Establishment Size Dynamics in the Aggregate Economy*

Esteban Rossi-Hansberg
Princeton University

Mark L.J. Wright
Stanford University

October 24, 2005

Abstract

Why do growth and net exit rates of establishments decline with size? What determines the size distribution of establishments? This paper presents a theory of establishment dynamics that simultaneously rationalizes the basic facts on economy-wide establishment growth, net exit, and size distributions. The theory emphasizes the accumulation of industry specific human capital in response to industry specific productivity shocks. It implies that establishment growth and net exit rates should decline faster with size, and that their size distribution should have thinner tails, in sectors that use human capital less intensively, or correspondingly, physical capital more intensively. In line with the theory, we document substantial sectoral heterogeneity in US establishments dynamics and establishments size distributions, which is well explained by variation in physical capital intensities.

*We would like to thank the editor and two anonymous referees for very helpful comments. We also thank Liran Einav, Gene Grossman, Bob Hall, Brian Headd, Tom Holmes, Boyan Jovanovic, Gueorgui Kambourov, Pete Klenow, Narayana Kocherlakota, Per Krusell, Erzo Luttmer, Steve Redding and numerous seminar participants for helpful comments, Tim Bresnahan and CEEG for financial support, Trey Cole of the US Census Bureau for his help in constructing the database, and Adam Cagliarini and Raphael Godefroy for outstanding research assistance. We aknowledge financial support from the National Science Foundation Grant # SES-0453125.

1. INTRODUCTION

Establishment size dynamics are *scale dependent*: small establishments grow faster than large establishments and net exit rates decline with size. Scale dependence in growth and net exit rates is also systematically reflected in the size distribution of establishments. In this paper we propose an explanation of this scale dependence that relies on the response of production decisions to the allocation and accumulation of industry specific human capital. In addition, the theory implies that differences in the importance of industry specific human capital, and therefore also physical capital, across sectors should lead to cross-sectoral variation in the degree of scale dependence within a sector. We present evidence from a new data-set to document these facts for the US economy. We find that, as implied by our theory, US sectors with larger physical capital shares exhibit significantly more scale dependence in establishments dynamics, net exit rates, and size distributions.

Understanding the economic forces leading to the scale dependence in establishment dynamics is important because it is informative as to the efficiency of the resource allocation mechanism in the economy. Indeed, the existence of scale dependence is sometimes cited as evidence of inefficiencies, and consequently as evidence in favor of the need for various forms of policy intervention. For example, the fact that small establishments grow faster than large establishments is often cited as evidence for the presence of imperfections in financial markets, and as justification for subsidies aimed at small businesses. Although some of the observed scale dependence may be the result of inefficiencies, an evaluation of these distortions must be done by measuring deviations from an efficient theory of establishment dynamics. These efficient theories may also imply scale dependence, and variation in scale dependence across sectors, thereby decreasing, or eliminating, the magnitude of the inefficiencies consistent with observed establishment dynamics and the size distribution. Thus, how certain we are about the mechanisms underlying establishment dynamics and the distribution of firms sizes is fundamental in evaluating the efficiency of the resource allocation mechanism in the economy.

We construct a theory of establishment dynamics that relies only on the efficient accumulation and allocation of industry specific human capital. To the extent that this approach is successful, it serves to cast doubt on the use of scale dependence to support various policies. By extension, we hope to use this theory as a way of bounding the application of such policy proposals. Our emphasis on aggregate forces such as the accumulation and allocation of resources naturally requires an aggregate model. This has a number of additional advantages. First, by emphasizing the accumulation of industry specific human capital, a force that is present across different sectors, countries, and time periods, the mechanism we propose is in turn robust to the enormous variety of institutions, technologies and regulations observed in practice. Similarly, and importantly given the ubiquity of the scale dependence facts in the data, our approach also allows us to assess the aggregate implications of our explanation.

Our basic approach is simple, and starts by noting that all of the above facts are essentially manifestations of mean reversion in the economy; indeed, the fact that small firms grow faster than large firms is an explicit statement of mean reversion. However, mean reversion in factor accumulation is a general result in macroeconomic models. We focus upon an aggregate theory of establishment dynamics based on the accumulation of industry specific human capital, because industry specific human capital is, as a result of on-the-job training and learning-by-doing, more closely tied to production conditions in the industry itself than any other factor. Under standard conditions, an abundance of human capital leads to low rates of return and slower accumulation of human capital. Conversely, if the stock of human capital is relatively low, rates of return are high and accumulation is fast. This process, which is at the heart of the resource allocation mechanism in the economy, leads to mean reversion in the stock of industry specific human capital. As long as establishments sizes respond monotonically to fluctuations in factor prices driven by the stock of human capital, mean reversion in these stocks leads to mean reversion in establishment sizes. This results in small establishments growing faster than large establishments.

The same process also implies that net exit rates decline with size. To see this, note that, given the level of employment in the industry, increases in average establishment sizes imply that some establishments exit and, conversely, decreases in establishment sizes imply that some establishment enter. The extent to which employment in the industry varies with industry shocks depends on the degree of substitutability in consumption determined by preferences. As long as the degree of substitutability is not too large, employment at the industry level does not increase/decrease enough to offset the larger/smaller establishment sizes, and so establishments exit/enter. Since small establishments grow faster than large establishments, the net exit rate is largest for small establishments and we have scale dependence in net exit rates. We can then combine the implications of the model for growth and net exit to show that in the long run the distribution of establishment sizes in a sector converges to an invariant distribution that displays scale dependence in the sense that it has thinner tails than the Pareto distribution with coefficient one.

Our emphasis on the accumulation and allocation of specific human capital in turn implies that establishment growth and exit rates should decline faster with size in sectors that use human capital less intensively. This is intuitive: the less intensively human capital is used, the faster diminishing returns to scale set in, and the faster the rate of mean reversion. In turn, this implies that the tails of the size distribution of establishments should be thinner the smaller the human capital share. Hence, the degree of mean reversion decreases with human capital intensity, just as in the neoclassical growth model the speed of convergence decreases with the physical capital share. We show that the process of entry and exit of establishments ensures that industry production will display constant returns to scale, and so physical capital intensities are negatively related to human capital intensities. This implies that the intensity of physical capital in production is positively related to the degree of mean reversion in human capital and, hence, to the degree of mean reversion in establishment sizes.

We assess the relationship between capital shares and establishment scale dependence using a new data-set commissioned from the US Census Bureau on establishment growth and net exit rates, as well as establishment size distributions, for very fine size categories and 2 digit SIC (or 3 digit NAICS) sectors. We first test the implication on growth rates and show that, as predicted by the theory, there is a positive and significant relationship between scale dependence in growth rates and physical capital shares. We then proceed to show that this same relationship is reflected in net exit rates and in significant differences in the size distribution of establishments across sectors. The differences are economically large. As one example, a doubling of the size of an establishment in the physical capital intensive manufacturing sector results in a decline in growth rates of more than half a percentage point per year, while it has little effect in the educational services sector. As another, in order to make the size distribution of establishments in the physical capital intensive manufacturing sector conform to the size distribution of establishments in the labor intensive educational services sector, we would need to take roughly three million employees (about twenty per-cent of total manufacturing employment) from medium size manufacturing establishments (between 50 and 1000 employees), and reallocate two million to very large establishments and one million to very small establishments. To the best of our knowledge, this is the first study to make use of detailed establishment size data for the entire non-farm private sector. This allows us to uncover these novel empirical regularities predicted by our theory.[1]

[1]Relatively little work has examined cross-industry differences in firm sizes. In terms of firm

Most recent theoretical attempts to explain the dynamics and size distribution of establishments generate scale dependence via selection mechanisms: unsuccessful establishments decline and exit. In Hopenhayn (1992), Ericson and Pakes (1995) and Luttmer (2004), this selection occurs as a result of sequences of bad productivity shocks, while in Jovanovic (1982) it occurs as establishments learn about their fixed productivity, and in Kortum and Klette (2003) as establishments adjust product lines in response to their own and competitors' investments in research and development. In contrast, while acknowledging that this type of effect is important for small, young establishments, we argue below that they are less relevant for the scale dependence observed across more mature, medium sized and large establishments, and abstract from them in our theory.

Another mechanism that has its main impact on small establishments is the presence of inefficiencies in financial markets as in Cabral and Mata (2003), Clementi and Hopenhayn (2002), Albuquerque and Hopenhayn (2002) and Cooley and Quadrini (2001). Still other models, for example Lucas (1978) and Garicano and Rossi-Hansberg (2004), produce a size distribution for establishments that inherits the properties of the distribution of managerial ability in the population. In contrast, our approach endogenously produces the dynamics and size distribution of establishments as the result of the efficient accumulation and allocation of factors of production.

Our theory is not the first one to successfully produce the scale dependence in establishment dynamics, net exit rates, and the size distribution observed for all establishments in the US economy. However, many of these alternate theories have very different implications for welfare and government policy. Consequently, it is necessary to find new dimensions of the data which we can use to discriminate between these theories. In this paper we propose another dimension: the variation in scale dependence with technology across sectors. We derive the empirical predictions of our theory and show that, consistent with the theory, scale dependence in growth rates, net exit rates and the size distribution increases with capital shares. None of the alternative theories has developed this prediction. Paraphrasing Jovanovic (1982), many of the mechanisms in the literature undoubtedly contribute towards an explanation of establishment dynamics. This paper shows, we believe, that the

growth rates, Audretsch et al (2002) found that Gibrat's Law is a better approximation for the Dutch services sector than it is for the manufacturing sector. In terms of entry and exit, Geroski (1983) found that gross entry and exit rates of firms are positively correlated across industries, while Geroski and Schwalbach (1991) found that turnover rankings were common across countries. Orr (1974), Gorecki (1976), Hause and Du Rietz (1984) and MacDonald (1986) all found that firm exit rates were negatively related to measures of physical capital intensity by industry.

accumulation of industry specific human capital matters too.

The rest of this paper is structured as follows. Section 2 develops our theory in detail for the case in which establishments act competitively and derives the key empirical predictions of our theory. A number of extensions, designed to show the robustness of our mechanism and its predictions to changes in the institutional environment, are presented in Section 3, along with a discussion of this link between our theory and the empirical work on specific human capital by Kambourov and Manovskii (2002). Section 4 describes our data, and presents results that show that establishment growth and net exit rates, as well as the establishment size distribution, vary with physical capital shares in precisely the way predicted by our theory. Section 5 concludes, while an appendix contains proofs of the propositions contained in the text.

2. THE MODEL

We present a stochastic dynamic aggregate model in which establishments are perfectly competitive. Labor is mobile across all industries, while both physical and human capital are specific to each industry. The model of an establishment is standard: fixed costs plus increasing marginal costs of production imply a U-shaped average cost curve, while free entry and exit of establishments ensures that all establishments in an industry operate at the bottom of their average cost curves. As the focus is upon the allocation of factors across establishments and industries, the demand side of the model is kept as simple as possible by assuming logarithmic preferences. This assumption, combined with Cobb-Douglas production functions and log-linear depreciation, ensures that we are able to solve the entire model in closed form.

2.1 Households

The economy is populated by a unit measure of identical small households. At the beginning of time, the household has N_0 members, and over time the number of members of the household N_t grows exogenously at rate g_N. Households do not value leisure and order their preferences over state contingent consumption streams $\{C_t\}$ of the single final good according to

$$(1-\delta)E_0\left[\sum_{t=0}^{\infty}\delta^t N_t \ln\left(\frac{C_t}{N_t}\right)\right], \tag{1}$$

where δ is the discount factor of the household, and E_0 is an expectation operator conditioned on information available to the household at the beginning of time. This function reflects the fact that at any point in time, each of the N_t members of the household consumes an equal share of the households consumption bundle, and that the household as a whole sums the valuations of each of its members.

The household produces the final good by combining quantities of J different intermediate goods $\{Q_{tj}\}$ according to the constant returns to scale production function

$$C_t + \sum_{j=1}^{J} X_{tj} = B \prod_{j=1}^{J} (Q_{tj})^{\theta_j} . \tag{2}$$

The final good can be used for consumption, as well as for investment in physical capital in each of the J intermediate good industries X_{tj}. We distinguish these intermediates by what we refer to as a *sector* and an *industry*. In particular, we assume that there are S sectors in this economy, and that each sector contains J_s industries, where $s = 1, ..., S$. Each industry produces a single distinct good so that there are $J = \Sigma_{s=1}^{S} J_s$ goods being produced in this economy. Sectors differ according to the methods by which output is produced and factors are accumulated; within a sector, the parameters governing production and accumulation of factors for each industry are the same. We also assume that each industry within a sector has the same share in production of the final good so that $\theta_j = \theta_i$ for all i, j in sector s. Importantly, each industry within a sector receives its own productivity shock and accumulates its own stocks of human and physical capital. This is useful below: because each industry within a sector evolves separately, according to a process governed by the same parameters, we will be able to characterize the invariant distribution of establishment sizes within each sector. In thinking about the data, we define our sectors to be roughly comparable to the list of 3 digit NAICS classifications, while our industries map into NAICS industries at a much finer level of disaggregation.

In each period, each member of the household is endowed with one unit of time which the household can allocate to work in any one of the J industries, so that the amount of time worked in industry j at time t, N_{tj} is constrained by

$$\sum_{j=1}^{J} N_{tj} \leq N_t. \tag{3}$$

Households also rent out their stocks of each of the J industry-specific physical and human capital stocks, which we denote by K_{tj} and H_{tj} respectively. Physical capital accumulates according to the log-linear form

$$K_{t+1j} = K_{tj}^{\lambda_j} X_{tj}^{1-\lambda_j}. \tag{4}$$

This log-linear form for physical capital accumulation has grown increasingly popular as a device for modelling adjustment of physical capital while still admitting closed form solutions. Here λ_j captures the importance of past physical capital stocks to the amount of capital next period: if λ_j is one, capital does not evolve and is a fixed factor; if λ_j is zero, physical capital depreciates fully each period.

Human capital is also assumed to accumulate according to a log-linear function

$$H_{t+1j} = A_{t+1j} H_{tj}^{\omega_j} I_{tj}^{1-\omega_j}.$$

Here, A_{t+1j} is an industry specific shock that is assumed to be i.i.d. with compact support $\left[\underline{A}_j, \overline{A}_j\right]$ and is designed to capture the random accumulation of knowledge within an industry, while I_{tj} is an investment in human capital accumulation. These industry specific productivity shocks are the only source of randomness in our model.[2]

We assume that I_{tj} is denominated in terms of the output of the industry itself, in order to capture the idea that industry specific learning requires some industry specific inputs, so that the resource constraint for output of industry j, Y_{tj}, is

$$Q_{tj} + I_{tj} = Y_{tj}.$$

In our framework there are no externalities: Human capital investments are paid by households, and they rent the new human capital for use in production. In Section 3, below, we also present an extension of the model which allows for learning-by-doing externalities and show that it has similar properties. Moreover, with learning-by-doing externalities, households do not appropriate the rewards to industry-specific learning, which is consistent with empirical evidence on industry specific human capital (for example, Kambourov and Manovskii 2002). The assumption that human capital accumulation responds to industry-specific production levels is essential for our results as it will serve as the primary source of industry specific mean reversion.

