THE NATURE OF FIRM GROWTH*

Benjamin W. Pugsley
Federal Reserve Bank of New York

Petr Sedláček
University of Bonn

Vincent Sterk
University College London

February 2017

Abstract

There are vast differences in the growth patterns of firms: high-growth, young businesses, or "gazelles", account for the vast majority of employment growth at incumbent firms. Based on a large administrative panel data set for the United States, this paper shows that most of the size heterogeneity of firms, at a given age, is driven by ex-ante differences rather than ex-post shocks. We reach this conclusion after documenting the autocovariance structure of firm-level employment and estimating a reduced-form process that captures this structure. Next, we explore macroeconomic implications by matching a firm dynamics model to the empirical evidence. We show that, due to strong prevalence of ex-ante heterogeneity, firm selection creates sizeable gains in aggregate productivity. Nearly all of these gains derive from selection that takes place at the very beginning of firms' life cycles.

*VERY PRELIMINARY.** We thank Richard Blundell, Christian Bayer, Urban Jermann, Greg Kaplan, Fabien Postel-Vinay and Kjetil Storesletten for helpful comments.

1 Introduction

There are vast differences in the growth patterns of firms: high-growth, young businesses, sometimes called “gazelles”, account for the vast majority of employment growth at incumbent firms (see e.g. Haltiwanger *et al.* (2014)). One potential explanation for these growth differences is that, following entry, firms are exposed to idiosyncratic shocks, to for example productivity or demand. According to this view, a firm outgrows its peers when it is hit by relatively benign shocks during its lifetime. An alternative view is that there are ex-ante differences between firm startups, with some types poised for growth and others destined to stay small. Under this view, heterogeneity in firms' growth paths are predictable, given their initial characteristics.

While it seems plausible that both views on firm growth are –to some extent– grounded in reality, little is known about their relative empirical relevance. However, the nature of growth differences may have important consequences for aggregate outcomes. For example, if there are large ex-ante differences in the growth potential of firms, then the process which selects those aspiring startups with the most potential to become actual producers, may have a large positive effect on aggregate productivity. By contrast, if a firms' growth paths are mostly determined by post-entry shocks, then the gains from selection may be much smaller.

In this paper, we present direct empirical evidence on the importance of a deterministic component in firms' growth patterns, vis-à-vis post-entry shocks. We then use this evidence to discipline a firm dynamics model, designed to quantify the impact of firm selection on aggregate productivity. We find that the a large fraction of the differences in firm size, conditional on age, can be attributed to ex-ante heterogeneity, ranging from 85 percent in the year of startup and 47 percent at old age. At the macro level, we find that firm selection increases aggregate productivity by nearly one quarter and that the bulk of this increase is driven by selection that takes at the very beginning of firms' life cycles, before they may have even started to produce.

The key piece of empirical evidence we present is the autocovariance matrix of employment at the firm (and establishment) level, up to 19 years after startup. We estimate this matrix from the Longitudinal Business Dynamics (LBD) database, which contains administrative information on the population of employers in the United States. In order to summarize the information contained in the autocovariance matrix, we propose a reduced-form employment process. Based on the estimated process, we then quantify the importance of ex-ante heterogeneity versus ex-post shocks. Moreover, we use the reduced-form process as a guideline for the type idiosyncratic shock process to be integrated into the structural model.

The proposed reduced-form process allows for heterogeneity in both initial and long-run "steady-state" employment levels, as well as heterogeneity in the speed at which this transition takes place. In addition, it allows for post-entry shocks. Moreover, the process nests various specifications that are commonly used in the firm dynamics literature to model the idiosyncratic shock process that firms are exposed to. However, we find that standard processes do not capture very well the autocovariance structure that we observe. For example, our process nests a simplified case, found in for example Melitz (2003) among many studies, in which there are no ex-post shocks and all heterogeneity is modelled as a firm-level that is drawn ex-ante and remains constant over the life cycle, so that firms immediately reach their steady-state employment levels. This implies a flat autocovariance function, regardless of the age and horizon. In the data, however, autocovariances decline with the horizon, implying some role for ex-post shocks. Moreover, they increase with age, indicating a role for transitional dynamics. Another popular specification for the shock process is an AR(1) with a homogeneous constant, as found in for example Hopenhayn and Rogerson (1993). Such a process implies that firms gravitate towards the same steady-state levels and hence that ex-ante heterogeneity ultimately dies out. Thus, the implied autocovariances decline towards zero as the horizon is increased. In the data, however, autocovariances appear to stabilize at positive levels at longer horizons, suggesting an important role for heterogeneity in steady-state levels.[1]

[1] Our process also nests specifications with heterogeneity in the constant, as commonly allowed for in the econometrics literature on panel data models. However, our process is still more general, since we allow different components of the process to have different persistence parameters. Allowing for the the latter turns out to be important to fit the autocovariance matrix well and in Section XXX we

Not surprisingly, the generalization of the process turns out to be critical when quantifying the importance of ex-ante versus ex-post heterogeneity. A practical disadvantage of our process, however, is that it contains various state variables which may create computational difficulties when integrating the process into a structural model. The most general process introduces five exogenous state variables, versus only one in either Hopenhayn and Rogerson (1993) or Melitz (2003). To address this issue, we present restrictions that reduce the number of state variables from five to two, while preserving most of the dramatic improvement in fit, relative to the more standard processes.

Our next step is to explore implications of the empirical results for the macroeconomy using a structural firm dynamics model. While the nature of the idiosyncratic process is likely to matter in many applications of firm dynamics models, we focus on the effects of firm selection on aggregate productivity. Many studies, including Jovanovic (1982), Hopenhayn (1992), Hopenhayn and Rogerson (1993), Melitz (2003), Foster *et al.* (2008) have studied the selection of firms and have emphasized the importance of selection in the determination of aggregate outcomes. Our contribution is to demonstrate, qualitatively and quantitatively, the importance of the nature of the firm-level growth process in this dimension. The model we use for this purpose is an extension of the popular framework of Melitz (2003), but with an enriched idiosyncratic process which is flexible enough for the model to fit the empirical autocovariance structure of employment with reasonable accuracy.

Before conducting any quantitative exercises, we present two simplified cases which illustrates why the nature of the firm growth process is critical determinant of the strength of the selection channel. First, we consider a case in which all heterogeneity is permanent and determined ex ante, as in Melitz (2003). We show that the effect of selection on aggregate productivity depends positively on the amount of (ex-ante) heterogeneity. Second, we illustrate a polar case in there is no ex-ante heterogeneity and all shocks are drawn ex-post and are purely transitory. In this example, selec-

show that, interestingly, standard panel data estimators may produce quite misleading results in our application.

tion effects are small (or even completely absent). Combining the two cases, only the ex-ante heterogeneity matters for the selection effects. Intuitively, when differences between startups are large and permanent, there are large productivity gains to be made from selecting the best startups. By contrast, the effect of transitory shocks of firm productivity is only short-lived and therefore have a very limited effect effects on a firm's expected value. Hence, firm's exit decision, which are based on expected firm values, are not much affected and selection effects are therefore small.

Finally, we take the model to the data. Having integrated an idiosyncratic shock process of the type proposed earlier, the model can provide a good fit of the observed autocovariance matrix of employment, as well as the profiles of average size and exit by age. We then use the model to quantify the effect of selection on aggregate productivity, by comparing the model to a counterfactual version in which selection effects are shut off. We find sizeable productivity gains from selection, in the order of 20 percent in the aggregate. This gain is almost entirely driven by selection in the very first period, at a point at which firms have observed their ex-ante parameters but have not actually started to produce. Interestingly, this is true even though a substantial amount of endogenous exit takes place in subsequent years. The reason why this subsequent exit has relatively limited effects on aggregate productivity, is that these firms tend to be close to indifferent between exit and continuation, whereas many of the startups who exit immediately are on average further away from the indifference point.

