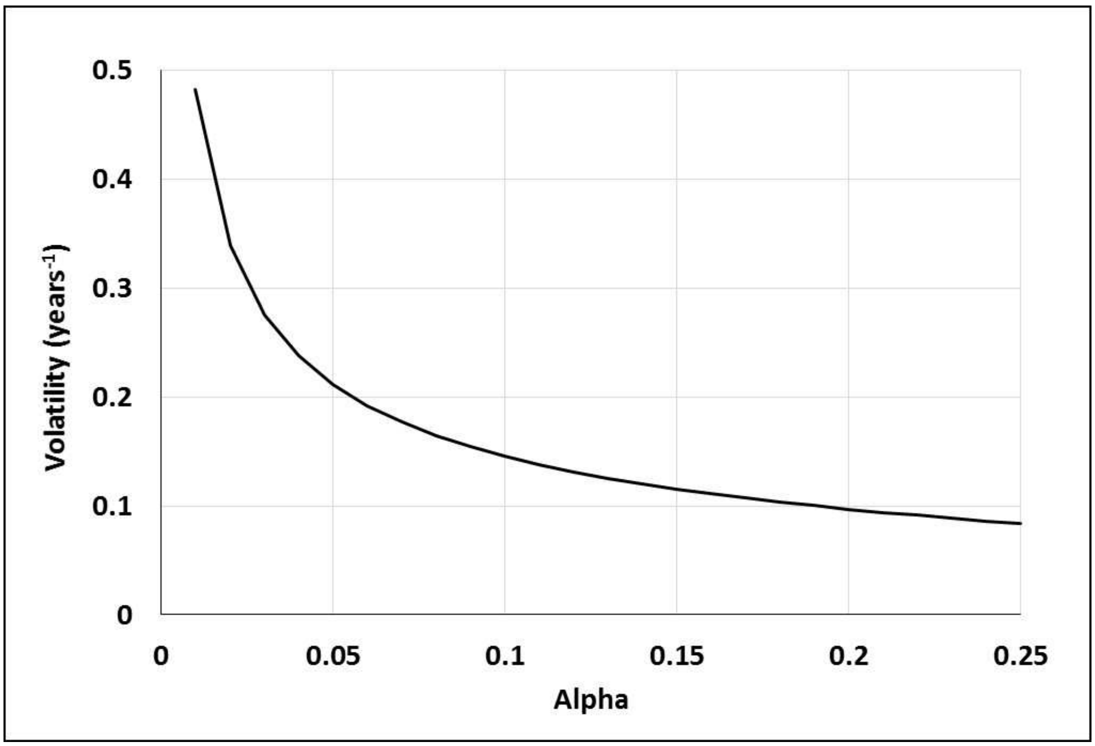

Figure 5: ***Calibration to U.S. Growth.*** *The solid line depicts the tradeoff between* α *and* σ *using Equation (27) given two observables for the U.S. economy:* $g_w = 0.0194$ *years*$^{-1}$ *and* $L_T = 77.8$ *million (geometric averages over the twentieth century: Maddison, 2009) . The labor participation rate is assumed to be* 0.5 *(Bureau of Labor Statistics: http://www.bls.gov/data).*

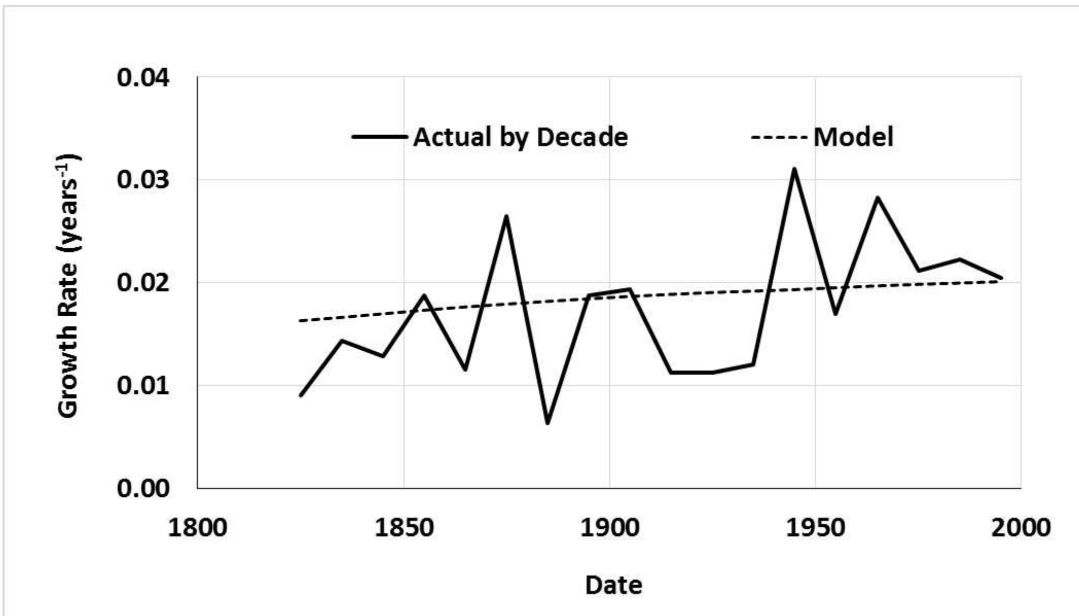

Figure 6: ***U.S. Per-Capita Income Growth: 1820-2000.*** *The dashed line represents the predicted scale effect for the U.S. economy. Sources: Maddison (2009) for population and income growth rates, Broadberry & Irwin (2006) for historical labour participation rates.*