Finally, as noted above, we assume that the accumulation parameters are identical across all industries within a sector; that is, $\omega_j = \omega_i$ and $\lambda_j = \lambda_i$ for all i, j in sector s. The household begins with initial stocks of these specific factors denoted by K_{0j} and H_{0j}.

[2]We could have added industry specific TFP shocks, instead of shocks to the human capital accumulation equations. This would not change any of our substantive results, but comes at the cost of some substantially more complicated algebra.

2.2 Establishments

Production within each industry takes place in production units that we call establishments. To begin, for simplicity, we abstract from establishment specific heterogeneity and assume that each establishment in industry j at time t has access to the same production technology; we relax this assumption in Section 3 below. To produce in a period, the establishment must pay a fixed cost F_j that period. Once the flow fixed cost has been paid, the establishment hires industry-j-specific physical capital k_{tj}, in combination with an industry-j-specific labor input that is, in turn, produced by combining raw labor n_{tj} with industry-j-specific human capital, h_{tj}, and produces according to

$$y_{tj} = \left[k_{tj}^{\alpha_j} \left(h_{tj}^{\beta_j} n_{tj}^{1-\beta_j} \right)^{1-\alpha_j} \right]^{\gamma_j}. \tag{5}$$

Here $\gamma_j < 1$ captures the extent of decreasing returns to production which, in combination with the fixed cost, ensures that average costs are "U-shaped" and serves to pin down the size of the establishment. The parameter α_j governs the share of physical capital in value added, while β_j captures the share of human capital in the labor aggregate. Both production parameters and the process governing evolution of the productivity shock are assumed to be common across all industries within a sector: $\alpha_j = \alpha_i$, $\beta_j = \beta_i$ and $\gamma_j = \gamma_i$ for all i, j in sector s.

None of our results depend upon the denomination of the fixed cost, and so to begin we assume that it is denominated in the units of the establishments output. This has the expositional advantage of pinning down the scale of production of the plant (measured in terms of output), so that we can easily analyze the effects of changes in factor prices on the size of the establishment (measured in terms of the number of employees); we return to this assumption below.

2.3 Capital accumulation and labor allocation

To complete the characterization of the evolution of establishment sizes in this economy, all that is necessary is to characterize the evolution of productivity and factors in equilibrium. If we allow for a non-integer number of establishments, μ_{tj}, this economy satisfies all of the assumptions of the welfare theorems. As we are primarily interested in allocations, and not prices, we proceed by solving the *Social Planning Problem* for this economy: Choose state contingent sequences $\{C_{tj}, X_{tj}, I_{tj}, N_{tj}, \mu_{tj}, H_{tj}, K_{tj}\}_{t=0,j=1}^{\infty,J}$ so as to maximize household welfare

$$(1-\delta)E_0\left[\sum_{t=0}^{\infty}\delta^t N_t \ln\left(\frac{C_t}{N_t}\right)\right], \tag{6}$$

subject to the resource constraint on the final good

$$C_t+\sum_{j=1}^{J}X_{tj}=B\prod_{j=1}^{J}\left(Y_{tj}-I_{tj}\right)^{\theta_j}, \tag{7}$$

for all dates and states, the resource constraint on each intermediate good

$$Y_{tj}=\left[K_{tj}^{\alpha_j}\left(H_{tj}^{\beta_j}N_{tj}^{1-\beta_j}\right)^{1-\alpha_j}\right]^{\gamma_j}\mu_{tj}^{1-\gamma_j}-F_j\mu_{tj}, \tag{8}$$

for each industry, date and state, the accumulation equations for each industry-specific factor

$$K_{t+1j}=K_{tj}^{\lambda_j}X_{tj}^{1-\lambda_j}, \tag{9}$$

and

$$H_{t+1j}=A_{t+1j}H_{tj}^{\omega_j}I_{tj}^{1-\omega_j}, \tag{10}$$

for all industries, dates and states, and the constraint on labor allocation

$$N_t=\sum_{j=1}^{J}N_{tj}, \tag{11}$$

for all dates and states.

Inspection of this problem reveals that the choice of the number of establishments is entirely static: μ_{tj} only appears in the resource constraint for industry j at time t. This implies that we can first solve for the optimal number of establishments before solving for the dynamics of the economy. The first order condition with respect to μ_{tj} is given by

$$F_j=\left(1-\gamma_j\right)y_{tj}=\left(1-\gamma_j\right)\left[\left(\frac{K_{tj}}{\mu_{tj}}\right)^{\alpha_j}\left(\left(\frac{H_{tj}}{\mu_{tj}}\right)^{\beta_j}\left(\frac{N_{tj}}{\mu_{tj}}\right)^{1-\beta_j}\right)^{1-\alpha_j}\right]^{\gamma_j},$$

which implies

$$\mu_{tj}=\left[\frac{1-\gamma_j}{F_j}\right]^{\frac{1}{\gamma_j}}K_{tj}^{\alpha_j}\left(H_{tj}^{\beta_j}N_{tj}^{1-\beta_j}\right)^{1-\alpha_j}.$$

This leads to an equilibrium establishment size that depends on the amount of factors in the industry according to

$$n_{tj} = \frac{N_{tj}}{\mu_{tj}} = \left[\frac{F_j}{1-\gamma_j}\right]^{\frac{1}{\gamma_j}} \left(\frac{N_{tj}}{K_{tj}}\right)^{\alpha_j} \left(\frac{N_{tj}}{H_{tj}}\right)^{\beta_j(1-\alpha_j)}. \tag{12}$$

If the stock of specific factors is high relative to the amount of labor employed in the industry (which corresponds to the case of relatively cheap specific factor prices), establishments size measured in terms of the number of employees will be small. Similarly, mean reversion in the stock of relative specific factor stocks will drive mean reversion in establishment sizes. Importantly, the qualitative nature of the relationship between factor stocks and establishment size can be reversed without changing the result that mean reversion in these stocks produces mean reversion in establishments sizes. In the next section, we show that the incentive to accumulate specific factors produces precisely the required mean reversion in the general equilibrium of our model.

Substituting for the optimal number of establishments into the resource constraint gives

$$Q_{tj} + I_{tj} \leq \gamma_j \left[\frac{1-\gamma_j}{F_j}\right]^{\frac{1-\gamma_j}{\gamma_j}} K_{tj}^{\alpha_j} \left(H_{tj}^{\beta_j} N_{tj}^{1-\beta_j}\right)^{1-\alpha_j}.$$

This is our first main result: by varying the number of establishments, each of which produces at the bottom of its average cost curve, the industry behaves as though it has constant returns to scale. Hence, at the industry level (but not at the firm level) increases in physical capital shares are related to decreases in human capital shares. The result is an entirely standard log-linear multi-sector growth model with a new constant returns to scale production function.[3] As a result of the log-linear assumptions, we get the well-known result (see, for example, the appendix to Rossi-Hansberg and Wright (2004)) that income and substitution effects offset to ensure that a fixed proportion of the labor supply is allocated to each industry, a fixed proportion of the final good is consumed, while fixed proportions are invested in each industry, and a fixed proportion of the output of each intermediate input is used for investment in human capital specific to that industry.

[3] In a related paper Jones (2004) shows how a Pareto size distribution of firms leads to an aggregate Cobb-Douglas production function.

2.4 Implications for Survivor Establishment Growth, Net Exit, and the Establishment Size Distribution

With these results in hand, we can now characterize the evolution of establishment sizes in the economy. Taking natural logarithms and differences of the expression for establishment size (12) we find that the growth rate of an establishment in industry j that survives from one period to the next is given by

$$\begin{aligned}\ln n_{t+1j} - \ln n_{tj} &= \left(\alpha_j + \beta_j (1-\alpha_j)\right) g_N - \alpha_j \left[\ln K_{t+1j} - \ln K_{tj}\right] \\ &\quad -\beta_j (1-\alpha_j) \left[\ln H_{t+1j} - \ln H_{tj}\right],\end{aligned}$$

and substituting for the evolution of human capital we get

$$\begin{aligned}\ln n_{t+1j} - \ln n_{tj} &= \left(\alpha_j + \beta_j (1-\alpha_j)\right) g_N - \alpha_j \left[\ln K_{t+1j} - \ln K_{tj}\right] \\ &\quad -\beta_j (1-\alpha_j) \left[\ln A_{t+1j} + (\omega_j - 1) \ln H_{tj} + (1-\omega_j) I_{tj}\right].\end{aligned}$$

This equation reveals that the growth rate of a surviving establishment in industry j is driven by three factors. The first is the deterministic growth in the aggregate labor supply g_N which, other things equal, encourages establishments to expand in size over time. We will often assume that either population growth is zero, or that establishments growth rates are being measured relative to trend, in order to abstract from this term. The second factor is the growth in industry specific physical capital. However, as physical capital investment in each industry is a constant proportion of the aggregate production of the final good, this is also determined by aggregate forces. Over time, if the number of industries is large so that industry-specific randomness washes out in the aggregate, the aggregate economy converges to a a steady state and this term will be a constant. In what follows we assume this is the case in order to focus on industry specific variation; in general, the results that follow can be thought of as being conditioned upon the state of the aggregate economy. Finally, we have the contribution of industry specific variability, which works through the shock to human capital accumulation, and the level of industry output which affects human capital accumulation through I_{tj}: if industry output is high, human capital accumulation proceeds, on average, at a faster pace.

Before turning to a discussion of scale *dependence* in growth rates, it is useful to begin by examining the conditions under which we get scale *independent* growth or, in other words, the conditions under which we get Gibrat's Law. First, suppose we eliminate human capital as a factor of production by either reducing the importance of labor as a whole (that is, reducing $(1-\alpha_j)$) or reducing the importance of human capital in producing labor services (that is, reducing β_j). In this case, the establishment grows at a deterministic rate that is independent of scale. This is due to the fact that the only source of industry-specific randomness comes from shocks to the accumulation of human capital.[4] Second, suppose that human capital is accumulated exogenously, or that $\omega_j = 1$: this ensures that output in an industry has no effect on the pace of its human capital accumulation.[5] With the aggregate economy in steady state, the growth rate of the establishment becomes

$$\ln n_{t+1j} - \ln n_{tj} = \left(\alpha_j + \beta_j (1 - \alpha_j)\right) g_N - \beta_j (1 - \alpha_j) \ln A_{t+1j},$$

which is a constant plus an i.i.d. random variable: the growth rate of the establishment is independent of the size of the establishment.

To see how the growth rates of surviving establishments depend upon establishment size in general, assume as before that population growth is zero and that the aggregate economy is in steady state so that physical capital is constant in all industries. Then using equation (12) the growth rate of the establishment, after substituting for I_{tj}, can be written as

$$\ln n_{t+1j} - \ln n_{tj} = n_j^C - (1 - \omega_j) \left(1 - \beta_j + \alpha_j \beta_j\right) \ln n_{tj} - \beta_j (1 - \alpha_j) \ln A_{t+1j}, \quad (13)$$

where n_j^C is a constant term that depends on the physical capital stock. That is, abstracting from the population growth in steady state the theory implies that natural logarithm of firm size is given by an AR(1) process with an autoregressive coefficient given by $1 - (1 - \omega_j) \left(1 - \beta_j + \alpha_j \beta_j\right) < 1$.

We summarize the results of this discussion in the following proposition in which we emphasize the effect of changes in physical capital intensity, an observable parameter which we focus upon in our empirical analysis.

Proposition 1 *Growth rates of surviving establishments are weakly decreasing in size. The higher is the physical capital share, the faster growth rates decline with*

[4] One way to retain randomness in production while still eliminating human capital as a factor is to scale up the shock to human capital by the inverse of the elasticity of human capital in production $\beta_j (1 - \alpha_j)$. In this case, the growth rate of the firm also satisfies Gibrat's Law and becomes $\ln n_{t+1j} - \ln n_{tj} = \alpha_j g_N - \ln \hat{A}_{t+1j}$, where $\hat{A}_{t+1j}$ is the scaled shock process.

[5] If $\omega_j = 1$, human capital in industry j, and consequently also output, is difference stationary. If industry j is of positive measure, the aggregate physical capital stock will not in general converge to a steady state under this assumption. As long as $1 - \omega_j$ is positive, no matter how small, the existence of a steady state is preserved. When we refer to the case of $\omega_j = 1$ below, we shall think of $1 - \omega_j$ arbitrarily small but positive.

size. The growth rate of surviving establishments is independent of size only if either human capital is not a factor of production (in the limit as β_j or $(1-\alpha_j)$ are equal to 0), or human capital evolves exogenously (in the limit as ω_j approaches one).

The log-linearity of the model was shown above to imply that the employment allocation across industries was constant over time. Combined with the result of the above proposition, this has strong implications on net exit rates: net exit is positive whenever establishment sizes grow on average and negative when they decline. In a more general model in which the labor allocation varies in equilibrium this result continues to hold as long as the elasticity of substitution in consumption of each good is not too large. This is sufficient to guarantee that the labor allocation to the industry does not change by as much as establishment sizes. Moreover, the above proposition implies that the higher the physical capital share, the faster the net exit rate decreases with establishment size.

Corollary 2 *Establishment net exit rates are weakly decreasing in size. The higher is the physical capital share, the faster net exit rates decline with size. The net exit rate of establishments is independent of size only if either human capital is not a factor of production (in the limit as β_j or $(1-\alpha_j)$ are equal to 0), or human capital evolves exogenously (in the limit as ω_j approaches one).*

These implications for the relationship between physical capital shares, establishment growth rates and net exit can be tested directly using longitudinal data. In combination with the assumption that the distribution of establishment sizes has converged to its long-run distribution, we can also test this implication with data on the size distribution of establishments. Rossi-Hansberg and Wright (2004) showed that the combination of scale independent growth for a finite number of industries, combined with this form of entry and exit, is sufficient to generate an invariant distribution that satisfied Zipf's law: the size distribution is Pareto with coefficient one. An analogous result holds for the current framework.

Proposition 3 *(Zipf's Law) If either human capital is not a factor of production (in the limit as β_j or $(1-\alpha_j)$ are equal to 0), or human capital evolves exogenously (in the limit as ω_j approaches one), the size distribution of establishments converges to a Pareto distribution with shape coefficient one.*

Away from these limits, when there is mean reversion in establishment growth rates, the existence of a unique invariant distribution, as well as its properties, can also be established. We do this for two cases. First, we examine a case in which productivity shocks are unbounded and are drawn from a lognormal distribution. In this special case, the invariant distribution of establishment sizes can be derived in closed form, and we can study the way its variance changes with physical capital shares. Second, we characterize the invariant distribution of firm sizes for arbitrary productivity shock processes with bounded support. Here we study how the amount of dispersion in the establishment size distribution – measured by the amount of mass in the tails of the distribution – varies with the capital share. This alternate measure of dispersion has the advantage that it is less sensitive to the sizes of the very largest establishments. This is especially important for combinations of parameters that are close to the limiting cases studied in Proposition 3, where the long run variance of establishment sizes diverges.