Relation to the literature. The importance of ex-ante conditions has been highlighted by Hurst and Pugsley (2011) who present survey evidence that many nascent entrepreneurs do not expect their business to grow large. Barseghyan and Dicecio (2011b) present evidence that firm growth is partly predictable based on observable characteristics at the time of startup, e.g. whether or not the company is named after its owner or whether it is a corporation or limited liability company. Abbring and Campbell (2005) estimate an industry model with both transitory and persistent shocks, using data on 305 bars in Texas. They find that ex-ante decision account for about 40 percent of the variation in ex-post outcomes. Sedláček and Sterk (2016) document the presence of strong cohort effects in employment data and estimate a firm dynamics model with ex-ante demand heterogeneity and aggregate shocks. They find that much of the differences across cohorts born can be attributed to the state of the economy in the year of startup, suggesting that cohorts differ in their composition with respect to ex-ante characteristics.[2] In the present paper, by contrast, we quantify the importance of ex-ante heterogeneity directly by exploiting within-cohort variation.

Our reduced-form analysis is inspired by a large empirical literature on earnings dynamics of workers, which traditionally derives identification from the autocovariance structure of earnings, see e.g. MaCurdy (1982), Abowd and Card (1989). A common assumption in this literature that earnings are the sum of an individual fixed effect, an age fixed effect, an AR(1) process with zero mean, and an i.i.d. shock. Some authors, however, have argued that allowing in addition for individual-specific trends helps to capture the autocovariance structure of earnings, see Guvenen (2009) for a discussion of this branch of the literature. The possible presence of such "Heterogeneous Income Profiles" (Guvenen (2007)), has received much attention since they may have large implications for the extent to which income changes should be expected to transmit to consumption, from the perspective of standard life-cycle models (see e.g. Guvenen and Smith (2014)). Somewhat surprisingly, the literature on firm-level employment dynamics does not have a similar tradition of estimating reduced-form processes. To the best of our knowledge, even the basic autocovariance structure of employment dynamics has not been systematically documented.

Our structural model builds on a large literature which uses firm dynamics models to understand the determinants of aggregate productivity. Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) quantify the effects of frictions that reduce aggregate productivity by creating misallocation of resources, but abstract from selection effects. Bartelsman *et al.* (2009) consider a framework but allow for both misallocation and selection effects. Importantly, they discipline their model using the observed covariance between firm size and productivity. Barseghyan and Dicecio (2011a) use present a

[2] The importance of the composition of the firm population is also emphasized by **?**, who document a strong trend in the U.S. towards older firms, which is the result of accumulating startup deficits.

model to quantify the aggregate effects of variations in entry costs observed across countries,but abstract from post-entry shocks and restrict ex-ante heterogeneity to be constant, as in Melitz (2003). Our analysis complements these studies by highlighting the importance of matching the observed autocovariance matrix of employment when quantifying selection effects. To the best of our knowledge, our study is the first to quantify the effect of selection on aggregate productivity in isolation, by making the comparison to a counterfactual economy without selection.

2 Reduced-form estimations

2.1 Data description

We use data on firm-level employment in the United States, take from the from Census Longitudinal Business Database. [*to be completed*]

We construct a balanced panel of employment in firms surviving up to age 15. Prior to the analysis, we take out a fixed effect for the birth year of the firm and the firm's 4 digit industry.

2.2 The autocovariance structure of firm-level employment

Figure 1 presents basic cross-sectional moments of logged employment by firm age.[3] The upper left panel shows that, at younger ages, average log employment is an increasing of age, growing from an average of 0.25 (1.30 in levels) in the year of startup to 4.1 at age fifteen (59.1 in levels). After age ten, there is relatively little average size growth. The upper right panel present shows the standard deviation by age. During the first 10 years, employment levels fan out. After this age, the cross-sectional standard deviation stabilizes as a function of age.

Observed jointly the upper panels reveal two important patterns. First, differences between firms of the same age grow larger during the first ten years of their lifetime.

[3] We present moments for logged employment rather than levels, since we will use logged employment in the estimations.

Second, by age 10, the cross-sectional employment distribution appears close to a stationary end point.

The lower left panels of Figure 1 present the autocorrelation functions by age. Two important patterns stand out. First, autocorrelations are high, and decay only slowly as the lag length is increased. As a result, differences in employment across firms are very persistent, even at very long lags. Second, for a given lag length, the degree of persistence increases with firm age. For example, the correlation of logged employment at age 0 and 10 is 0.42, whereas the correlation between age 5 and age 15 is 0.88. Finally, the lower right panel presents the autocovariance function by age. Although the information in this panel is already summarized by the standard deviation an autocorrelation, we present it for the sake of completeness and because we will use the autocovariance function directly in the estimations to be presented below (XXX update figure).

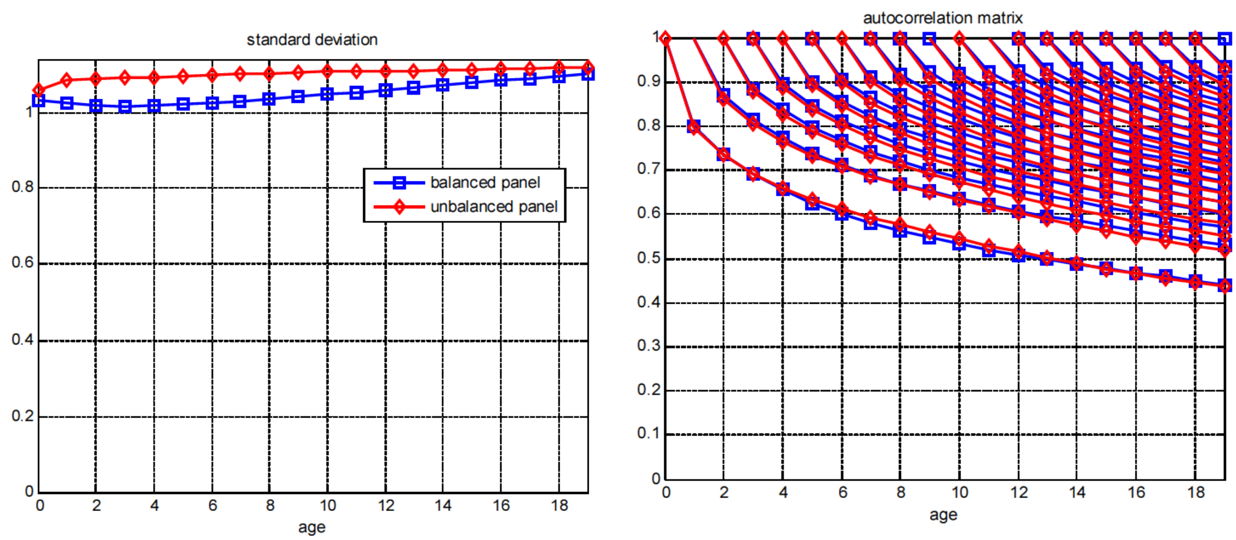


Figure 1

2.3 Employment process

We now estimate a reduced-form process for firm employment. Let $n_{i,a}$ be the employment level of an individual firm of age a. We consider the following process for logged employment:

$$\ln n_{i,0} = \rho \ln \widetilde{n}_i + \theta_i, \tag{1}$$

$$\ln n_{i,a} = \rho \ln n_{i,a-1} + \theta_i + \varepsilon_{i,a}, \text{ for } a = 1, 2, 3, ... \tag{2}$$

Here, $\rho \in (-1, 1)$ is a persistence parameter which is common across firms, and $\ln \widetilde{n}_i$ is an i.i.d. shock, drawn before entry from a distribution with mean $\mu_{\widetilde{n}}$ and standard deviation $\sigma_{\widetilde{n}}$. Further, θ_i is a firm-specific parameter, drawn before entry, from a distribution from an i.i.d. distribution with mean μ_θ, and standard deviation σ_θ. Finally, $\varepsilon_{i,a}$ is a post-entry shock which is drawn at age a from an i.i.d. distribution with mean zero and standard deviation σ_ε.