To begin, assume that the logarithm of the productivity shock A_{tj} is distributed normally with mean M_{A_j} and variance $S^2_{A_j}$. Given the AR(1) form of the equation governing the evolution of surviving firms in (13), it is straightforward to see that the invariant distribution of representative (or average) establishment sizes in a sector, in logarithms, will be normal with mean $M_j = \beta_j (1 - \alpha_j) M_{A_j}$ and variance given by

$$var(\ln n_j) \equiv S_j^2 = \frac{\left[\beta_j (1-\alpha_j)\right]^2 S^2_{A_j}}{1 - \left[1 - (1-\omega_j)\left(1 - \beta_j (1-\alpha_j)\right)\right]^2}. \tag{14}$$

To obtain the size distribution of establishments and its variance, it is necessary to also account for the process of entry and exit or, more specifically, adjust for the fact that an industry in this sector has precisely $\mu_j = N_j / n_j$ firms of size n_j. However, the resulting actual size distribution of firms turns out also to be lognormal.

Proposition 4 *(Lognormal) If the productivity shock A_{tj} is distributed lognormally with mean M_{A_j} and variance $S^2_{A_j}$, then the long-run size distribution of establishments is lognormal with mean and variance given by*

$$e^{M_j - \frac{S_j^2}{2}} \text{ and } e^{2M_j + S_j^2}\left(e^{S_j^2} - 1\right),$$

respectively. The long run variance of the size distribution of establishments is decreasing in α_j.

In practice, for sectors with small capital shares, the empirical variance of establishments in a sector is quite sensitive to the measured size of the largest establishments, and hence to measurement error in their sizes. This should not be surprising given that, by Proposition 3, the size distribution of establishments should be close to a Pareto distribution with shape coefficient one, for which the variance diverges. This suggests we should look at other measures of dispersion in the size distribution. The assumption of lognormal shocks is also arguably quite strong. The following series of propositions characterize the invariant distribution of establishment sizes for the class of probability distributions for A with compact support, and presents a different measure of dispersion in the size distribution. In particular, the assumption that log productivity levels lie in the compact set $\left[\ln \underline{A}, \ln \overline{A}\right]$ for some $\underline{A}$ suitably small and $\overline{A}$ suitably large, and that establishment sizes are measured relative to trend (or equivalently that population growth is zero), is sufficient to guarantee that establishment sizes lie in the compact set

$$\ln n_{tj} \in LN \equiv \frac{\beta_j \left(1-\alpha_j\right)}{\left(1-\omega_j\right)\left(1-\beta_j\left(1-\alpha_j\right)\right)}\left[-\ln \overline{A}, -\ln \underline{A}\right].$$

Under this assumption, we get the following proposition.

Proposition 5 *If log productivity levels are bounded, then for any $\alpha_j, \beta_j, \omega_j \in (0,1)$, there exists a unique invariant distribution over establishment sizes in sector j.*

We also want to establish how the size distribution of establishments in a sector varies with the capital share of a sector. For any $\alpha_j, \beta_j, \omega_j \in (0,1)$, we have already established that the invariant distribution of establishments sizes has thinner tails than the Pareto distribution with coefficient one. Now we establish that we can order distributions in terms of the thinness of their tails, and can show that industries with higher physical capital shares have thinner tails. We make these notions precise in the following definition and proposition.

Definition 6 *Let λ and ψ be probability measures on $\left[\underline{b}, \overline{b}\right]$. The probability measure λ has* ***thinner tails*** *than ψ if there exists $\underline{x}$ and $\overline{x} \in \left[\underline{b}, \overline{b}\right]$ such that for all $\underline{b} \leq x \leq \underline{x}$, $\lambda\left(\left[\underline{b}, x\right]\right) \leq \psi\left(\left[\underline{b}, x\right]\right)$, for all $\underline{x} \leq x \leq \overline{x}$, $\lambda\left(\left[\underline{x}, x\right]\right) \geq \psi\left(\left[\underline{x}, x\right]\right)$, and for all $\overline{x} \leq x \leq \overline{b}$, $\lambda\left(\left[\overline{x}, x\right]\right) \leq \psi\left(\left[\overline{x}, x\right]\right)$.*

In order to apply this definition, we need to standardize the support of the size distributions produced by our model. This is also necessary to contrast the implications of our model with the data where the size categories are the same for all industries. If we scale the productivity process A_{tj} by

$$\frac{(1-\omega_j)\left(1-\beta_j(1-\alpha_j)\right)}{\beta_j(1-\alpha_j)}$$

the support of the establishment size distribution is unchanged across industries and is equal to $\left[-\ln \overline{A}, -\ln \underline{A}\right]$. Under this scaling, we prove the following proposition.

Proposition 7 *For any $\alpha_j, \beta_j, \omega_j \in (0,1)$, the invariant distribution of establishment sizes has thinner tails than the Pareto distribution with coefficient one. Other things equal, if $\alpha_j > \alpha_k$, the invariant distribution of establishments in sector j has thinner tails than the invariant distribution of establishments in sector k.*

In this section, we established that the process of accumulating industry specific human capital alone is sufficient to generate many observed properties of establishment size dynamics and establishment size distributions. In particular, mean reversion in the stock of industry specific human capital will cause small establishments to grow faster than large establishments and net exit rates of establishments to decline with size. Moreover we were also able to establish that the invariant distribution of establishment sizes would have thinner tails than the Pareto distribution with coefficient one.

As a consequence of using the accumulation of industry specific human capital to explain scale dependence, our theory also predicts that the degree of scale dependence varies with the physical capital intensity of the industry. In Section 4 below we examine this implication using US data. Before turning to the data, the next section establishes that these implications are robust to a number of different modelling assumptions that were adopted above either for simplicity or expositional reasons.

3. ROBUSTNESS OF THE MECHANISM

In the introduction we argued that it is essential that any proposed explanation for the documented patterns in establishment dynamics and size distribution be robust to the wide variety of differences in institutions and market structures for which these patterns have been observed. In this section, we establish that the mechanism described above in a particular setup survives generalization to environments in which the specification of establishment costs are different, to the introduction of establishment level heterogeneity, to alternative mechanisms for the accumulation of human capital such as learning by doing, and to an environment in which competition amongst establishments is monopolistic. In each case, we show how the general pattern of mean reversion in industry specific human capital stocks leads to mean reversion in establishments sizes, and so all our results continue to hold.

3.1 Establishment Costs

The basic mechanism of our paper relies on mean reversion in the stock of industry specific human capital of production. Mean reversion in turn leads to the mean reverting characteristics that we emphasized for establishment dynamics and size distributions. Nothing about this argument depends upon the qualitative relationship between the relative stock of factors, and the relative size of the establishment. In the model presented above, we assumed for simplicity that the establishments cost structure combined decreasing returns to scale with a fixed cost denominated in terms of the establishment's output. This combination implied that the output of the establishment was constant, so that establishments reduced employment (and hence size in terms of employment) when the stock of specific human capital grew. In other words, reversion to the mean in the stock of specific factors *from above*, produces reversion to the mean in establishment sizes *from below*.

Changes in the specification of the cost structure have the potential to reverse the qualitative relationship between factor supplies and establishment size. To see this, assume as before that each establishment in industry j at time t produces output according to equation (5). Now, however, assume that fixed costs depend on the average number of workers hired in industry j at time t, $\bar{n}_{tj}$. In particular, assume that fixed costs are given by $F_j \bar{n}_{tj}^{\xi_j}$. We have in mind institutional or organizational costs (for example dealing with unions or other industry organizations) that depend on the average size of establishments in the industry. Individual establishments do not take into account the effect of their hiring decisions on the fixed costs, so the problem of the establishment is identical to the one presented above. The problem of the establishment is to maximize profits

$$\max_{k_{tj}, h_{tj}, n_{tj}} \Pi \equiv \max_{k_{tj}, h_{tj}, n_{tj}} y_{tj} - r_{tj} k_{tj} - s_{tj} h_{tj} - w_{tj} n_{tj} - F_j \bar{n}_{tj}^{\xi_j},$$

where r_{tj}, s_{tj}, w_{tj} denote the corresponding factor prices. We assume that $0 \leq \xi_j < 1$

and so if $\xi_j = 0$ we have the same case studied above. Taking first order conditions and allowing for free entry and exit so that profits are zero implies

$$\left(1-\gamma_j\right) y_{tj} = F_j \bar{n}_{tj}^{\xi_j}.$$

Now output changes with the average level of employment in the industry and, since in equilibrium all establishments are identical, also with the employment level of the establishment. Given this symmetry, equilibrium in factor markets implies that the size of the typical establishment in the industry is given by

$$n_{tj} = \frac{N_{tj}}{\mu_{tj}} = \left[\frac{\left(1-\gamma_j\right)}{F_j}\right]^{\frac{1}{\xi_j-\gamma_j}} \left(\frac{N_{tj}}{K_{tj}}\right)^{\frac{\alpha_j\gamma_j}{\gamma_j-\xi_j}} \left(\frac{N_{tj}}{H_{tj}}\right)^{\frac{\beta_j\left(1-\alpha_j\right)\gamma_j}{\gamma_j-\xi_j}}.$$

This equation is analogous to the case considered above with a pure fixed cost. The main differences are that now both employment and output respond to changes in factor supplies.[6] Moreover, the direction of the change can differ: for $\xi_j < \gamma_j$, the behavior of employment is as before, declining with the industry physical and human capital stocks; for $\xi_j > \gamma_j$ this pattern is reversed and the size of establishments depends positively on the stock of both types of capital but negatively with industry employment. In either case, the main properties for establishment growth and exit rates, and the size distribution, are preserved: regardless of whether establishments in industries with large human capital stocks are large or small they revert to the mean. The example illustrates that the necessary property of establishment sizes is that they respond monotonically to the stock of human capital in the industry. The direction of this response is not important: in the case where $\xi_j > \gamma_j$, reversion to the mean in the stock of specific factors *from above*, produces reversion to the mean in establishment sizes *from above*. Mean reversion in the stock of human capital then leads to the same arguments and results we presented in previous sections.

3.2 Within Industry Establishment Heterogeneity

In the theory presented above, we abstracted from heterogeneity amongst establishments *within* an industry in order to focus our attention on heterogeneity *across* industries. This allowed us to emphasize the contribution of the accumulation of industry specific human capital to the evolution of establishment sizes. Clearly, there

[6] Notice that because the fixed costs entail an external cost, the equilibrium will not be Pareto optimal. However, one can set up a pseudo-social planner problem that yields the same aggregate implications than the problem discussed in the Section 2 (see also Section 3.4).

exist differences in establishment sizes even within narrowly defined industries. While this may be caused by aggregation (data is rarely available beyond the three or four digit SIC levels), it is probable that some establishment specific heterogeneity remains. In this section we demonstrate how establishment specific heterogeneity can be added to our framework, and show that it does not change the key empirical implications of our theory for the differences in establishment dynamics and size distributions across industries.

Consider the model of Section 2, where we suppress time and industry subscripts. Suppose that after having decided to produce in a period (that is, after paying the fixed cost F) each establishment $i \in [0, \mu]$ observes a establishment specific productivity shock z_i. This shock is assumed to be i.i.d. over time, establishments, and industries within a sector. After observing this shock, the establishment i can then hire labor n_i and industry-j-specific physical, k_i, and human capital, h_i, to produce output according to

$$y_i = z_i \left(k_i^{\alpha} \left[h_i^{\beta} n_i^{1-\beta} \right]^{1-\alpha} \right)^{\gamma}.$$

To see how this affects the results, we consider once again the social planners problem. To begin, suppose that the planner has decided that there are μ establishments in the industry employing N workers. The amounts of industry specific physical and human capital are fixed at K and H. The planner then observes the identities of the establishments that receive each productivity shock. The problem of the planner is then to allocate factors across establishments in the industry to maximize industry output

$$\int_0^{\mu} z_i \left(k_i^{\alpha} \left[h_i^{\beta} n_i^{1-\beta} \right]^{1-\alpha} \right)^{\gamma} di,$$

subject to

$$\int_0^{\mu} k_i di \leq K, \qquad \int_0^{\mu} h_i di \leq H, \qquad \int_0^{\mu} n_i di \leq N.$$

We assume that we can index the productivity shock by the unit interval with density ϕ with mean normalized so that $\int_0^1 z_i^{\frac{1}{1-\gamma}} \phi\,(di) = 1$ and that the appropriate Law of Large Numbers holds for continua of i.i.d. random variables. Then this problem becomes one of maximizing

$$\mu \int_0^1 y_i \phi\,(di)\,,$$

subject to

$$\mu \int_0^1 k_i \phi\,(di) \leq K, \qquad \mu \int_0^{\mu} h_i \phi\,(di) \leq H, \qquad \mu \int_0^1 n_i \phi\,(di) \leq N.$$

The first order conditions for this problem imply that the relative levels of output and factors are given by

$$\frac{y_i}{y_j} = \frac{k_i}{k_j} = \frac{h_i}{h_j} = \frac{n_i}{n_j} = \left(\frac{z_i}{z_j}\right)^{\frac{1}{1-\gamma}}.$$

That is, establishments within an industry with a higher shock use more of both inputs and produce more output. Actual amounts used in each establishment can be determined from the resource constraint so that

$$\frac{k_i}{K} = \frac{h_i}{H} = \frac{n_i}{N} = \frac{z_i^{\frac{1}{1-\gamma}}}{\mu \int_0^1 z_i^{\frac{1}{1-\gamma}} \phi(di)} = \frac{z_i^{\frac{1}{1-\gamma}}}{\mu}.$$

With these results, we can characterize the level of output in the industry given the initial choice of the number of establishments μ, the choice of labor N, and previously accumulated physical and human capital K and H as

$$\mu \int_0^1 z_i \left(k_i^\alpha \left[h_i^\beta n_i^{1-\beta}\right]^{1-\alpha}\right)^\gamma \phi(di) = \left(K^\alpha \left[H^\beta N^{1-\beta}\right]^{1-\alpha}\right)^\gamma \mu^{1-\gamma}$$

From this equation, it is easy to see that the form of the industry production function is exactly the same as for the original problem, and consequently that the choices of N and μ, as well as investment in both types of capital, are analogously determined.

Clearly, the addition of an i.i.d. productivity shock has no effect on the mean growth and net exit rates of establishments in that industry. Consequently, the model has the same implications for growth and net exit at the sector level. Further, the distribution of *average* establishment sizes is unchanged, and so the relationship between factor intensities and the shapes of the establishment size distribution is unchanged. One implication that can be affected is the range of cases under which Zipf's Law exactly holds: when the conditions of Proposition 3 hold, we observe Zipf's Law for average establishment sizes, but only for actual establishment sizes if either all establishments are identical within an industry, or if the distribution within an industry is also Pareto with coefficient one. We might think of the latter as being produced by a similar mechanism as the one laid out in this paper, working through *enterprise specific* human capital.

3.3 Learning-by-Doing Externalities

In the model of Section 2, we assumed that human capital accumulation required some industry specific inputs. The dependence on industry-specific inputs was important for our model, as it allows human capital accumulation to vary with output in the industry, and is the primary source of mean reversion at the industry level.

In that model, the inputs to learning were purchased by consumers, and the resulting level of human capital was rented out by consumers, so that there was no externality. An alternative assumption that has similar effects is the assumption that human capital is accumulated from *learning-by-doing externalities* of the form

$$H_{t+1j} = A_{t+1j} H_{tj}^{\omega_j} Y_{tj}^{1-\omega_j},$$

which states that the higher is output in the industry, the higher is accumulation of human capital. Importantly, this involves no resource cost to the economy. Suppose also that production occurs according to

$$Y_{tj} + F_j \mu_{tj} = \left[K_{tj}^{\alpha_j} \left(H_{tj} N_{tj} \right)^{1-\alpha_j} \right]^{\gamma_j} \mu_{tj}^{1-\gamma_j},$$

so human capital operates exactly like labor augmenting technological progress.

Although we can no longer use the social planners problem to solve for equilibrium allocations in this model, we can use a pseudo-planner problem to solve it as we do in Subsection 3.4. Similar reasoning then produces an expression for the normalized rate of growth of an establishment given by

$$\ln n_{t+1} - \ln n_t = n^C - \alpha_j \left(1 - \omega_j\right) \ln n_t - \left(1 - \alpha_j\right) \ln A_{t+1},$$

where n^C again denotes a constant specific to this formulation. If there is no learning by doing, or $\omega_j = 1$, there is no mean reversion in human capital stocks, and establishment growth rates satisfy Gibrat's Law. As before, increases in the capital intensity of an industry increase the rate of mean reversion in establishment sizes.

This extension emphasizes that it is not industry specific human capital *per se*, but rather the sensitivity of current production decisions to past output in the industry, that is important for our results on mean reversion. This is important in the light of recent research by Kambourov and Manovskii (2002) who argue that there is little evidence for industry-specific human capital in individual earnings data.[7] However, this evidence is consistent with industry-specific learning-by-doing externalities where individual workers do not appropriate the returns to industry-specific human capital.

[7] Our model makes no distinction between workers within an industry, and so cannot distinguish between industry-specific human capital and the occupation-specific human capital emphasized by Kambourov and Manovskii (2002).

3.4 Monopolistic competition

The previous model uses an extremely simple theory of the establishment to derive conclusions on the size distribution of establishments. In this section we use a different theory of the establishment to show that the conclusions derived above are not specific to that particular theory of the organization of production in establishments. For this we use the Dixit-Stiglitz monopolistic competition model with taste for variety. In this model substitution for varieties in the same industry limits demand for a particular variety in an industry and therefore determines the size of the establishment. The model includes naturally the two margins we have emphasized so far, the number of establishments in an industry and the size of these establishments. We need a version of this theory where both margins react to factor accumulation. In particular, a theory that includes the three factors that we introduced in the model above. Now, physical and human capital are specific to an industry but mobile across varieties within that industry.

3.4.1 Households.—
As above, we assume that there are J industries divided into sectors with similar technologies. Now, however, we assume that each industry consists of a continuum of potential varieties which we index by ϖ. Households provide labor and industry-specific (but *not* variety-specific) physical and human capital to each variety within an industry. Output of each variety D_{tj}^{ϖ} is combined by the household using a constant elasticity of substitution production function with parameter $\sigma_j > 1$ to produce a composite industry good that is used for investment in human capital and as an input to production of a final good (in combination with the composite goods of other industries) that is consumed and invested in physical capital.

That is, the problem of a consumer is to purchase goods and accumulate industry specific capitals to maximize lifetime utility, or

$$\max_{D_{tj}^{\varpi}, N_{tj}, C_{tj}, X_{tj}} (1-\delta) E_0 \left[\sum_{t=0}^{\infty} \delta^t N_t \ln\left(\frac{C_t}{N_t}\right) \right]$$

subject to

$$E_0 \left[\sum_{t=0}^{\infty} \sum_{j=1}^{J} \int_{0 \le \varpi \le \Omega_{tj}} p_{tj\varpi} D_{tj\varpi} d\varpi \right] \le E_0 \left[\sum_{t=0}^{\infty} \sum_{j=1}^{J} r_{tj} K_{tj} + s_{tj} H_{tj} + w_{tj} N_{tj} \right],$$

$$K_{t+1j} = K_{tj}^{\lambda_j} X_t^{1-\lambda_j}, \qquad H_{t+1} = A_{t+1j} H_{tj}^{\omega_j} I_{tj}^{1-\omega_j}$$

$$Q_{tj} + I_{tj} \equiv E_{tj} \leq \left\{ \int_{0 \leq \varpi \leq \Omega_{tj}} \left(D_{tj\varpi} \right)^{\frac{\sigma_j - 1}{\sigma_j}} d\varpi \right\}^{\frac{\sigma_j}{\sigma_j - 1}},$$

$$C_t + X_t = \prod_{j=1}^{J} \left(Q_{tj} \right)^{\theta_j}, \qquad \sum_{j=1}^{J} N_{tj} \leq N_t.$$

for all t and all j, where E_{tj} is total demand for the final good from industry j, and Q_{tj} is the amount of the final good in industry j used to produce consumption and physical capital investment in combination with the goods in other industries. The consumer takes as given the prices of intermediate inputs and factors, as well as the range of varieties of goods available.

In order to solve the establishment's problem below, it is useful to record that the first order conditions of the consumer's problem with respect to a variety implies a demand for variety ϖ in industry j of

$$D_{tj}^{\varpi} \left(p_{tj}^{\varpi} \right) = E_{tj}^{\varpi} \frac{\left(p_{tj}^{\varpi} \right)^{-\sigma_j}}{\int\limits_{0 \leq \varpi \leq \Omega_{tj}} \left(p_{tj}^{\varpi} \right)^{1-\sigma_j} d\varpi},$$

where Ω_{tj} is the measure of varieties that make positive profits and therefore produce in equilibrium in industry j at time t, which consumers take as given.

3.4.2 Establishments and industry equilibrium.—

An establishment producing a variety ϖ uses a constant returns to scale Cobb-Douglas technology with labor, physical, and human capital as factors of production, given by

$$y_{\varpi} = k_{\varpi}^{\alpha} \left[h_{\varpi}^{\beta} n_{\varpi}^{1-\beta} \right]^{1-\alpha},$$

We suppress the time and industry subscripts whenever this does not lead to confusion. The first stage of the problem of the establishment is to minimize costs,

$$\begin{aligned} C\left(r, s, w, D_{\varpi}, F\right) &\equiv \min_{K_{tj}^{\varpi}, L_{tj}^{\varpi}} rK_{\varpi} + sH_{\varpi} + wN_{\varpi} \\ s.t.\ D_{\varpi} + F &= k_{\varpi}^{\alpha} \left[h_{\varpi}^{\beta} n_{\varpi}^{1-\beta} \right]^{1-\alpha}, \end{aligned}$$

where D_{ϖ} is the quantity demanded of the variety and F is a fixed cost of production. The cost function of the problem then becomes

$$C\left(r, s, w, D_{\varpi}, F_j\right) = \lambda \left(D_{\varpi} + F \right),$$

where

$$\lambda = \left(\frac{r}{\alpha}\right)^{\alpha} \left(\frac{s}{\beta\left(1-\alpha\right)}\right)^{\beta(1-\alpha)} \left(\frac{w}{\left(1-\beta\right)\left(1-\alpha\right)}\right)^{(1-\beta)(1-\alpha)}.$$

Notice that average costs $C\left(r, s, w, D_{\varpi}, F_j\right)/D_{\varpi}$ are a decreasing function of D_{ϖ}.

The second stage of the establishment problem is to maximize profits

$$\Pi\left(r, s, w, F\right) = \max_{p_{\varpi}} D_{\varpi}\left(p_{\varpi}\right) p_{\varpi} - C\left(r, s, w, D_{\varpi}\left(p_{\varpi}\right), F\right),$$

where $D_{\varpi}\left(p_{\varpi}\right)$ is derived from the consumers problem and stated above.

The first order conditions of the establishment problem then imply that $p_{\varpi} = \lambda\sigma/\left(\sigma-1\right)$. Hence in equilibrium the levels of production and profits by establishments are given by

$$D_{\varpi}\left(p_{\varpi}\right) = \frac{E}{\Omega\lambda}\frac{\sigma-1}{\sigma} \qquad \text{and} \qquad \Pi\left(r, s, w, F\right) = \frac{E}{\sigma\Omega} - F\lambda.$$

Zero profits then implies that the number of varieties (or establishments since only one establishment produces each variety) is given by $\Omega = E/\left(\sigma F\lambda\right)$ and so

$$D_{\varpi}\left(p_{\varpi}\right) = F\left(\sigma-1\right).$$

The equilibrium conditions in factor markets are given by

$$K = \frac{E\alpha}{r}, \qquad H = \frac{E\beta\left(1-\alpha\right)}{s}, \qquad N = E\frac{\left(1-\beta\right)\left(1-\alpha\right)}{w},$$

which implies that

$$\lambda = EK^{-\alpha}H^{-\beta(1-\alpha)}N^{-(1-\beta)(1-\alpha)} \text{ and so } \Omega = \frac{K^{\alpha}H^{\beta(1-\alpha)}N^{(1-\beta)(1-\alpha)}}{\sigma_j F_j}.$$

Output in the industry is given by

$$Y = \Omega D_{\varpi}\left(p_{\varpi}\right) = \frac{\sigma-1}{\sigma}K^{\alpha}H^{\beta(1-\alpha)}N^{(1-\beta)(1-\alpha)}.$$

Notice that this function is constant returns to scale, with TFP given by a function of the elasticity of substitution.

The size of establishments in terms of employees is given by

$$n_{\varpi} = F\sigma\left(\frac{N}{K}\right)^{\alpha}\left(\frac{N}{H}\right)^{\beta(1-\alpha)},$$

which has a very similar form to the one derived for the case of perfect competition above. As a result, the model has identical implications for the dynamics and size distribution of establishment sizes.

3.4.3 Capital accumulation, labor allocation and establishment sizes.— All that remains is to calculate the accumulation decisions of agents. Although this can be done directly from the agents decision problem, it is instructive to compute them in an analogous way to the allocations for the perfectly competitive economy discussed above. Although the welfare theorems do not hold for this economy, the fact that the markup of these monopolistic establishments is constant combined with the log-linearity of the model means that the equilibrium allocations can be obtained as the solution of an *equivalent optimum problem* that is identical to the social planners problem used above, except that the resource constraint is now

$$C_t + X_t \leq \prod_{j=1}^{J} \left(\frac{\sigma_j - 1}{\sigma_j} K_{tj}^{\alpha} H_{tj}^{\beta(1-\alpha)} N_{tj}^{(1-\beta)(1-\alpha)} - I_{tj} \right)^{\theta_j},$$

for all t and j (see Chapter 18 of Stokey, Lucas and Prescott (1989) for another example of this pseudo-economy approach). As before, the solution of this model has the household accumulating a fixed proportion of the output of each industry to produce investment in physical and human capital. The allocation of labor in each industry is fixed at the same levels as before. From these results it is straightforward to show that the evolution of establishment sizes in the model with monopolistic competition is identical (with $\gamma = 1$) to the evolution of establishment sizes in the model with perfect competition. In particular, analogues of Propositions 1, 3, 4, 5, and 7 and of Corollary 2 continue to hold.

4. EVIDENCE ON SCALE DEPENDENCE

The model above has a number of empirical implications consistent with findings in the empirical literature on scale dependence in establishment sizes and dynamics. Most notably, the model can reproduce the facts that establishment growth and net exit rates decline with size, and that the size distribution has thinner tails than the Pareto with shape coefficient one. However, and more importantly, our model has strong implications for the variation of scale dependence across sectors. In our theory the degree of reversion to the mean in human capital stocks, and therefore in establishments sizes, increases with the degree of diminishing returns in human capital, or equivalently decreases with the degree of diminishing returns in physical capital. A low physical capital share implies a high human capital share, hence a low degree of diminishing returns in human capital and, therefore, a low degree of reversion to the mean in establishment sizes. As the physical capital share increases from zero the degree of diminishing returns in human capital increases as does the reversion to the mean in establishment sizes. Hence, the model implies that the degree of scale dependence in growth rates and in net exit rates, and the thinness of the tails of the size distribution, are intrinsically determined by the importance of industry specific physical capital in technology. In this section we present results from a new dataset on establishment dynamics and size distributions for the private non-farm US economy. This dataset is novel in that it provides data for a wider range of sectors and industries than previously available, and so we begin by first verifying these facts for the entire US economy. We then turn to an examination of the sectoral predictions of our model using sectoral US data.

4.1 Data Sources and Economy-wide Scale Dependence

We investigate scale dependence and its variation across sectors using data on growth rates, exit and entry rates (and so net exit rates), and the distribution of establishment sizes. We use two data-sets constructed specially for this project by the US Census Bureau. These new data sets have several advantages for our purposes in comparison with the publicly available data sources. First, they provide the number of establishments per size category for the finest size categories that the US Census will release given their confidentiality restrictions. Because of our emphasis on the shape of the size distribution, this level of detail is crucial. Previous analysis of the size distribution of establishments have tended to use data for much larger size bins or only for a couple of sectors. Second, these data include all sectors in the private non-farm US economy, including both manufacturing and services. This is important for our study given that we want to understand the effect of sectoral differences in physical capital shares on the size distribution of establishments. Variations in physical capital shares are much larger across service and manufacturing sectors than within them. Third, the data refers to establishment sizes, and not enterprise sizes, which as we have argued is a better fit for our theory.

4.1.1 Cross Section Data and the Size Distribution.— The first source is a data-set from the Statistics of US Businesses (SUSB) program on establishment size distributions by sector at the two digit SIC level for 1990 and three digit NAICS level for 2000. These data are constructed from a number of sources including the annual County Business Profile (CBP) data files. We use these data to examine, following a large literature beginning with Gibrat (1931), the size distribution of establishments. Figure 1 illustrates the scale dependence in the size distribution of establishments by comparing the densities of establishment sizes (employment at operations at a single location) and enterprises (employment at operations under common ownership or control) for the US economy in 2000 to a commonly used benchmark: a Pareto distribution with shape coefficient one (see, for example, Axtell (2001)). The Pareto distribution is scale independent in the sense that the distribution is invariant to truncation of the left tail. The figure shows that the enterprise and establishment size distributions are similar, reflecting the fact that only the very largest enterprises possess more than a single establishment. Importantly, both distributions have thinner tails than the Pareto benchmark: there is scale dependence in the size distribution of both establishments and enterprises. In Figure 2, we present these data in a different format in order to emphasize the right tail of the distribution. If production units are distributed according to a Pareto distribution, the logarithm of the share of production units greater than a particular employment size varies linearly with the logarithm of employment. If the Pareto distribution has a shape coefficient of one, the slope of the line is minus one. If, however, the tails of the actual distribution are thinner than the tails of a Pareto distribution, as in Figure 1, the relationship is concave and not linear.

First note that the tails of the size distribution of establishments are clearly thinner than the Pareto distribution, since the corresponding curve in Figure 2 is clearly concave. The distribution of enterprises is closer to the Pareto, especially if we focus attention on enterprises with between 50 and 10000 employees. However, if we observe the whole distribution it exhibits thinner tails than the Pareto. To highlight the previous statement in Figure 2 we include data for enterprises with close to one million employees. The similarities and differences between these two distributions may shed some light on the forces that determine the boundaries of the firm. This topic, although fascinating, is beyond the scope of this study. The theory we develop refers to the technology of a single production unit and does not address questions of ownership or control. Consequently throughout the rest of the paper we focus solely on establishment data.