Given the above process, one can decompose the firm's logged employment level at age a into three components:

$$\ln n_{a,i} = \rho^{a+1} \ln \widetilde{n}_i + \sum_{k=0}^{a} \rho^k \theta_i + \sum_{k=1}^{a} \rho^{a-k} \varepsilon_k \tag{3}$$

The first term on the right-hand side of the above equation captures the effect of the pre-entry shock $\ln \widetilde{n}_i$. Given that ρ is smaller than one in absolute value, this component decays to zero as the firm ages. The second component captures the effect of the pre-entry parameter θ_i and does not die out with age. Letting age approach infinity, the second component converges to $\frac{\theta_i}{1-\rho}$. The third component captures the effect of the post-entry shocks, and has an unconditional mean of zero. Thus, $\frac{\theta_i}{1-\rho}$ is the steady-state level of employment of firm i, i.e. the mean of its old-age stationary employment distribution. It now follows that when $\sigma_\theta = 0$, there is no heterogeneity in steady state level across firms. That is, under this assumption all firms would fluctuate around the same employment levels at old age. Hence, σ_θ is a key parameter of interest.

Equation (3) directly leads to the following expression for the cross-sectional variance of logged employment, conditional on age:

$$Var(\ln n_a) = \left(\rho^{a+1}\right)^2 \sigma_{\widetilde{n}}^2 + \left(\sum\nolimits_{k=0}^{a} \rho^k\right)^2 \sigma_\theta^2 + \sum\nolimits_{k=1}^{a} \rho^{2(a-k)} \sigma_\varepsilon^2 \tag{4}$$

Note that the expression is a function of parameters only and that again it is the sum of the three components mentioned above. Thus, once the parameters of the process are known it is straight-forward to decompose the cross-sectional employment variance of firms at a certain age into the contributions of the pre-entry transitory draw, the steady-state parameter (also drawn pre-entry), and the post-entry shocks. Letting age approach infinity, we arrive at a particularly simple expression for the cross-sectional employment variance, which is a function of only three parameters, ρ, σ_θ^2 and σ_ε^2, as the transitory pre-entry component no longer plays a role:

$$Var(\ln n_\infty) = \frac{\sigma_\theta^2}{(1-\rho)^2} + \frac{\sigma_\varepsilon^2}{1-\rho^2}. \tag{5}$$

2.4 Parameter identification

We now show that key parameters ρ, σ_θ and σ_ε and $\sigma_{\tilde{n}}$ can be identified from the autocovariance matrix of logged employment. In the appendix, we derive the following moment conditions:

$$Cov\left(\ln n_{a,i}, \ln n_{a-j,i}\right) = \gamma_{j-1}\gamma_{a-j}\sigma_\theta^2 + \rho^j Var\left(\ln n_{a-j,i}\right), \tag{6}$$

$$Var\left(\ln n_{a,i}\right) = \rho^{2j} Var\left(\ln n_{a-j,i}\right) + \gamma_{j-1}^2\sigma_\theta^2 + \phi_{j-1}\sigma_\varepsilon^2 + 2\rho^j\gamma_{j-1}\gamma_{a-j}\sigma_\theta^2, \tag{7}$$

where $\gamma_j \equiv \sum_{k=0}^{j}\rho^k$ and $\phi_j \equiv \sum_{k=0}^{j}\rho^{2k}$. It now directly follows that the parameters ρ, σ_θ^2 and σ_ε^2 can be identified by three elements of the autocovariance matrix: $Var\left(\ln n_{a,i}\right)$, $Var\left(\ln n_{a-j,i}\right)$, and $Cov\left(\ln n_{a,i}, \ln n_{a-j,i}\right)$ for any arbitrary age $a > 0$ and lag length $j < a$. To identify $\sigma_{\tilde{n}}$ it is useful to the variance of the initial employment level as:

$$Var\left(\ln n_{0,i}\right) = \rho^2\sigma_{\tilde{n}}^2 + \sigma_\theta^2.$$

It follows that the four key parameters can be identified off the autocovariance matrix. Also, any element of the autocovariance matrix is useful for parameter identification. For his reason, we will exploit the entire matrix in the estimation.

While the autocovariance matrix of logged employment levels is sufficient to identify the parameters, other moments may be useful as well. In particular, a large literature on dynamic panel data estimation has developed moment estimators that can be used for the parameter ρ, see e.g. Anderson and Hsiao (1982), Arellano and Bond (1991). Moreover, given that in our case the initial employment levels do not have the same mean as the old stationary distribution, we can also identify ρ directly from following moment condition:

$$E\left(\ln n_{i,a}\right) - E\left(\ln n_{i,a-1}\right) = \rho E\left(\ln n_{i,a-1}\right) - \rho E\left(\ln n_{i,a-2}\right). \tag{8}$$

Intuitively, firms grow on average during the first years of their lives. The speed at which average growth decays is pinned down by the persistence parameter ρ.[4]

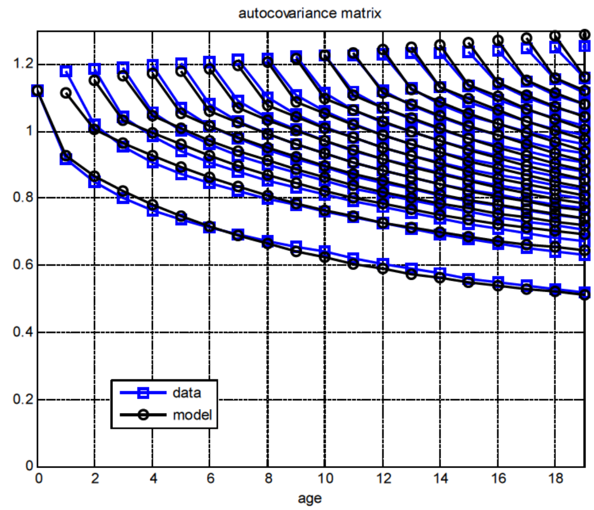
2.5 Estimation results

2.5.1 Baseline estimates

	1 (baseline)	2	3	4
		$\sigma_\theta = 0$	$\rho_s = \rho_u = \rho_v$	$\rho_u = \rho_v$
ρ_s	0.2059	0.7382	0.5096	0.4475
ρ_u	0.8415	0.9884	0.5096	0.9726
ρ_v	0.9489	0.9642	0.5096	0.9726
σ^2_θ	0.3637	-	0.2092	0.0688
$\sigma^2_{\tilde{s}}$	4.1864	0.4360	2.8363	1.5864
$\sigma^2_{\tilde{u}}$	0.5444	0.6761	0.2099	0.6020
σ^2_ε	0.0652	0.0669	0.2173	0.0570
σ^2_z	0.0688	0.0727	0.0000	0.0834
$RMSE$	0.0100	0.0276	0.0773	0.0184
$std(\overline{n}_i)$	0.7594	0	0.9327	0.4747

	1 (baseline)	5 (H&R)	6 (Melitz)
		$\rho_u = \rho_v$ $\sigma_{\tilde{s}} = \sigma_\theta = \sigma_z = 0$	$\rho_s = 0$ only $\sigma_\theta > 0$
ρ_s	0.2059	-	0
ρ_u	0.8415	0.9752	-
ρ_v	0.9489	0.9752	-
σ^2_θ	0.3637	-	0.8519
$\sigma^2_{\tilde{s}}$	4.1864	-	-
$\sigma^2_{\tilde{u}}$	0.5444	0.8225	-
σ^2_ε	0.0652	0.0681	-
σ^2_z	0.0688	-	-
$RMSE$	0.0100	0.0387	0.1575
$std(\overline{n}_i)$	0.7594	0	0.9230