4.1.2 Longitudinal Data, Growth and Net Exit.—

To examine establishment dynamics, the second data-set, from the Business Information Tracking System (BITS), contains data on growth rates of establishments between 1990 and 2000, and deaths and births of establishments by size category for 1995-1996. The unique aspect of the longitudinal data-set is that it tracks the size of establishments for several years, and, for exiting/entering establishments, for three years before/after they exit/enter. First, we examine the well-known stylized fact that small establishments grow faster than large establishments, when attention is restricted to those establishments that remain in operation.[8] This is illustrated in Figure 3 which plots growth rates by establishment size for the US over both one and ten year intervals. This figure shows that the difference in growth rates between small and large establishments can be as large as twenty per-cent within a year, and that the accumulated effect of this pattern over a decade leads to differences of more than one-hundred per-cent between small and large establishments. Moreover, this scale dependence in growth rates is not limited to the smallest establishments, and is significant throughout the size distribution. Note that Figure 3 presents data on establishments that survived the relevant period, hence selection may be a relevant force explaining the exhibited scale dependence. This is an important issue that we address in detail below.

In a typical period, a substantial fraction of production units turn over: some units exit, while new ones are created.[9] The second fact that we examine with this dataset, the scale dependence in net exit rates, is illustrated in Figure 4 which plots the cohort of establishments that turned over between 1995 and 1996. Clearly, net exit rates decline with size, even for establishments with more than 1000 employees.[10]

Figures 1 through 4 show important degrees of scale dependence in the size distribution, growth and net exit rates in the US economy. This scale dependence is documented using data aggregated by size category. Our theory predicts that we should observe this scale dependence for representative firms across industries. As we have argued, our notion of an industry is very narrow, since an industry includes only

[8]This fact was most forcefully demonstrated by Mansfield (1962) in his study of firms in the steel, petroleum, tire, and automobile industries. More recent work by Hall (1987) and Evans (1987a,b) using data on firms, and by Dunne, Roberts and Samuelson (1989a,b) on manufacturing plants, has confirmed this finding. See also the surveys by Scherer (1980), Geroski (1995), Sutton (1997), and Caves (1998), who document the robustness of these results across time, industries, and countries.

[9]Mansfield (1962) was one of the first to emphasize the importance of turnover and to find scale dependence in exit rates: small establishments are more likely to exit than large establishments.

[10]Figure 4 only shows establishments between 50 and 2,500 employees. Smaller establishments exhibit a somewhat more erratic behavior and we have less than 10 exits/entries per size bin for larger establishments (we include all sizes in the empirical exercises across sectors below). Figure 4 shows that the logarithmic trend of net exit rates is below zero for all size categories. This means that throughout the US economy on average there was net entry of establishments with between 50 and 2500 employees. Our theory could explain this as the result of economic or population growth.

firms that produce the same goods, and use exactly the same technology and physical and human capital. We do not have data disaggregated at this level. Hence, throughout this section we interpret each establishment in our data set as a representative establishment in a narrowly defined industry with the number of establishments per industry given by our theory.[11]

4.1.3 Selection and Age Effects.—

The theory outlined in Section 3 makes specific predictions for the growth rate of establishments, *conditional upon their survival*. It also makes predictions about the behavior of the net exit rate of establishments, and about the size distribution of establishments. Consequently, in the empirical analysis below, we focus, separately, on conditional growth rates, net exit rates, and size distributions. The focus on conditional growth rates contrasts with the empirical literature testing Gibrat's law which has emphasized establishment growth rates that do not condition on survival, and in particular the role of exit in reducing the unconditional growth rate of small firms. There are three reasons why we do not take this approach here. First, as noted, our theory makes specific predictions for both growth rates conditional upon survival, and on net exit, and so we examine both directly. Second, the implications of all theories are sensitive to the precise way in which the growth rate of exiting firms is treated, and whether or not entering firms are also included. Moreover, there is no clear consensus as to the appropriate way to include entry, as evidenced by the alternative empirical methodologies of Dunne, Roberts and Samuelson (1989) and Davis and Haltiwanger (1999). The theory of the current paper continues to predict scale dependent under either of these methodologies. However, the fact that the same mechanism causes the scale dependence in conditional growth rates and net exit rates means that there exists yet further treatments of entry and exit that give rise to the prediction that unconditional growth rates display no scale dependence. This leads to our third reason: by focusing upon these facts separately, it is possible to directly examine whether or not the degree of scale dependence in both net exit and conditional growth rates varies across sectors as our theory predicts.

This focus also serves to further distinguish our approach form studies that emphasize selection mechanisms in producing scale dependent growth. Although we acknowledge that selection effects may be important for small establishments, we interpret the evidence as suggesting that they are less relevant for the scale dependence

[11] A theoretically consistent empirical decomposition between industry and firm heterogeneity requires unit record data which is not available for a broad sample like ours.

observed across medium sized and large establishments. For example, one important prediction of selection theories, which is at variance with our own theory, is that firms should become smaller in the years prior to exit, which is often referred to as the "Shadow of Death" (Griliches and Regev, 1995). To assess this, Figure 5 plots exit rates in 1995 with respect to their size on the year of exit, as well as 1 and 3 years before they exit (note that these are *exit* and not *net exit* rates). The figure shows no evidence of the "Shadow of Death": establishments declining in size in the years leading up to their death. If it did, the curves in Figure 5 would show a downward shift as the year of exit approaches. Instead, the only difference between the curves is limited to quite small firms, and consists primarily of firms that survive less than 1 or 3 years. This suggests that selection may be important for small young establishments, but not for medium and large ones. In contrast, our theory predicts scale effects in net exit for establishments of all sizes.

Our theory emphasizes the role of size in establishment dynamics, but in common with many theories, it abstracts completely from age effects. In our theory young firms behave identically to old firms: size, but not age, matters. That is, our theory has been designed to capture the significant scale effects found in the empirical literature, without saying anything about age effects. Unfortunately, our dataset does not contain information on age and so we are not able to present results for given age cohorts. This would be a problem if the scale dependence that we document was all due to the fact that young firms are small and old firms are large. However, there are two reasons to believe that this is not the case. First, although some of the scale dependence we document may be the result of age effects, the preceding empirical literature on establishment dynamics has found that scale effects are important even after controlling for age (Evans (1987a,b), Hall (1987), Davis and Haltiwanger (1999)[12]). Moreover, the magnitude of the age dependence documented by David and Haltiwanger (1999) for the US is much smaller than the scale dependence we found in our data, which suggest that a large part of it comes from actual scale and not age dependence.[13] Second, age effects seem to diminish quite quickly with age. For example, Davis and Haltiwanger (1999) find that the 6 and 11 year-old cohorts behave very similarly. The scale dependence in growth rates that we document and use in the empirical exercises in the next section concern 10 year growth rates conditional on survival. Although this may not eliminate age effects completely, it should

[12] Davis and Haltiwanger (1999) study only unconditional growth rates.

[13] Using more aggregated data for enterprises with only one establishment one can verify that the predictions of our theory for sectoral variation in the thickness of the tails of the size distribution hold for firms younger and older than 5 years in the leading example used below.

reduce their importance. It is also important to note that there is no theory of age dependence that we know of that is consistent with the rich set of facts on scale dependence that we document, and in particular the sectoral variation in the amount of scale dependence in growth rates, size distributions, and net exit rates. Finally, it should be noted that our theory could easily be extended to include age effects. As it stands, it is silent about which establishment exit or enter when the number of establishment in an industry expands or contracts. Any theory of age effects could be used to determine the identity of these establishment and therefore add age effects to our theory, and indeed we view these theories as complementary to ours. Since our focus in this paper are scale and not age effects we leave this extension for future research.

4.1.4 Sectoral Capital Shares.—
In the next few sections, we turn to an examination of the implications of our model for cross-sectoral differences in mean reversion. For this purpose, we need data on physical capital shares which comes from the Bureau of Economic Analysis (BEA) Industry Accounts. We use data on labor costs and value added at basic prices to construct labor shares which include human capital. We then construct physical capital shares as one minus the labor share. This implies that the physical capital shares we use include everything that is not classified as labor. There are two potential problems with the physical capital shares we compute. First, the physical capital shares include land shares. Land is not an industry specific factor, but as its share is usually small, this should have a negligible effect on the physical capital shares we use. Second, we are using the physical capital share in value added, but our theory is abstracting from the use of intermediate inputs. To address the former, we focus upon industries with physical capital shares smaller than one half, although the result are similar if we consider all sectors. To address the latter, we also present results with physical capital shares adjusted for the share of value added and the share of materials purchased from the same industry.

4.2 Evidence on Sectoral Scale Dependence

On top of the economy-wide scale dependence documented in the previous section, our theory implies that scale dependence should be larger in sectors which use physical capital more intensively. This implication distinguishes our theory from other available theories that may also imply economy wide scale dependence. For example, theories that emphasize financial constraints in explaining scale dependence predict that scale dependence should be more pronounced in sectors in which establishments have less collateral. This plausibly corresponds to sectors in which the human capital share is relatively large and the physical capital shares relatively small, which is the opposite prediction of our theory. We now present evidence on sectoral variation in scale dependence for conditional growth rates, net exit rates, and size distributions and show that it corroborates the implications of our theory.

4.2.1 Growth Rates of Surviving Firms.—

We begin by examining the growth rates of surviving establishments. As a first step, consider an example with two sectors. Educational services is a very labor and human capital intensive sector with a physical capital share of 0.054, while manufacturing is much more physical capital intensive with a share of 0.397. If the theory is consistent with the data, given that manufacturing is more physical capital intensive, we should see growth rates of manufacturing establishments decline faster with size than growth rates of establishments in the educational sector (Proposition 1).

Figure 6 illustrates that this is the case, and shows that the differences are very large over a period of ten years. Not only do small establishments grow faster than large establishments in both sectors, but the scale dependence is significant for the entire range of establishment sizes. The difference between the growth rates in these two sectors increases with establishment size and is, for the largest establishments, more than 40 per-cent.

This evidence is not particular to the pair of sectors in the example. We examine next the implication of our theory that scale dependence in growth rates increases with physical capital shares (denoted by α_j) for all sectors. We use data on the growth of establishments, n_{t+1j}/n_{tj}, in a particular size category, n_{tj}, and estimate the regression specified by equation (13):

$$\ln\left(\frac{n_{t+1j}}{n_{tj}}\right) = \tilde{a}_j + \tilde{b}\ln n_{tj} + \tilde{e}\alpha_j \ln n_{tj} + \tilde{\varepsilon}_{tj},$$

where

$$\begin{aligned} \tilde{a}_j &= n_j^C, & \tilde{b} &= -\left(1-\omega_j\right)\left(1-\beta_j\right), \\ \tilde{e} &= -\left(1-\omega_j\right)\beta_j, & \text{and } \tilde{\varepsilon}_{tj} &= -\beta_j\left(1-\alpha_j\right)\ln A_{t+1j}. \end{aligned}$$

Notice that a full structural estimation of our model would require $\tilde{b}$ and $\tilde{e}$ to vary as β_j and ω_j vary by sector. Unfortunately, we do not have data on the share of specific human capital in labor services or the share of investments in human capital production. So we assume that these two shares do not vary across sectors or, if they do, that they are uncorrelated with capital shares. In the latter case our estimation strategy is not efficient, but the coefficients are still unbiased and consistent. Given that all the results presented below are significant at a 1% level the lack of efficiency of the estimator is not worrisome.[14] Our estimation procedure allows us to back out the average level of these parameters for the US (which we compute below). Apart from this caveat, our empirical exercise uses precisely the structure imposed by our model.

A requirement for a structural estimation of our model is to account for the heteroscedasticity that the model implies. In particular, the model predicts that the variance of the error term decreases with α_j, if we assume that the variance of the technology shocks is constant across sectors. We use generalized least squares (GLS) to take this effect into account and estimate the equation above with and without including the effect of capital shares on the variance.

Estimating the regression above amounts to fitting an exponential trend where the parameter varies linearly with physical capital shares by sector. We estimate this relationship using GLS to take into account the fact that there are many more establishments in the smaller size categories, as well as the heteroscedasticity predicted by our model. Of course, this should improve the efficiency of our estimation, but the results without taking the heteroscedasticity into account are still unbiased and consistent. We calculate the weights using data on the number of establishments in each size category and capital shares. Throughout this subsection we use average capital shares for the period 1990-2000. The theory predicts that the estimate of $\tilde{e}$ should be negative and significant.

The estimates of $\tilde{e}$ are presented in Table 1. The first and third columns present the

[14] If β_j and ω_j are correlated with α_j then the estimates of $\tilde{e}$ are biased. Let $\beta_j = \beta + \beta_j^e$ where β is the mean of β_j. Similarly, let $\omega_j = \omega + \omega_j^e$. Then the sign of the bias depends on the sign of the covariance between $\omega_j^e \beta + \beta_j^e (1-\omega) + 2\omega_j^e \beta_j^e$ and α_j. If $\omega_j^e = 0$ for all j then if $cov\left(\beta_j^e, \alpha_j\right) \geq 0$ our estimates are biased towards zero which reinforces our results. If $\beta_j^e = 0$ for all j then if $cov\left(\omega_j^e, \alpha_j\right) \geq 0$ again we get estimates biased towards zero which reinforces our results. In general in order for our estimates to be biased towards zero we need to assume

$$cov\left(\omega_j^e \beta + \beta_j^e (1-\omega) + 2\omega_j^e \beta_j^e, \alpha_j\right) \geq 0.$$

This restriction amounts to saying that human capital depreciates slower and amounts for a larger share in industries that are more capital intensive. This would be the case if, for example, industry specific human capital is needed to operate industry specific machines, and if those machines depreciate slower in industries that use more physical capital. We leave it as a restriction imposed on the empirical model. Note, however, that given the magnitude of our results this covariance would need to be very strongly negative to overturn them.

estimates using average capital shares for the period 1990-2000. The first two columns weight observations only by the number of establishments in a particular class bin. The last two adjust also for the heteroscedasticity predicted by our model. Although the estimates of $\tilde{e}$ are negative and strongly significant in all regressions, the results are the strongest when we use the exact specification given by our theory, namely, variance terms that depend on capital shares. Given the largest establishment size in our sample, a larger (in absolute value) coefficient implies more scale dependence for all establishment sizes. The results in Table 1 show that scale dependence increases significantly with sectoral physical capital shares: A doubling in the size of establishments in manufacturing ($\alpha_j \approx 1/3$) decreases average growth by about 5% while in educational services ($\alpha_j \approx 0$) the growth rate is roughly the same.

As mentioned before, the physical capital shares have been calculated as 1 minus the share of labor compensation in value added. Given that materials are an important fraction of gross output in an industry, this may result in physical capital shares that are too large relative to the ones in gross output. Since our theory does not include materials, it is not designed to address this distinction. To address these concerns we calculated the share of value added plus the share of inputs originating from the same sector using the input-output data provided by the BEA. We then multiply this share by the physical capital share to obtain an adjusted physical capital share. If all intermediate inputs originated in the same sector, the original physical capital shares would equal the adjusted physical capital shares. If the rest of the materials used in production are homogeneous, the adjusted physical capital shares would differ from the original shares, and the adjustment is theoretically exact. In general, even with this adjustment, we are abstracting from the effects of mean reversion in human capital stocks in other industries. However, one would expect the omission of these effects to bias our coefficients toward zero. Given the statistical significance of our results presented in columns two and four of Table 1, we believe that this does not undermine our empirical strategy.[15] The omission of intermediate inputs from other sectors may account for some of the unexplained variation in growth rates. Variation across sectors in other parameters of the model, such as the share of raw labor, the variance of productivity shocks, or the depreciation parameters, may account for some of the unexplained variation too.

Our estimation of $\tilde{b}$ and $\tilde{e}$ assumes that both β_j and ω_j are constant, or independent of α_j, across industries (call the average or constant values β and ω respectively). We

[15] Adjusting the physical capital shares increases the number of sectors in our sample with physical capital shares below one-half from 44 to 52.

can then use the estimates presented in Table 1, together with the estimates of $\tilde{b}$ and equation (13) to infer values for β and ω. The estimates of $\tilde{b}$ for the exercises in the first two columns of Table 1 are $-.146$ (s.e. .009) and $-.154$ (s.e. .008). When we use the capital share in the weights in the third and forth columns they become $-.134$ (s.e. .011) and $-.145$ (s.e. .008). These values imply a share of specific human capital in labor services between .432 and .556 (where $\beta = 1/(1+\tilde{b}/\tilde{e})$). That is, the model, and the estimation above, imply that the share of labor services related to specific human capital is roughly half. Thus, as we have argued in this paper, the share of specific human capital consistent with the scale dependence in establishment dynamics is very significant. Other forms of human capital that are not industry specific, and therefore are associated with individuals and not with an industry are, of course, not included in this share. These estimates also imply an average share of investments in human capital production ($1-\omega = \tilde{e}/\beta$) between .258 and .326. This share depends on the period over which we calculate growth rates, which in this case is 10 years.

The last ten years have witnessed a substantial decline in employment among large manufacturing establishments. A potential concern is that this phenomenon may be driving the larger scale dependence observed in capital intensive manufacturing sectors. To address this concern, we replicate the previous exercise for manufacturing and non-manufacturing sectors separately. The results presented in Table 2 show that this phenomenon is not driving the results in Table 1. The point estimates for non-manufacturing are close to the ones for the whole economy and strongly significant. For manufacturing the estimates are much less precise reflecting the smaller variation in physical capital shares among these sectors. This was precisely our original justification for using all sectors in the economy.

4.2.2 Net Exit Rates.—

Our mechanism, which emphasizes mean reversion in stocks of specific factors, when combined with an assumption on the level of the elasticity of substitution, also implies that net exit rates should decline with establishment size. Furthermore, the rate of decline should increase with physical capital shares. Figure 7 examines this prediction using BITS data for US manufacturing and educational services in 1995-1996. We plot net exit rates in manufacturing and educational services, together with their logarithmic trend, for establishments between 50 and 2000 employees. For these data the theory does well. Net exit rates decline clearly faster with size for manufacturing than for educational services, as is particularly clear in the trend lines. Overall, the logarithmic trend in manufacturing is steeper than in educational services.[16]

The results hold very strongly across all sectors in the economy. We run the following regression

$$\ln\left(1 + NER_j\right) = \breve{a}_j + \breve{b} \ln n_j + \breve{e} \alpha_j \ln n_j + \breve{\varepsilon}_{tj},$$

where NER_j denotes the net exit rate in sector j and size bin n_j. This amounts to estimating the exponential relationship between net exit rates and sizes implied by the model. The results are presented in Tables 3 for net exit rates in 1995-1996.[17] The top panel shows the results using GLS where the weights in the variance-covariance matrix includes only the number of firms.[18] The lower panel presents the same exercise when we adjust the weights to take into account the heteroscedasticity predicted by the theory. Again, we present all results using average capital shares for the period 1990-2000, along with results with adjusted capital shares. On top of this we also present the results if we measure size one year before/after exit/entry. All regressions include all firms size bins for which we have data (that is, in contrast to the graphs above the regression results include also establishment with less than 50 and more than 2000 employees).

The results are consistent with our theory: All of the estimates are negative and significant. The results are also economically significant: A doubling of establishment size decreases net exit rates by around 1% in manufacturing while net exit rates decline little with size in educational services. The results are still very significant but smaller in magnitude when we measure size one year before/after entry/exit. Finally, note that the theory implies that the estimates of $\breve{e}$ should be identical to the estimates of $\tilde{e}$ for conditional growth rates in Section 4.2.1, once we correct for the fact that the share of investments in human capital accumulation $(1 - \omega)$ should be smaller for one year (as in the case of net exit rates) than for 10 years (as in the case of growth rates). We can use one year growth rates in 1990 and net exit rates in 1993-1994 (using size

[16] Orr (1974), Gorecki (1976), Hause and Du Rietz (1984) and MacDonald (1986) found that firm exit rates were negatively related to measures of physical capital intensity by industry. Given that these studies do not distinguish among firms with different sizes, the negative relationship may be the result of the dependence predicted by our theory. This would be the case if firms in physical capital intensive sectors are larger on average.

[17] The results are very similar if we use net exit rates in 1993-1994.

[18] The measure of size we use is given by $\breve{\mu}_j = \left(\left(\left(\mu_{jy_1} + \mu_{jy_2}\right)/2\right) / \left(\mu_{jy_1} + E_{jy_1}\right)\right)^2$ where μ_{jy_2} is the number of establishments in year y_i of a given size indexed by j and E_{jy_1} is the number of establishments that entered in y_2 of a given size j. The reason we use this measure is that in contrast with the growth rate regressions we should not just use the number of surviving firms but the sum of all firms alive before exit and after entry.

also in 1993-1994) and estimate both equations jointly, restricting the parameters to be the same (the difference in years is due to data limitations). In this case we obtain a coefficient for $\tilde{e} = \breve{e}$ equal to -0.0507 (s.e. 0.006) or -0.0418 (s.e. 0.007) with adjusted capital shares. If we take into account the heteroskedasticity implied by the model we obtain -0.0496 (s.e. 0.007) and -0.0372 (s.e. 0.008) respectively. Hence, restricting the parameters to be the same yields results that are in line with the results in Table 3. This confirms yet another of the theories predictions: that the amount of scale dependence in both growth and net exit rates should be similar within a given sector.

4.2.3 Size Distributions.—

We now turn to the implication of our theory for the size distribution of establishments. As we have argued above, the variance of growth rates and net exit rates, and therefore establishment sizes, within a sector should be smaller the higher the capital share. In fact in some exercises presented above, we used this prediction to calculate efficient estimates of the variation of scale dependence with capital shares. It is natural therefore to look directly at the relationship between the standard deviation of establishment sizes and capital shares across sectors. Figure 8 plots the standard deviation of establishment sizes against adjusted capital shares for the years 1990 and 2000. We also plot linear trends for both years. As the theory predicts, the relationship between the standard deviation and adjusted capital shares is clearly negative, and is similar for both years.

In Section 2.4 above, we also examined an alternate measure of the amount of dispersion in establishment sizes based on the thinness of the tails of the size distribution of establishments. This measure has the advantage of being less sensitive to the size of the largest firms in the sample, which is particularly important for industries where the size distribution is well approximated by a Pareto distribution for which second moments are not defined. To compare thinness of tails, we use data from the SUSB to calculate the share of establishments in sector j with employment larger than n_j, which we denote by P_j. If the distribution of establishment sizes is Pareto with coefficient one, or growth rates are scale independent, the relationship between $\ln P_j$ and $\ln n_j$ should be linear with slope minus one. If growth rates depend negatively on scale, the tails of the distribution are thinner than the tails of a Pareto with coefficient one, and the relationship is concave. Our theory states that the degree of concavity should be positively related with physical capital shares (Proposition 7).

A first look at the data is presented in Figure 9 where we plot $\ln P_j$ and n_j for educational services and manufacturing. This representation of the size distribution emphasizes the degree of concavity and makes differences between the two distributions particularly clear for large establishment sizes. The differences between the distribution are also clear if we look at the density functions (with normalized means) plotted in Figure 10. It is clear that the distribution of establishment sizes in the educational sector has more mass for very small and large establishments, and less mass for intermediate establishments than in the manufacturing sector. This is particularly clear for small establishments in the graph. The figure also compares these distributions with the Pareto distribution with coefficient one (that corresponds to a straight line with slope -1 in Figure 9). The Pareto distribution with coefficient one has even more mass at the tails and less at the center, consistent with Proposition 7 as long as $\beta_j, \omega_j, (1-\alpha_j) > 0$. Both industries have thinner tails than the benchmark, but as the theory predicts, the difference is larger for the manufacturing sector. Moreover, the differences between these distributions are economically large: in order to transform the size distribution of the manufacturing sector to that of the educational services sector, around 20% of the labor force that currently works in medium sized manufacturing establishments would need to be reallocated to establishments with less than 50, or more than 1000, employees.

In order to test the relationship between physical capital shares and the size distribution of establishments *for all sectors*, we use our new data set on the size distributions of establishments for 1990 and 2000. We estimate the following regression

$$\ln P_j = \hat{a}_j + \hat{b}_j \ln n_j + \hat{d} \left(\ln n_j\right)^2 + \hat{e}\alpha_j \left(\ln n_j\right)^2 + \hat{\varepsilon}_j,$$

where $\hat{a}_j$ and $\hat{b}_j$ are industry specific coefficients. This amounts to constraining the quadratic term to vary linearly with the physical capital share. The model now predicts that $\hat{e}$ should be negative and significant. The results are presented in Table 4, which shows that the estimate of $\hat{e}$ for both 1990 and 2000, is negative and strongly significant. The physical capital shares used in this regression are the ones corresponding to 1990 and 2000 respectively. The results with adjusted physical capital shares are presented in the third and forth columns of Table 4, which further confirms the empirical significance of the mechanism in our theory.

5. CONCLUSION

In this paper we have constructed a theory that is consistent with some well known facts on scale dependence in establishment dynamics and establishment size distributions. The mechanism emphasizes the role of the accumulation of industry specific human capital, and we have shown that this mechanism is robust to institutional and economic differences across sectors and countries. We claim that the ubiquitous presence of these facts has to be the result of a mechanism, like ours, that is general enough to be present in a variety of circumstances. The central role of accumulation of industry specific human capital in the theory led us to focus on cross sectoral differences in the importance of human, and therefore physical, capital in production, and in particular physical capital intensity. Increases in the importance of industry specific physical capital lead to an increase in the degree of diminishing returns in human capital, and hence more scale dependence in growth, net exit rates, and establishment size distributions. Since it was the theory that guided our focus on this particular dimension of the data, the available evidence in the empirical literature is only indirect. Consequently, we take this prediction to the data and show that it is a surprisingly good description for the cross-section of US sectors. Any theory of establishment dynamics has to confront this new evidence.

In the introduction we commented on different studies that have emphasized financial frictions as well as other types of market inefficiencies as explanations for the observed scale dependence. What we show in this paper is that even though these frictions may be important for very small firms, they are not needed or necessarily consistent with the degree of scale dependence and its variation across sectors. This points to frictions that may be present for very small firms and that might be alleviated with particular policies. It is important, however, that these policies do not interfere with the growth and net exit of larger and existing establishments, which are processes well described by our efficient economy. Our results are, in general, not sensitive to government policies that affect establishments independently of their size. Scale dependent policies may affect some of our implications and Restuccia and Rogerson (2004) argue that scale dependent policies may have large effects on efficiency. International evidence on establishments dynamics and the size distribution of establishments, when combined with our benchmark, could shed some light on the empirical significance of scale dependent policies.

Our empirical findings should also be of interest to researchers interested in assessing differences in the efficiency of resource allocation across countries. For example, in recent work Guner and Ventura (2005) use differences in the size distribution of retail establishments between Japan and the US to argue for the presence of inefficiencies in the resource allocation mechanism. Similarly, other authors (see, for example, the survey by Tybout (2000) and the empirical evidence presented by Sleuwaegen and Goedhuys (2002)) have concluded that the thicker tails of the size distribution of establishments in sub-Saharan African countries, relative to the US, is evidence of corruption in these countries as establishments either stay small, to avoid official notice, or grow until they are large enough to co-opt the system to their own benefit. Our empirical finding, that the size distributions of industries with lower physical capital intensities display thicker tails, may strengthen this conclusion, given the greater concentration of the US economy in human capital intensive sectors.

By emphasizing the accumulation of specific human capital, our theory also makes predictions for the future evolution of the establishment size distribution. The ongoing specialization of developed economies in services will have important consequences on establishment sizes and establishment dynamics. Our theory predicts that this will lead to a more dispersed distribution of establishment sizes, where we will see more small and more very large establishments. These arguments suggest that we are moving towards an economy in which the dominance of large establishments in some industries, like Walmart, will coexist increasingly with large numbers of small establishments in different industries within the same sector, like bakeries or tailors. This trend is the natural result of the efficient division of an industry's production among establishments.

REFERENCES

[1] Albuquerque, R. and H. Hopenhayn (2002) "Optimal Lending Contracts and Firm Dynamics." University of Rochester - Center for Economic Research Working Paper 493.

[2] Audretsch, D.B., L. Klomp, and A.R. Thurik (2002) "Gibrat's Law: Are the Services Different?" ERIM Report Series 2002-04.

[3] Axtell, R. L. (2001) "Zipf distribution of US Firm sizes", *Science*, 293, 1818-1820.

[4] Cabral, L. and J. Mata. (2003). "On the Evolution of the Firm Size Distribution: Facts and Theory." *American Economic Review*, 93(4): 1075-1090.

[5] Aw, B., S. Chung, and M. Roberts. (2003) "Productivity, Output and Failure: A Comparison of Taiwanese and Korean Manufacturers." working paper, Penn State University.

[6] Caves, R. E. (1998). "Industrial Organization: New Findings on the Turnover and Mobility of Firms." *Journal of Economic Literature*, 36: 1947-1983.

[7] Clementi, G. and H. Hopenhayn (2002) "A Theory of Financing Constraints and Firm Dynamics." University of Rochester - Center for Economic Research Working Paper 492.

[8] Cooley, T. and V. Quadini (2001) "Financial Markets and Firm Dynamics", *American Economic Review*, 91(5):1286-1310.

[9] Davis, S. and J. Haltiwanger (1999). "Gross Job Flows." Handbook of Labor Economics. O. Ashenfelter and D. Card. Amsterdam, North Holland. 3:2711-2805.

[10] Dunne, T., M. Roberts, and L. Samuelson, (1989). "Plant Turnover and Gross Employment Flows in the U.S. Manufacturing Sector." *Journal of Labor Economics*, Vol. 7 (1) pp. 48-71.

[11] Dunne, T., M. Roberts, and L. Samuelson, (1989). "The Growth and Failure of U.S. Manufacturing Plants." *Quarterly Journal of Economics*, Vol. 104 (4) pp. 671-98.