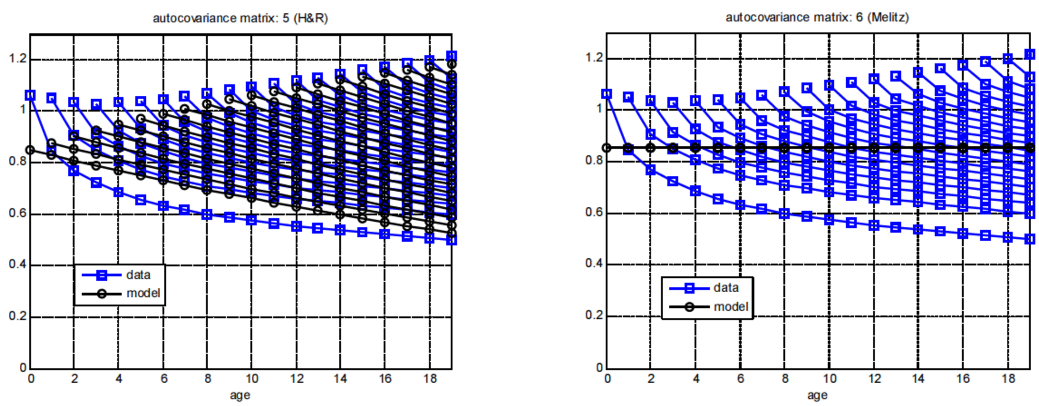
[4] For age groups old enough for employment to have reached its stationary distribution, the condition is not useful to identify ρ, as stationarity implies that $E\left(\ln n_{i,a}\right) - E\left(\ln n_{i,a-1}\right) = E\left(\ln n_{i,a-1}\right) - E\left(\ln n_{i,a-1-2}\right) = 0$. For younger firms, however, the data feature substantial employment growth, see Figure 1.

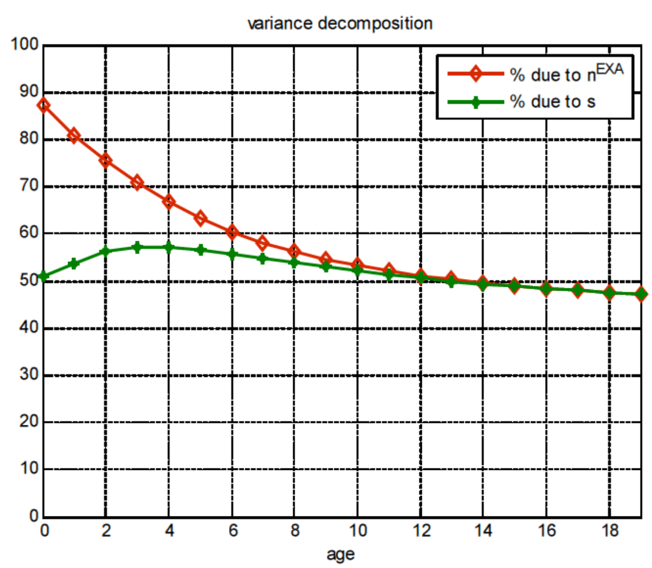


3 Firm growth and aggregate productivity: a structural model

We now explore macroeconomic implications of the nature of the firm growth process, by evaluating how this process impacts on the effects of firm selection on aggregate productivity. We do so by enriching a standard model of firm dynamics with a firm-level shock process that entails both ex-ante and ex-post heterogeneity. We use stylized examples to illustrate how the importance of ex-ante versus ex-post heterogeneity has a critical impact on aggregate productivity, via its effect on firm selection. Specifically, we show that aggregate productivity gains from firm selection are particularly large in an economy in which firm size heterogeneity is mostly driven by ex-ante factors. By contrast, in an economy with only ex-post shocks such gains may be completely absent.

Next, we quantify the effects by matching the model to the empirical evidence presented in the previous subsection. We find that firm selection elevates aggregate productivity by about twenty percent. Moreover, we find that nearly all of this productivity gain is due to selection that happens at the very beginning of firms' life cycles, before they may have even started to produce.


autocovariance matrix: 5 (H&R)
autocovariance matrix: 6 (Melitz)
data
model
age


variance decomposition
% due to n^{EXA}
% due to s
age

3.1 The model

The model is an extension of the closed-economy model presented in Melitz (2003), and features heterogeneous and monopolistically competitive firms and endogenous entry and exit. Unlike Melitz, however, we allow not only for heterogeneity in a fixed, ex-ante productivity parameter, but also for heterogeneity in ex-ante growth profiles (depending on age) and on ex-post shocks. This extension allows the model to match the autocovariance structure of firm-level employment, as well as the age profiles of average size and exit. We consider an economy without aggregate uncertainty and a stationary population of firms. There is an endogenous measure of firms, denoted by Ω, who each produce a unique variety. An infinitely-lived representative household owns the firms and supplies a fixed amount of labor in each period given by $\overline{N}$. Within each period, the household has Dixit-Stiglitz preferences over the varieties. Specifically, household utility derived from consumption is given by

$$Y = \left(\int_{i \in \Omega} a_i^{\frac{1}{\eta}} y_i^{\frac{\eta-1}{\eta}} \, di \right)^{\frac{\eta}{\eta-1}} .$$

elasticity of substitution η over the varieties, a_i is an exogenous but time-varying demand shifter that is specific to good i. In each period of production, a firm must pay a fixed cost f, denominated in units of labor. Each firm $i \in \Omega$ faces a profit function:

$$\pi_i = p_i \left(\frac{p_i}{P} \right)^{-\eta} a_i Y - W(n + f),$$

denominated in units of the aggregate consumption bundle. Here, p_i is the firm's price, P is an aggregate price index, Y is aggregate demand, W is the real wage, σ is the elasticity of substitution between goods varieties. The profit function follows from the optimal consumption decisions of the households and assumes a production function given by $y_i = n_i$, where y_i denotes the firm's output. Profit maximization implies that firms set their price as $p_i = \frac{\sigma}{\sigma-1} WP$ and their employment level as $n_i = a_i \left(\frac{p_i}{P} \right)^{-\sigma} Y$. Hence, firm-level employment moves one-for-one with a_i and heterogeneity across firms fully derives from heterogeneity in demand. However, an alternative interpretation of

a_i is that it represents the productivity of the firm.[5]

The fundamental $a_i \in [a_{\min}, \infty)$ follows a firm-specific stochastic process, governed by a state vector s_i, which in turn follows a Markov process. We can therefore write all firm-level variables as functions of s_i. At the beginning of each period, before observing the new level of s_i, a firm must decide whether to exit or not. This timing assumption follows Hopenhayn (1992). If the firm decides to continue, it is forced to exit exogenously with a probability δ. Let $V(s_i)$ denote the firm value, i.e. the expected present value of firm profits, after observing the current level of s_i. We can express the firm value recursively as:

$$V(s_i) = \pi(s_i) + \beta(1-\delta)\max\left\{\mathbb{E}\left[V(s_i')\middle| s_i\right], 0\right\},$$

where s_i' denotes the value of s_i after the new shocks have been revealed.