[12] Ericson, R. and A. Pakes (1995). "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, 62(1): 53-82.

[13] Evans, D. S. (1987a). "Tests of Alternative Theories of Firm Growth." *Journal of Political Economy*, 95(4): 657-674.

[14] Evans, D. S. (1987b). "The Relationship Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries." *Journal of Industrial Economics*, 35(4): 567-81.

[15] Garicano, L. and E. Rossi-Hansberg (2004). "Inequality and the Organization of Knowledge." *American Economic Review P&P*, 94(2):197-2002.

[16] Geroski, P. (1995). "What do we Know About Entry?" International Journal of Industrial Organization, 13:421-440.

[17] Geroski, P. (1983). "The Empirical Analysis of Entry: A Survey." University of Southampton, mimeo.

[18] Geroski, P. and J. Schwalbach (1991). Entry and Market Contestability. Oxford, Blackwell.

[19] Gibrat, R. (1931). *Les inégalités économiques*, Libraire du Recueil Sirey.

[20] Gorecki, P. K. (1976). "The Determinants of Entry by Domestic and Foreign Enterprises in Canadian Manufacturing." *Review of Economics and Statistic*s, 58: 485-488.

[21] Griliches, Z. and H. Regev (1995). "Firm productivity in Israeli industry 1979-1988." *Journal of Econometrics*, 65: 175-203.

[22] Guner, N., G. Ventura, and X. Yi. (2005). "Macroeconomic Implications of Size-Dependent Policies." Unpublished paper, Pennsylvania State University.

[23] Hall, B. H. (1987). "The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector." *Journal of Industrial Economics*, 35(4): 583-606.

[24] Hause, J. C. and G. Du Rietz (1984). "Entry, Industry Growth, and the Microdynamics of Industry Supply." *Journal of Political Econom*y, 92: 733-757.

[25] Hopenhayn, H. A. (1992). "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, 60(2): 1127-1150.

[26] Jones, C.I. (2004). "The Shape of Production Functions and the Direction of Technical Change," NBER working paper no. 10547.

[27] Jovanovic, B. (1982). "Selection and the Evolution of Industry." *Econometrica*, 50(3): 649-670.

[28] Kambourov, G. and I. Manovskii (2002). "Occupational Specificity of Human Capita." Unpublished paper, University of Pennsylvania.

[29] Klette T. and S. Kortum (2004) "Innovating Firms and Aggregate Innovation." *Journal of Political Economy*, 112(5):986-1018.

[30] Luttmer, E. (2004). "The Size Distribution of Firms in an Economy with Fixed and Entry Costs." Working Paper 633, Federal Reserve Bank of Minneapolis.

[31] MacDonald, J. M. (1986). "Entry and Exit on the Competitive Fringe." *Southern Economic Journal,* 52: 640-658.

[32] Mansfield, E. (1962). "Entry, Gibrat's Law, Innovation, and the Growth of Firms." *American Economic Review*, 52(5): 1023-51.

[33] Orr, D. (1974). "The Determinants of Entry: A Study of the Canadian Manufacturing Industries." *Review of Economics and Statistics*, 56: 58-66.

[34] Restuccia, D. and R. Rogerson (2004). "Policy Distortions and Aggregate Productivity with Heterogeneous Plants." Unpublished paper, Arizona State University.

[35] Rossi-Hansberg, E. and M. L. J. Wright (2004). "Urban Structure and Growth." Unpublished paper, Stanford University.

[36] Scherer, F. M. (1980). *Industrial Market Structure and Economic Performance*. Boston, Houghton Mifflin.

[37] Schmalensee, R. (1989). "Inter-Industry Studies of Structure and Performance." *Handbook of Industrial Organizatio*n. R. Schmalensee and R. Willig. Amsterdam, North Holland. 2: 951-1009.

[38] Sleuwaegen, L. and M. Goedhuys (2002), "Growth of Firms in Developing Countries, Evidence from Côte d'Ivoire," *Journal of Development Economics*, 68: 117-135.

[39] Sutton, J. (1997). "Gibrat's Legacy." *Journal of Economic Literatur*e, 35(1): 40-59.

[40] Tybout, J. R. (2000). "Manufacturing Firms in Developing Countries: How Well do They do, and Why?" *Journal of Economic Literature*, 38 (1): 11-44

[41] van Ark, B. and E. Monnikhof (1996). "Size Distribution of Output and Employment: A Data Set for Manufacturing Industries in Five OECD Countries, 1960s-1990." OECD Economics Department Working Paper 166.

APPENDIX: PROOFS

Proof of Proposition 3. See Rossi-Hansberg and Wright (2004) Proposition 4. ■

Proof of Proposition 4. As noted in the text, if we detrend the growth rate of surviving establishments, the invariant distribution of representative or average establishment sizes, in logs, is normal with mean M_j and variance as in (14). If we let $y = \ln n$, the distribution of actual establishment sizes has to be weighted by N_j/e^y. If we normalize it to be a probability distribution, then clearly it has to be proportional to

$$\frac{1}{S_j\sqrt{2\pi}}\frac{N_j}{e^y}e^{-\frac{\left(y-M_j\right)^2}{2S_j^2}}.$$

To work out the proportionality, we require the proportion k to satisfy

$$\frac{1}{S_j\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{kN_j}{e^y}e^{-\frac{\left(y-M_j\right)^2}{2S_j^2}}dy = 1.$$

Rearranging the left hand side and dropping the subscript j for convenience, we get

$$\begin{aligned} kN\frac{1}{S\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-y}e^{-\frac{(y-M)^2}{2S^2}}dy &= e^{\frac{S^2}{2}-M}kN\frac{1}{S\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{\left(y-\left(M-S^2\right)\right)^2}{2S^2}}dy \\ &= e^{\frac{S^2}{2}-M}kN \end{aligned}$$

where the last line follows from the fact that the integrand is that of a normal pdf. Hence

$$k = \frac{e^{-\frac{S^2}{2}+M}}{N},$$

and the pdf of the size distribution is

$$e^{-\frac{S^2}{2}+M}\frac{1}{S\sqrt{2\pi}}e^{-y}e^{-\frac{(y-M)^2}{2S^2}} = \frac{1}{S\sqrt{2\pi}}e^{-\frac{\left(y-\left(M-S^2\right)\right)^2}{2S^2}}.$$

That is, the actual size distribution of establishments is also lognormal with mean

$$e^{M-\frac{S^2}{2}}$$

and variance

$$e^{2M_j+S_j^2}\left(e^{S_j^2}-1\right).$$

That the variance is decreasing in α_j follows from the fact that both the mean and variance of representative establishment sizes is decreasing in α_j. ■

Proof of Proposition 5. The proof is independent for each sector so we drop j from the notation. The size of a establishment at time $t+1$ is given by

$$\ln n_{t+1} = g\left(n_t, A_{t+1}\right) \equiv -\ln A_{t+1} + \left(1 - \left(1-\omega_j\right)\left(1-\beta_j\left(1-\alpha_j\right)\right)\right)\ln n_t,$$

where we have assumes that the population size is fixed (alternatively, we could work with variations from trend). This lies in the compact set LN defined above. Let μ be the probability measure over A. Then, the probability of a transition from a point n to a set S is given by

$$Q\left(n, S\right) = \mu\left(A : g\left(n, A\right) \in S\right).$$

For any function $f : LN \to \mathbb{R}$ define the operator T by

$$\left(Tf\right)\left(n\right) = \int_{LLN} f\left(n'\right) Q\left(n, dn'\right) = \int_{\underline{A}}^{\overline{A}} f\left(g(n, A)\right) d\mu\left(A\right).$$

Define also the operator T^*, that maps the probability of being in a set S next period given the current distribution, say λ, as

$$\left(T^*\lambda\right)\left(S\right) = \int_{LLN} Q\left(n, S\right)\lambda\left(dn\right).$$

Since the set LN is compact, we are able to use Theorem 12.12 in Stokey, Lucas and Prescott (1989) to prove that there exists a unique invariant distribution, if we can show that the transition probability function Q satisfies the Feller property, is monotone, and satisfies the mixing condition.

To see that it satisfies the Feller Property, note that the function g is continuous in $\ln n$, and $\ln A$. Since g is continuous and bounded, if f is continuous and bounded, $f\left(g(\cdot)\right)$ will be continuous and bounded and therefore so is Tf. Hence T maps the space of bounded continuous functions into itself, $T : C(\bar{S}) \to C(\bar{S})$. To see that it is monotone, we need to prove that if $f : LN \to \mathbb{R}$ is a non-decreasing function, then so is Tf. But this follows from the fact that the g is non-decreasing in n. Hence $f\left(g(n, A)\right)$ is non-decreasing in n and therefore so is Tf.

Finally, to show that it satisfies the mixing condition, we need to show that there exists $c \in LN$ and $\eta > 0$ such that

$$Q\left(\frac{-\ln \overline{A}\beta_j\left(1-\alpha_j\right)}{\left(1-\omega_j\right)\left(1-\beta_j\left(1-\alpha_j\right)\right)}, \left[c, \frac{-\ln \underline{A}\beta_j\left(1-\alpha_j\right)}{\left(1-\omega_j\right)\left(1-\beta_j\left(1-\alpha_j\right)\right)}\right]\right) \geq \eta,$$

and

$$Q\left(\frac{-\ln \underline{A}\beta_j\left(1-\alpha_j\right)}{\left(1-\omega_j\right)\left(1-\beta_j\left(1-\alpha_j\right)\right)}, \left[\frac{-\ln \overline{A}\beta_j\left(1-\alpha_j\right)}{\left(1-\omega_j\right)\left(1-\beta_j\left(1-\alpha_j\right)\right)}, c\right]\right) \geq \eta.$$

Let $c = 0$. As g is continuous and decreasing in A, there exists an A' such that for all $A \leq A'$, $g(n, A) > 0$. Let $\eta' = 1 - \mu(A')$. Similarly there exists an A'' such that for all $\varepsilon \leq A''$, $g(n, A) < 0$. Let $\eta'' = 1 - \mu(A'')$. Call the minimum of these probabilities η. Then $c = 0$ and η guarantee that the mixing condition holds. Theorem 12.12 in Stokey, Lucas and Prescott (1989) then guarantees that there exists a unique invariant distribution, and that the iterates of T^* converge weakly to that invariant distribution. ■

Proof of Proposition 7. The first claim is immediate form the discussion above. To see the second, for each α denote the unique invariant probability measure of establishment sizes (see Proposition 5) by $\lambda_\alpha : \mathcal{LN} \to [0, 1]$, where $\mathcal{LN}$ denotes the Borel $\sigma-$algebra associated with LN, with associated transition function Q_α and operator T^*_α. Since λ_α is an invariant distribution

$$\begin{aligned} \lambda_\alpha\left(\left[-\ln \overline{A}, \ln n\right]\right) &= (T^*_\alpha \lambda_\alpha)\left(\left[-\ln \overline{A}, \ln n\right]\right) = \int Q_\alpha\left(z, \left[-\ln \overline{A}, \ln n\right]\right) \lambda_\alpha(dz) \\ &= \int \mu\left(A : g_\alpha(z, A) \in \left[-\ln \overline{A}, \ln n\right]\right) \lambda_\alpha(dz), \end{aligned}$$

where $g_\alpha(z, A)$ denotes the log establishment size growth rate. We saw above that

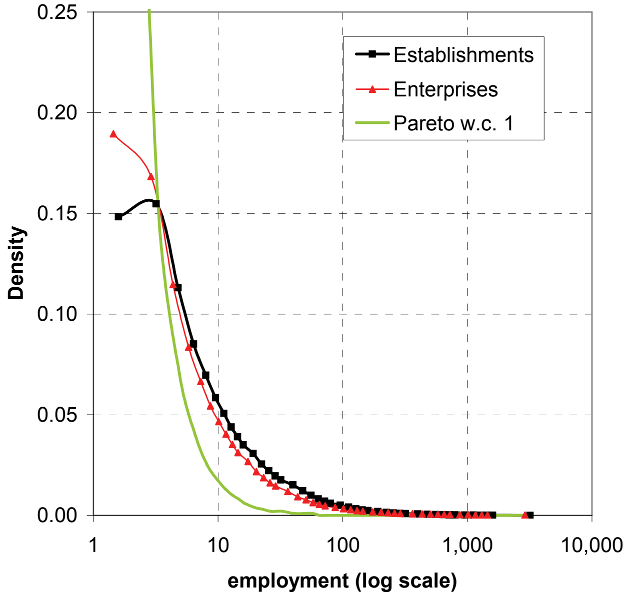
$$\frac{dg_\alpha(z, A)}{d\alpha} < 0.$$

Then, for n small enough, we know that

$$\begin{aligned} \lambda_{\alpha_k}\left(\left[-\ln \overline{A}, \ln n\right]\right) &= \int \mu\left(A : g_{\alpha_k}(z, A) \in \left[-\ln \overline{A}, \ln n\right]\right) \lambda_{\alpha_k}(dz), \\ &> \int \mu\left(A : g_{\alpha_j}(z, A) \in \left[-\ln \overline{A}, \ln n\right]\right) \lambda_{\alpha_k}(dz), \end{aligned}$$

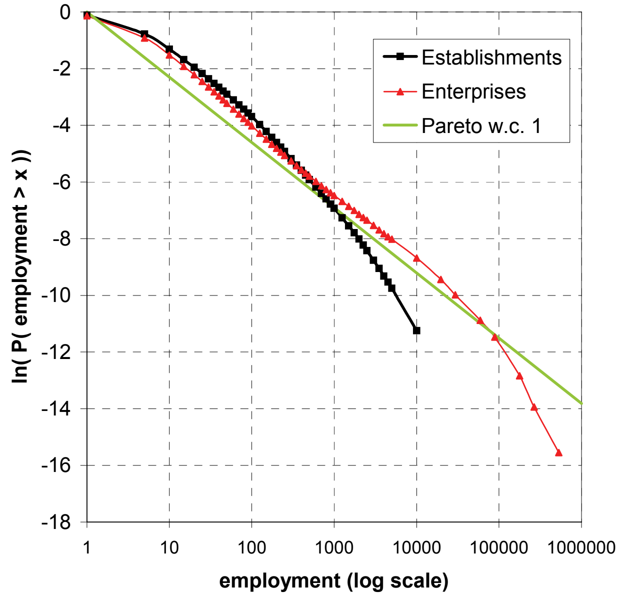
and hence λ_{α_k} is not the invariant distribution α_k, and the operator $T^*_{\alpha_j}$ maps the λ_{α_k} into distributions with thinner left tails. The case for intermediate and high $\ln n$ are analogous. ■

Figure 1: Density Function of Establishments and Enterprises in 2000



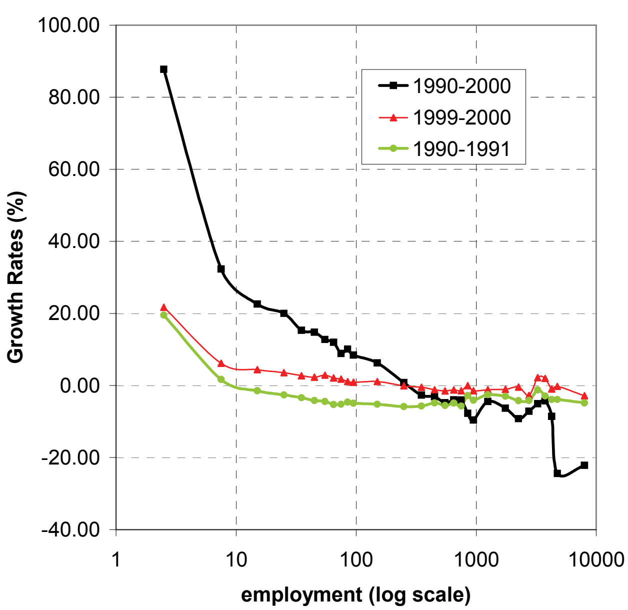
The figure presents the density of establishment and enterprise sizes in 2000 normalized so that the resulting distributions have the same mean. It also presents a pareto density with coefficient one. The data on the number of enterprises is aggregated in 50 bins and 43 bins for establishments. Source: US Census, Statistics of US Businesses.