Firm entry requires the payment of an entry cost f_e, denominated in units of labor. Following Hopenhayn (1992), we assume that after paying the entry cost, the firm observes an initial level of s_i. The firm can then choose to exit immediately, or continue to pay the fixed cost f, observe the s_i and produce. Free entry requires that:

$$Wf_e \leq \int V(s_i)dG(s_i),$$

where G is the CDF of the distribution from which the initial levels s_i are drawn. Labor market clearing requires that:

$$\overline{N} = M\int n(s_i)\mu(s_i)ds_i,$$

where M is the total measure of active firms and $\mu(s_i)$ is the distribution over the state vector. Aggregate output, aggregate productivity, and the aggregate price index are,

[5] Under this interpretation, preferences would be given by $Y = [\int_{i\in\Omega} c_i^{\frac{\eta-1}{\eta}} di]^{\frac{\eta}{\eta-1}}$, and the production function by $y_i = a_i^{\frac{1}{\eta-1}} n_i$. In our application, both interpretations are observationally equivalent, given that we do not observe prices. Hottman, Redding and Weinstein (2016) use price data to quantify the relative importance of productivity and demand in driving heterogeneity across firms, and find that most differences are due to demand factors.

respectively, given by $Y = M \int a_i(s_i)n(s_i)\mu(s_i)ds_i$, $Y/\overline{N}$, and $P = [\int_{i\in\Omega} a_i n_i^{1-\eta} di]^{\frac{1}{1-\eta}}$ (CHECK). Finally, we normalize the nominal wage WP to one, so that $P = W^{-1}$. We can now define the equilibrium:

Definition. *A stationary equilibrium consists of a real wage W, an aggregate output level Y, a measure of firms, M, and a distribution μ that are consistent with the evolution of the exogenous state s_i, with the firms' optimal decisions regarding exit, employment and price setting, with the free entry condition, with clearing of the labor market, and with clearing of the goods markets.*

3.2 Selection and aggregate productivity in two simplified cases

Before we evaluate the model quantitatively, we study two extreme cases which illuminate the importance of nature of the exogenous process in the determination of aggregate productivity.

Simple case 1: only ex-ante heterogeneity. In the first case, we assume that the firm-level fundamental is time-invariant and drawn ex-ante from a distribution. That is, $s_i = a_i$ is a scalar which is drawn from the ex-ante distribution with CDF G . This is precisely the assumption made by Melitz (2003). For simplicity, we set $\beta = 1$ in this example, as in Melitz (2003). The equilibrium can now be characterized in a simple way, by defining a cutoff level a^* such that any firm exits if and only if $a_i < a^*$. As a result, the productivity distribution of active firms is given by $\mu(a_i) = \frac{G(a_i)}{1-G(a^*)}$ for $a_i \geq a^*$, and $\mu(a_i) = 0$ for $a_i < a^*$. The free-entry condition can now be expressed as a relation between average profits, $\overline{\pi} \equiv \int \mu(a)\pi(a)dG(a)$, and the cutoff a^*:

$$\overline{\pi} = \frac{f_e \delta}{1 - G(a^*)}.$$

Now define $\widetilde{a}\,(a^*) \equiv \left[\int a^{\sigma-1}\mu(a)ds\right]^{\frac{1}{1-\eta}}$, i.e. a weighted average of firm-level productivity. Given that $\mu(a)$ is determined by the cutoff a^*, $\widetilde{a}$ is implicitly a function of a^*. As shown by Melitz (2003), $\widetilde{a}\,(a^*)$ coincides with aggregate productivity (CHECK), so the cutoff directly pins down aggregate productivity. We can now express the exit condition as another relation between $\overline{\pi}$ and a^*:

$$\overline{\pi} = k(a^*) f,$$

where $k(a^*) \equiv \left(\frac{\widetilde{a}(a^*)}{a^*}\right)^{\eta-1} - 1$. The equilibrium is at the intersection of the curves defined by the exit condition and the free-entry condition.

Combining, the two equations, the equilibrium must satisfy $\frac{f_e}{f/\delta} = (1 - G(a^*))\, k(a^*)$, which makes clear that the equilibrium cutoff is determined as a function of entry cost f_e, *relative* to the present value fixed costs to be paid, i.e.f/δ. Note also that if $\frac{f_e}{f/\delta}$ is sufficiently high, the cutoff may be driven down to the point that it hit its minimum i.e. $a_{\min} = a^*$.

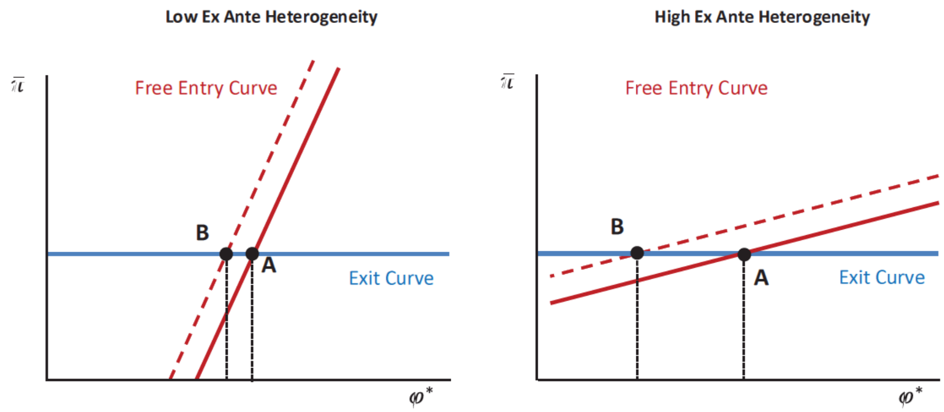
For simplicity, let us further assume that the ex-ante draw comes from a Pareto distribution with scaling parameter α. In this case, it can be shown that $k(a^*) = \left(\frac{\alpha}{\alpha-1}\right)^{\eta-1} - 1$. Thus, the exit condition then pins down $\overline{\pi}$ independently of the level of a^*. The slope of the free-entry condition is given by:

$$\frac{d\overline{\pi}}{da^*} = \frac{\delta f_e}{1 - G(a^*)} \frac{G'(a^*)}{1 - G(a^*)} = \overline{\pi}\alpha.$$

Thus, the slope of the free-entry curve is proportional to the Pareto parameter α. Letting α increase towards infinity, the variance of the Pareto distribution reduces to zero. Thus, the lower the variance of the productivity distribution (and hence the firm size distribution), the steeper the slope of the free-entry condition. This is illustrated by the left and right panel of Figure X.

The figure also illustrates what happens after an increase in the entry cost, which shifts up the free-entry curve.[6] In equilibrium, the cutoff declines, which reduces aggregate productivity. Thus, selection effects increase aggregate productivity in this model. Intuitively, an increase in the entry cost increases the cost of sampling from the distribution of ex-ante draws of the fundamental. This induces firms to sample less

[6] As mentioned above, a decrease in the fixed cost f or an increase in the exogenous exit rate δ as the same effect on the cutoff as an increase in the entry cost f_e.



often, i.e. entry decreases, which in turn has a positive effect on firm profits. This, however, makes firms more willing to continue operation under relatively low draws of the fundamental. That is, exit declines and the cutoff a^* shifts down, which pushes down firm profits. Under a Pareto distribution, the latter effects completely offsets the increase increase in firm profits induced by the decline in entry. The decline in the cutoff, in turn, reduces aggregate productivity.

As illustrated by Figure X, however, the magnitude of the decline depends critically on the slope of the free-entry condition. In the case with a large amount of ex-ante heterogeneity (right panel), the reduction in the cutoff is particularly large. Given that average profits are pinned down by the exit condition, the free-entry condition implies that an increase in the entry cost f_e must be offset by a proportional decline in the probability of successful entry $1-G(a^*)$ must adjust downwards. Under a large amount of heterogeneity, the distribution is spread out which means that a given change in the cutoff a^* has a relatively small effect on the probability of successful entry. Thus, a large decline in the cutoff is required to push down the entry success probability sufficiently in order to restore equilibrium.

Simple case 2: only ex-post heterogeneity. We now consider an opposite case in which there is no ex-ante heterogeneity. Specifically, we now assume that a_i is now determined as an ex-post i.i.d. shock. Recall that, in each period, the exit decision is made before observing the new shock. Since shocks are i.i.d, this implies that no firm has any specific information on its productivity when they make their exit decision. It follows that in equilibrium no firm voluntarily exits, given that new entrants pay the entry cost plus the fixed cost and that exits do still occur for exogenous reasons. As a result, there is no selection in the model with only i.i.d. ex-post shocks. Hence, aggregate productivity is not affected at all by firms' exit decisions.

It is possible combine Case 1 and 2, that is, to consider a model with both permanent ex-ante heterogeneity and i.i.d. ex-post shocks. It is straightforward to verify (Appendix?) that in this case only the ex-ante component of the process affects firm selection. Hence, the strength of selection effects is critically determined by the degree to which overall heterogeneity is determined by mix of ex-ante vis-à-vis the ex-post component. The literature typically makes ad hoc assumptions and extreme assumptions on this mix, like the ones we made above. Our empirical results suggest, however, that reality is more subtle. In the next section, we quantify the strength of selection effects taking into account a realistic mix between ex-ante and ex-post heterogeneity, as revealed by the autocovariance structure of firm-level employment.

3.3 Quantifying the effects of selection on aggregate productivity

We now integrate a more realistic shock process into the model. In line with the reduced-form we postulate the following process:

$$
\begin{aligned}
\ln a_{i,t} &= u_{i,t} + v_{i,t} + w_{i,t} + z_{i,t} \\
u_{i,t} &= \rho_u u_{i,t-1} + \theta_i,\ u_{i,-1} \sim iid(0, \sigma_u),\ \ \theta_i \sim iid(\mu_\theta, \sigma_\theta) \\
v_{i,t} &= \rho_v v_{i,t-1},\ v_{i,-1} \sim iid(0, \sigma_v) \\
w_{i,t} &= \rho_w w_{i,t} + \varepsilon_{i,t},\ w_{i,-1} = 0,\ \ \varepsilon_{i,t} \sim iid(0, \sigma_\varepsilon) \\
z_{i,t} &\sim iid(0, \sigma_z)
\end{aligned}
$$

Note that the firm-level state is given by $s_{i,t} = [u_{i,t}, v_{i,t}, w_{i,t}, z_{i,t}]$. The components $u_{i,t}$ and $v_{i,t}$ jointly capture the ex-ante component of the process, whereas $w_{i,t}$ and $z_{i,t}$

capture the ex-post shocks. Given the parameter values, we solve for the equilibrium using the following algorithm, which follows Hopenhayn and Rogerson (1993):

1. Solve for W from the free-entry condition. Given any value of W, one can compute firm values and exit decision rules. These can then be used to compute the average value of an entrant. Both sides of the free-entry condition can then be evaluated.

2. Simulate the model, normalizing $M = 1$.

3. Solve for M from the labour market clearing condition.

3.3.1 Calibration

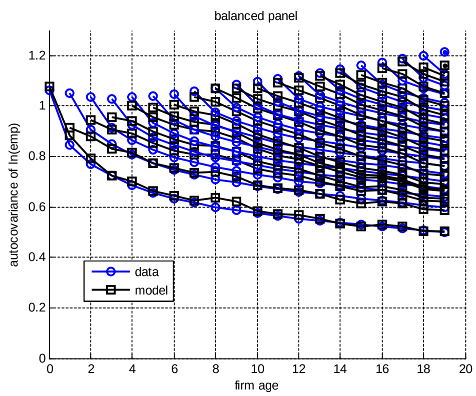
The model period is one year. The parameter values are displayed in the table below.

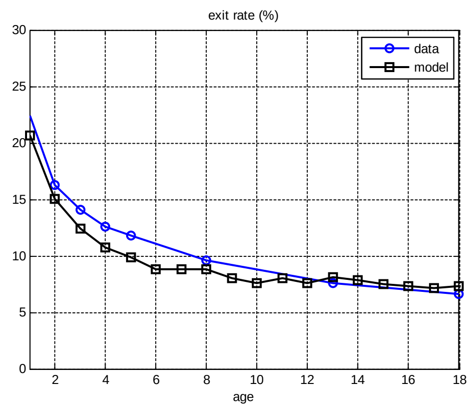
parameter	*value*	**parameter**	*value*	**parameter**	*value*
β	0.96	σ_u	3.2554	μ_θ	-1.7624
η	10	σ_θ	1.3500	ρ_u	0.4432
f	0.3139	σ_v	0.8775	ρ_v	0.9812
f_e	0.2574	σ_ε	0.2844	ρ_w	0.4432
δ	0.0582	σ_f [re-scale?]	0.4380		

The discount factor, β, is set to imply a real interest rate of four percent. The elasticity of substitution between goods, η, implies a markup of 11 percent over marginal costs. The ratio of the entry cost to the fixed cost, $\frac{f_e}{f}$ is set to 0.82, following an empirical estimate reported by Barseghyan and DiCecio (2011). Regarding the shock process, we ease the computational burden by setting $\sigma_z = 0$ and assuming $\rho_v = \rho_w$. The reduced-form evidence suggests that these restrictions are not very costly in terms of the ability to match the empirical autocovariance.

The remaining parameters are chosen to target jointly the autocovariance matrix of employment and the exit and average size profile by age. In order to fit the exit rate profile better, we introduce an iid shock to the entry cost, with mean zero and standard deviation σ_f.

The figures below illustrate the model fit. The model fits very well the autocovariance function and the profile of the exit rate by age. Specifically, the model reproduces the fact that the exit rate initially declines convexly with age and then stabilizes. The model also reproduces the increasing profile of average size, by age, although the profile predicted by the model is steeper than its empirical counterpart.[7]





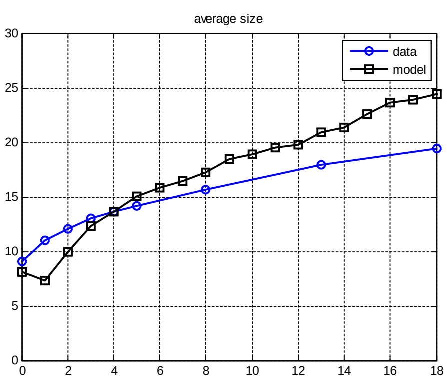


Table XXX further shows exit rates by average size, which was not targeted in the calibration procedure. The model reproduces quite well the exit rate among small age

[7] One possible way to improve the fit would be to allow for age-fixed effects, which would introduce additional state variables. Moreover, the age-fixed effects would be common across firms, and hence would have no direct impact on the amount of heterogeneity across firms at a given age.

bins, and in particular the steep decline going from 1-4 employees and 5-9 employees. For bins of larger firms, the model overpredicts the exit rate.

	data (BDS)	**model**
1 to 4	14.5	15.1
5 to 9	6.5	6.2
10 to 19	5.4	5.9
20 to 49	4.7	5.8
50 to 99	4.2	5.7
100 to 249	4	5.9
250 to 499	3.9	6.4
500 to 999	2.9	5.4

Table. Exit rates by size (in percentages).

3.3.2 Results

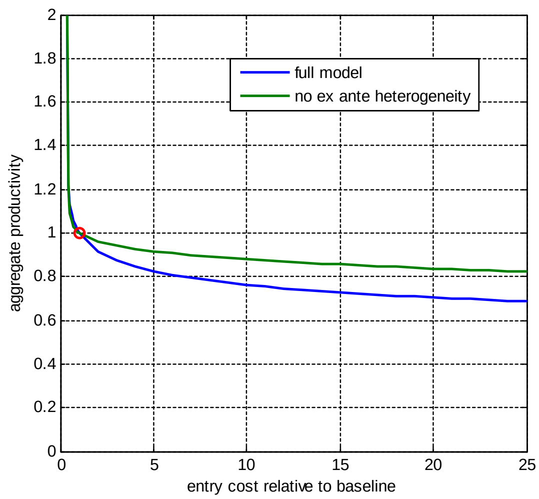
We now quantify the strength of the selection channel. We do so by computing the level of aggregate productivity in the baseline model, and various counterfactual versions. In the first counterfactual, we shut off completely the selection channel by prohibiting firms from exiting endogenously. We then re-compute the wage implied by the free-entry condition, from which we can compute aggregate productivity. As shown in the table below, aggregate productivity is 24 percent lower in the absence of the selection channel. Thus, the aggregate effects from selection are substantial.

	aggregate productivity	**relative to baseline**
Baseline	1.025	-
No selection	0.778	-24.1%
Only selection at startup	0.980	-4.3%

In the second counterfactual, we do allow for endogenous exit, but *only* in the very first period of firms' lives, i.e. immediately after firms have paid their entry cost and observe their ex-ante parameters. In any later period, firms are not allowed to exit endogenously. The table above shows that in this counterfactual, aggregate productivity is only 4.3 percent lower than in the baseline. This implies that 82 percent of the aggregate productivity effect derives from firm selection that takes place before firms even start to produce, at a point in time when firms have only observed their ex-ante parameters. The relative importance of ex-ante heterogeneity in the selection process is thus overwhelming, in line with the simple examples given previously.