Figure 2: Distribution of Establishments and Enterprises Sizes in 2000



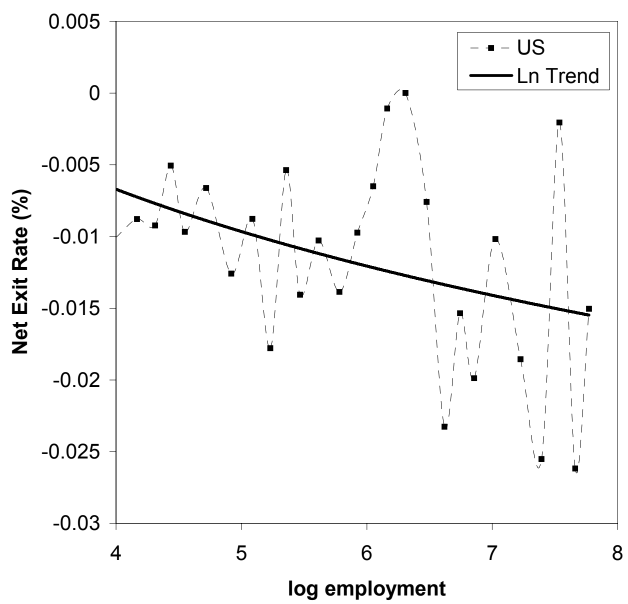
The figure presents the probability that establishments and enterprises are larger than a particular size against that size in 2000. The figure also presents the same probability for a pareto density with coefficient one. The data on enterprises is aggregated in 50 bins and in 43 bins for establishments. Source: US Census, Statistics of US Businesses.

Figure 3: Establishment Conditional Growth Rates, 1990-2000



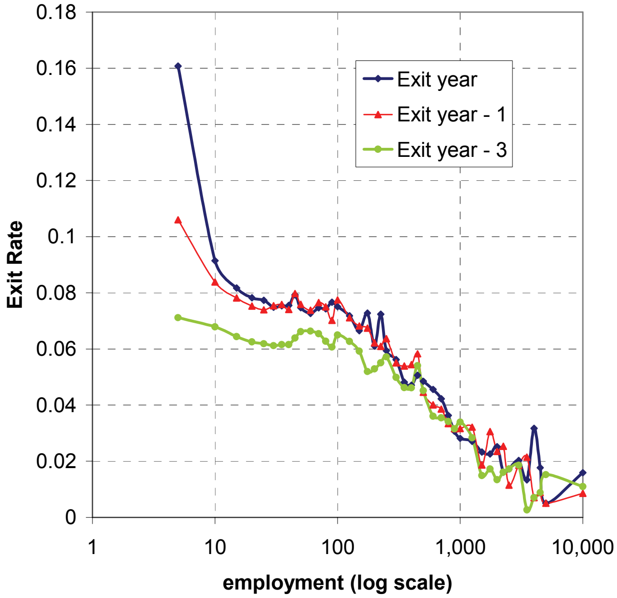
The figure presents average establishment employment rates by size bin of establishments that where alive between 1990 and 2000, 1999 and 2000, or 1990 and 1991 (respectively). Employment sizes are divided in 29 size bins. Source: US Census Bureau, Business Information Tracking System

Figure 4: Establishment Net Exit Rates, 1995-1996



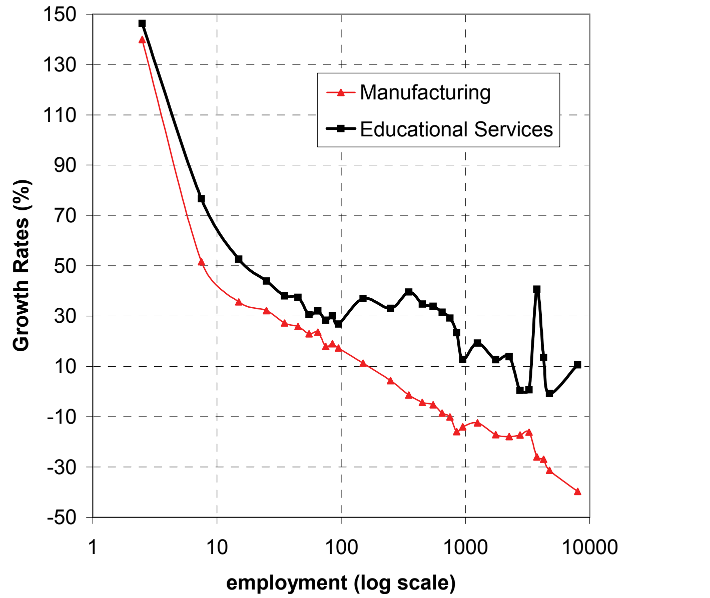
The figure presents the net exit rate (exit -entry rates) of establishments between 1995 and 1996 of firms with between 50 and 2500 employees by size. It also presents the series' log trend. The data is aggregated in 27 size bins. Source: US Census Bureau, Business Information Tracking System.

Figure 5: Establishment Exit Rates, 1995-1996



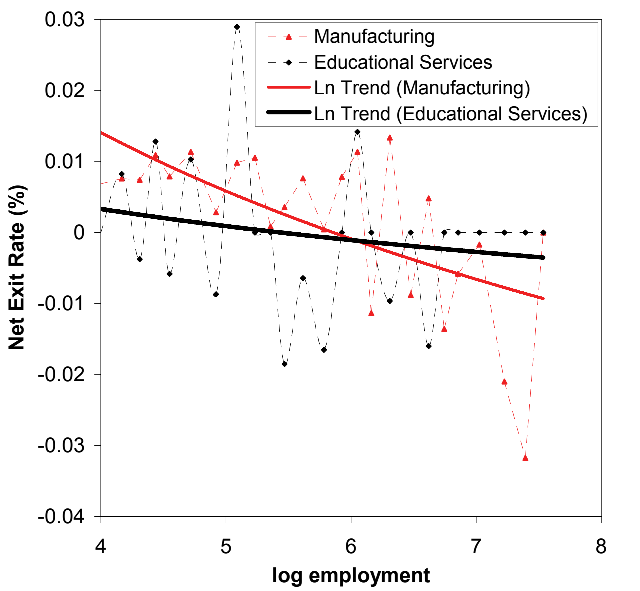
The figure presents the exit rate of establishments between 1995 and 1996 by size in the year of exit (1995) one year before exit (1994) and 3 years before exit (1992). The data is aggregated in 44 size bins. Source: US Census Bureau, Business Information Tracking System.

Figure 6: Establishment Conditional Growth Rates by Sector, 1990-2000



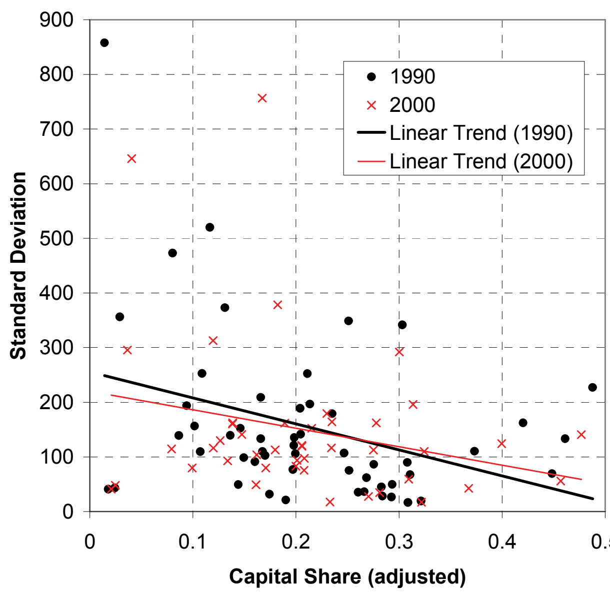
The figure presents average establishment employment rates by size bin of establishments that where alive between 1990 and 2000 in the educational services and manufacturing sectors. In both industries employment sizes are divided in 29 size bins. Source: US Census Bureau, Business Information Tracking System.

Figure 7: Establishment Net Exit Rate by Sector, 1995-1996



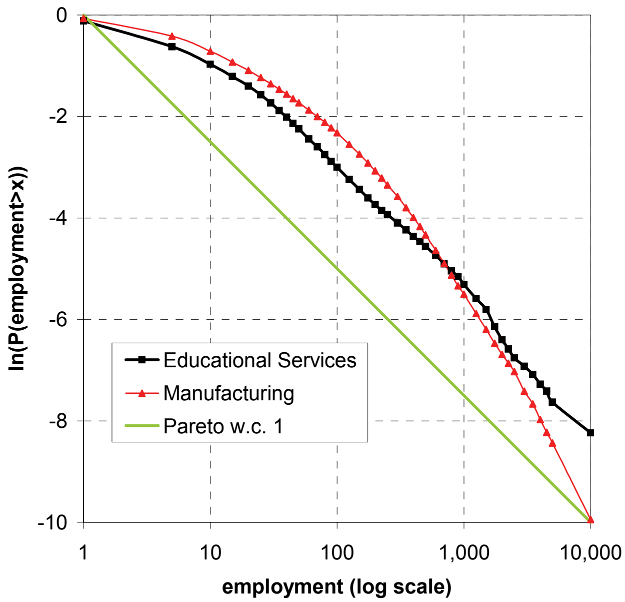
The figure presents the net exit rate (exit -entry rates) of establishments in the educational services and manufacturing sectors between 1995 and 1996 of firms with between 50 and 2000 employees by size. It also presentsave the log trends for both series. The data is aggregated in 24 size bins. Source: US Census Bureau, Business Information Tracking System.

Figure 8: Standard Deviation of Establishment Sizes and Capital Shares, 1990 and 2000



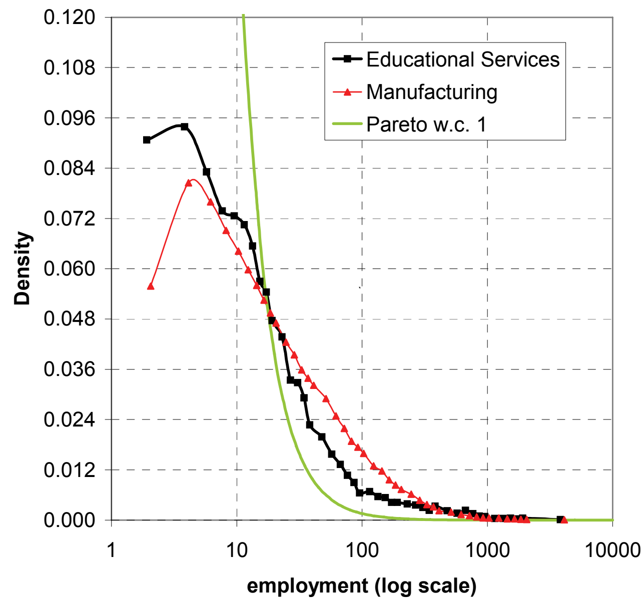
The figure presents the variance establishment sizes by sector and the corresponding capital shares for 1990 and 2000. The variance is . computed from data aggregated in size bins. Source: US Census Bureau, Statistics of United States Businesses.

Figure 9: Distribution of Establishment Sizes by Sector, 2000



The figure presents the probability that establishment in the educational services and manufacturing sector are larger than a particular size against that size in 2000. It also presents the same probability for a pareto density with coefficient one.The data on the number of establishments is aggregated in 43 bins. Source: US Census Bureau, Statistics of United States Businesses

Figure 10: Density Function of Establishments by Sector, 2000



The figure presents the density of establishmen t in the educational services and manufacturing sectors in 2000 normalized so that the resulting distributions have the same mean. It also presents a pareto density with coefficient one. The data on the number of enterprises is aggregated in 43 bins. Source: US Census Bureau, Statistics of United States Businesses

Table 1

	1990-2000			
	Variance $= 1/\mu_j$		Variance $= (1-\alpha_j)^2/\mu_j$	
		(adjusted)		(adjusted)
$\tilde{e}$	-0.1115	-0.1517	-0.1488	-0.1814
Standard error	0.0255	0.0314	0.0304	0.0325
P-value	0.0000	0.0000	0.0000	0.0000
# of Observations	940	1082	1137	1137

Table 2

	1990-2000			
	Variance $= 1/\mu_j$			
	Manufacturing		Non-Manufacturing	
		(adjusted)		(adjusted)
$\tilde{e}$	-0.0524	-0.0485	-0.1159	-0.1619
Standard error	0.0981	0.1213	0.0265	0.0329
P-value	0.5930	0.6900	0.0000	0.0000
	Variance $= (1-\alpha_j)^2/\mu_j$			
	Manufacturing		Non-Manufacturing	
		(adjusted)		(adjusted)
$\tilde{e}$	-0.0876	-0.0720	-0.1556	-0.1922
Standard error	0.0972	0.1295	0.0322	0.0342
P-value	0.3680	0.578	0.0000	0.0000
# of Observations	388	434	552	648

Table 3

Net Exit Rate 1995-1996, $\left(\frac{Exit96 - Entry95}{(\#\ Establishments96 + \#\ Establishments95)/2}\right)$

	Variance $= 1/\breve{\mu}_j$			
	Size in 1995-1996		Size in 1994-1997	
		(adjusted)		(adjusted)
$\breve{e}$	-0.0314	-0.0331	-0.0172	-0.0186
Standard error	0.0029	0.0034	0.0024	0.0028
P-value	0.0000	0.0000	0.0000	0.0000
	Weights $= (1-\alpha_j)^2/\breve{\mu}_j$			
	Size in 1995-1996		Size in 1994-1997	
		(adjusted)		(adjusted)
$\breve{e}$	-0.0324	-0.0280	-0.0164	-0.0151
Standard error	0.0036	0.0036	0.0029	0.0030
P-value	0.0000	0.0000	0.0000	0.0000
# of Observations	1733	2029	1682	1966

Table 4

	1990		2000	
		(adjusted)		(adjusted)
$\hat{e}$	-0.1015	-0.0402	-0.0730	-0.1309
Standard error	0.0152	0.0145	0.0167	0.0163
P-value	0.0000	0.0060	0.0000	0.0000
# of Observations	1864	2182	1486	1799