Figure XXX plots aggregate productivity as a function of the entry cost, scaled such that aggregate productivity is one under the baseline entry cost. Increasing the entry cost increases the cost of "trial-and-error" and hence it reduces the amount of positive selection, which in turn reduces aggregate productivity. As the entry cost is increased to a level that is 25 times its baseline value, aggregate productivity is reduced by over 30 percent.[8] As the entry cost is reduced, aggregate productivity asymptotes towards infinity, since it becomes costless to select the most productive firms, and productivity has no upper bound. Around the baseline entry cost, the elasticity of aggregate productivity with respect to the entry cost is -0.138.

[8] This reduction is larger than the productivity decline under the "no selection" counterfactual considered above. The reason is that entry costs themselve require labor inputs, but are not counted in aggregate output. Hence, everything else equal higher entry cost directly reduce productivity, on top of the selection channel.



The green line in Figure XXX repeats the exercise, but now using a version of the model in which all ex-ante heterogeneity is shut off. In this model, the selection channel is weaker to begin with and hence aggregate productivity declines by less as the entry cost is increase. Around the baseline entry cost, the elasticity of aggregate productivity with respect to the entry cost -0.068, which is only about half the value in the full model with ex ante heterogeneity.

References

Jaap Abbring and Jeffrey Campbell. A firm's first year, 2005. Tinbergen Institute Discussion Paper 05-046/3.

John Abowd and David Card. On the covariance structure of earnings and hours changes. *Econometrica*, 57(2):4111–445, 1989.

T.W. Anderson and Cheng Hsiao. Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18:47–82, 1982.

Manuel Arellano and Stephen Bond. Some tests of specification for panel data: Monte carlo evidence and application to employment equations. *Review of Economic Studies*, 58(2):277–297, 1991.

Levon Barseghyan and Riccardo Dicecio. Entry costs, industry structure, and cross-country income and tfp differences. *Journal of Economic Theory*, 146:1828–1851, 2011.

Levon Barseghyan and Riccardo Dicecio. Where is silicon valley? *Science*, 347(6222):606–609, 2011.

Eric Bartelsman, John Haltiwanger, and Stefano Scarpetta. Measuring and analyzing crosscountry differences in firm dynamics, 2009. In *Producer Dynamics: New Evidence from Micro Data*,NBER Chapters, pp. 15âĂŞ76. National Bureau of Economic Research, Inc.

Lucia Foster, John Haltiwanger, and Chad Syverson. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425, 2008.

Fatih Guvenen and Anthony Smith. Inferring labor income risk and partial insurance from economic choices. *Econometrica*, 82(6):2085–2129, 2014.

Fatih Guvenen. Learning your earning: Are labor income shocks really very persistent? *American Economic Review*, 97(3):687–712, 2007.

Fatih Guvenen. An empirical investigation of labor income processes. *Review of Economic Dynamics*, 12(1):58–79, 2009.

John Haltiwanger, Ron Jarmin, Robert Kulick, and Javier Miranda. High growth young firms: Contribution to job, output and productivity growth, 2014. working paper.

Hugo Hopenhayn and Richard Rogerson. Job turnover and policy evaluation: A general equilibrium analysis. *Journal of Political Economy*, 101(5):915–938, 1993.

Hugo Hopenhayn. Entry, exit and firm dynamics long run equilibrium. *Econometrica*, 60(5):1127–1150, 1992.

Chiang-Tai Hsieh and Peter J. Klenow. Misallocation and manufacturing tfp in china and india. *Quarterly Journal of Economics*, 4:1403–1448, 2009.

Erik Hurst and Benjamin Pugsley. What do small businesses do?, 2011. NBER working paper no. 17041.

Boyan Jovanovic. Selection and the evolution of industry. *Econometrica*, 50(3):649–670, 1982.

Thomas MaCurdy. The use of time-series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics*, 18:83–114, 1982.

Marc Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71:1695–1725, 2003.

Diego Restuccia and Richard Rogerson. Policy distortions and aggregate productivity with heterogeneous plants. *Review of Economic Dynamics*, 11(4):707–720, 2008.

Petr Sedláček and Vincent Sterk. The growth potential of startups over the business cycle, 2016. working paper.

4 Appendix

4.1 GMM estimation and overidentification from autocovariance

Repeating from above: (CHANGE NOTATION TO MATCH PAPER)

Generally, the autocovariance function for $a, j \geq 0$ is:

$$\begin{aligned}
\text{Cov}\left[\log z_a, \log z_{a+j}\right] &= \sigma_\varepsilon^2 \sum_{k=0}^{a-1} \rho^{2k+j} + \sigma_\theta^2 \sum_{k=0}^{a} \rho^k \sum_{k=0}^{a+j} \rho^k + \rho^{2(a+1)+j} \sigma_z^2 && (9) \\
&= \sigma_\varepsilon^2 \rho^j \frac{1-\rho^{2a}}{1-\rho^2} + \sigma_\theta^2 \frac{1-\rho^{a+1}}{1-\rho} \frac{1-\rho^{a+j+1}}{1-\rho} + \rho^{2(a+1)+j} \sigma_z^2 && (10)
\end{aligned}$$

Exact identification of $\rho, \sigma_\theta$and σ_z follows from the first three autocovariances, $\text{Cov}\left[\log z_0, \log z_a\right]$ $0 \leq a \leq 2$, which are a nonlinear function of only ρ, σ_θ and σ_z. These can be solved in closed form for the parameters:

$$\begin{aligned}
\frac{\text{Cov}\left[\log z_0, \log z_2\right] - \text{Cov}\left[\log z_0, \log z_1\right]}{\text{Cov}\left[\log z_1, \log z_0\right] - \text{Cov}\left[\log z_0, \log z_0\right]} &= \rho \frac{\rho\sigma_\theta^2 + \rho\sigma_z^2}{\rho\sigma_\theta^2 + \rho^3\sigma_z^2} \\
&= \rho
\end{aligned}$$

$$\text{Cov}\left[\log z_1, \log z_0\right] - \rho \text{Cov}\left[\log z_0, \log z_0\right] = \sigma_\theta^2$$

$$\frac{\text{Cov}\left[\log z_0, \log z_0\right] - \sigma_\theta^2}{\rho^2} = \sigma_z^2$$

Adding a fourth moment identifies σ_ε

$$\text{Cov}\left[\log z_1, \log z_1\right] - \rho^2 \text{Cov}\left[\log z_0, \log z_0\right] = \sigma_\varepsilon^2 + (1 + 2\rho)\, \sigma_\theta^2.$$

4.1.1 Nonlinear GMM estimation

Let $\theta = \left(\rho, \sigma_\theta^2, \sigma_z^2, \sigma_\varepsilon^2\right)'$ be an arbitrary parameter vector in compact parameter space $\mathbb{P}$. Since we use ages 0 to A, we define the $\frac{A*(A+1)}{2}$ length vector valued function

$$f\left(n_i, \theta\right) = \left[\left(\log n_{ia} - E\left[\log n_{ia}\right]\right) \log n_{ij} - \text{Cov}\left[\log n_{ia}, \log n_{ia+j}; \theta\right]\right]$$

where $a = 0, \ldots, A$ and $j = a, a + a, \ldots, A$. Let θ_0 be the true parameter vector, so that identification follows from $E[f(n_i; \theta)] = 0$ iff $\theta = \theta_0$. The term $\mathrm{Cov}[\log n_{ia}, \log n_{ia+j}; \theta]$ is a constant and equal to the formula from equation **??** computed for an arbitrary parameter vector θ.

Writing out some of the elements of f:

$$f(n_i; \theta_0) = \begin{bmatrix} (\log n_{i0} - E[\log n_{i0}]) \log n_{i0} - \frac{\sigma_{\theta 0}^2 + \rho^2 \sigma_z^2}{(1-\alpha)^2} \\ (\log n_{i0} - E[\log n_{i0}]) \log n_{i1} - \frac{(1+\rho_0)\sigma_\theta^2 + \rho^3 \sigma^2}{(1-\alpha)^2} \\ \vdots \\ (\log n_{i0} - E[\log n_{i0}]) \log n_{iA} - \frac{\sigma_\theta^2 \sum_{k=0}^{10} \rho^k + \rho^{12} \sigma_z^2}{(1-\alpha)^2} \\ (\log n_{i1} - E[\log n_{i1}]) \log n_{i1} - \frac{\sigma_\varepsilon^2 + (1+\rho)^2 \sigma_\theta^2 + \rho^4 \sigma_z^2}{(1-\alpha)^2} \\ \vdots \\ (\log n_{iA} - E[\log n_{iA}]) \log n_{i1} - \frac{\sigma_\varepsilon^2 \sum_{k=0}^{A-1} \rho^{2k+A} + \sigma_\theta^2 \sum_{k=0}^{A} \rho^k \sum_{k=0}^{2A} \rho^k + \rho^{2(A+1)+A} \sigma_z^2}{(1-\alpha)^2} \end{bmatrix}$$

Define the sample analog to $E[f(n_i; \theta)]$

$$g_N(\theta) \equiv \frac{1}{N} \sum_i f(n_i; \theta).$$

A law of large numbers implies $g_N(\theta) \to^p E[f(n_i; \theta)]$. Define the GMM estimator

$$\tilde{\theta}_N = \operatorname*{argmin}_{\theta \in \mathbb{P}} g_N(\theta)' W g_N(\theta)$$

for an arbitrary symmetric positive definite weighting matrix W. The asymptotic distribution of the estimator $\tilde{\theta}_N$ is:[9]

$$\sqrt{N}\left(\tilde{\theta}_N - \theta_0\right) \to_d N(0, \Sigma)$$

[9] To see this, write $g_N\left(\tilde{\theta}_N\right)$ as

$$g_N\left(\tilde{\theta}_N\right) \approx g_N(\theta_0) + \frac{\partial g_N(\theta_0)}{\partial \theta'}\left(\tilde{\theta}_N - \theta_0\right).$$

Multiplying through by $\frac{\partial g_N(\tilde{\theta}_N)}{\partial \theta'} W$ so that the LHS is equal to the first order condition where

$$\begin{aligned}\Sigma &\equiv \left(d'Wd\right)^{-1}\left(d'WVWd\right)\left(d'Wd\right)^{-1}\\ d &\equiv \frac{\partial Ef\left(n_i;\theta_0\right)}{\partial\theta'}\\ V &\equiv E\left[f\left(n_i;\theta_0\right)f\left(n_i;\theta_0\right)'\right].\end{aligned}$$

Note V is not a covariance matrix for $\log n_{ia}$ since $E[f]$ is the unique elements of the covariance matrix.

Estimation To operationalize the estimator we have to estimate both V and the means $E[\log n_i]$. Define

$$\tilde{f}\left(n_i,\theta\right) \equiv \left[\left(\log n_{ia} - \frac{1}{N}\sum_{i'}\log n_{ia}\right)\log n_{ij} - \text{Cov}\left[\log n_{ia}, \log n_{ia+j};\theta\right]\right]$$

$$\tilde{g}_N\left(\theta\right) = \frac{1}{N}\sum_{i=1}\tilde{f}\left(n_i,\theta\right),$$

where $a = 0, \ldots, 10$ and $j = a, a+a, \ldots, 10$. Define the $A\left(A+1\right)/2 \times A\left(A+1\right)/2$ moment covariance matrix

$$\tilde{V}_N = \frac{1}{N}\sum_{i=1}\left[\tilde{f}\left(n_i,\theta\right)\tilde{f}\left(n_i,\theta\right)'\right] = \frac{1}{N}\sum_{i=1}\left(h\left(n_i\right)-\bar{h}\right)h\left(n_i\right)'.$$

$\frac{\partial g_N\left(\tilde{\theta}_N\right)}{\partial\theta'}Wg_N\left(\tilde{\theta}_N\right) = 0$, then

$$\begin{aligned}0 &\approx \frac{\partial g_N\left(\tilde{\theta}_N\right)}{\partial\theta'}Wg_N\left(\theta_0\right) + \frac{\partial g_N\left(\tilde{\theta}_N\right)}{\partial\theta'}W\frac{\partial g_N\left(\theta_0\right)}{\partial\theta'}\left(\tilde{\theta}_N-\theta_0\right)\\ \left(\tilde{\theta}_N-\theta_0\right) &\approx -\left(\frac{\partial g_N\left(\tilde{\theta}_N\right)}{\partial\theta'}W\frac{\partial g_N\left(\theta_0\right)}{\partial\theta'}\right)^{-1}\frac{\partial g_N\left(\tilde{\theta}_N\right)}{\partial\theta'}Wg_N\left(\theta_0\right).\end{aligned}$$

Letting $N\to\infty$ then

$$\sqrt{N}\left(\tilde{\theta}_N-\theta_0\right)\to^d -\left(\frac{\partial g_N\left(\theta_0\right)}{\partial\theta'}W\frac{\partial g_N\left(\theta_0\right)}{\partial\theta'}\right)^{-1}\frac{\partial g_N\left(\theta_0\right)}{\partial\theta'}W\sqrt{N}g_N\left(\theta_0\right)$$

since $\tilde{\theta}_N\to^p\theta_0$. And from the CLT $\sqrt{N}g_N\left(\theta_0\right)\to^d N\left(0,V\right)$.

Note that since $\tilde{f}(n_i, \theta) = h(n_i) - q(\theta) = 0$ then $E\left[\tilde{f}_i \tilde{f}_i\right] = \text{Cov}\left[\tilde{f}_i, \tilde{f}_i\right] = \text{Cov}\left[h(n_i), h(n_i)\right]$.

To deal with missing data. We can create an indicator variable λ_{iaj} for whether or not the observation is missing and then define

$$\tilde{f}(n_i, \theta, \lambda_i) \equiv \left[\lambda_{iaj}\left(\left(\log n_{ia} - \frac{1}{N}\sum_{i'} \log n_{ia}\right)\log n_{ij} - \text{Cov}\left[\log n_{ia}, \log n_{ia+j}; \theta\right]\right)\right]$$

and use weighting matrix

$$A = \Pi^{-1}\Pi^{-1}$$

where

$$\Pi = \begin{bmatrix} \frac{N_{00}}{N} & & & \\ & \frac{N_{01}}{N} & & \\ & & \ddots & \\ & & & \frac{N_{AA}}{N} \end{bmatrix}$$

and N_{aj} is the number of non missing observations for that moment. This is equivalent to equally weighting but uses the correct number of observations when computing $\tilde{V}$

4.2 Autocovariance of employment

Provide the full autocovariance matrix for reference