NBER WORKING PAPER SERIES

ENTRY, EXIT, FIRM DYNAMICS, AND AGGREGATE FLUCTUATIONS

Gian Luca Clementi
Berardino Palazzo

Working Paper 19217
http://www.nber.org/papers/w19217

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2013

We are grateful to Dave Backus, Rudi Bachman, Susanto Basu, Russel Cooper, Jason Faberman, Shigeru Fujita, Ramon Marimon, Gianluca Violante, and Stan Zin, as well seminar attendants at Boston College, Boston University, European University Institute, Richmond Fed, New York University, University of Southern California, University of Texas at Austin, University of Virginia, University of Western Ontario, SED meeting, Canadian Macroeconomics Study Group, CREI-MOVE Workshop on Misallocation and Productivity, Greenline Macroeconomics Meeting, NY/Philly Macroeconomics Workshop, and Workshop on Macroeconomic Dynamics for their comments and suggestions. All remaining errors are our own responsibility. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Entry, Exit, Firm Dynamics, and Aggregate Fluctuations
Gian Luca Clementi and Berardino Palazzo
NBER Working Paper No. 19217
July 2013
JEL No. D92,E23,E32,L11

ABSTRACT

Do firm entry and exit play a major role in shaping aggregate dynamics? Our answer is yes. Entry and exit propagate the effects of aggregate shocks. In turn, this results in greater persistence and unconditional variation of aggregate time-series. These are features of the equilibrium allocation in Hopenhayn (1992)'s model of equilibrium industry dynamics, amended to allow for investment in physical capital and aggregate fluctuations. In the aftermath of a positive productivity shock, the number of entrants increases. The new firms are smaller and less productive than the incumbents, as in the data. As the common productivity component reverts to its unconditional mean, the new entrants that survive become more productive over time, keeping aggregate efficiency higher than in a scenario without entry or exit.

Gian Luca Clementi
Department of Economics
Stern School of Business
New York University
44 West Fourth Street
New York, NY 10012
and NBER
clem@nyu.edu

Berardino Palazzo
Department of Finance
Boston University School of Management
595 Commonwealth Avenue
Boston, MA 02215
bpalazzo@bu.edu

1 Introduction

During the last 25 years or so, empiricists have pointed out a tremendous amount of between–firm and between–plant heterogeneity, even within narrowly defined sectors. Yet, for most of its young life the modern theory of business cycles has completely disregarded such variation. What is the loss of generality implied by this methodological choice?

There are many reasons why heterogeneity may matter for aggregate fluctuations, some of which have received a substantial attention in the literature.[1] Our goal is to contribute to the understanding of the role played by entry and exit. What are, if any, the costs of abstracting from firm entry and exit when modeling aggregate fluctuations?

Our main result is that entry and exit propagate the effects of aggregate shocks. We find this to be a feature of the equilibrium allocation in Hopenhayn (1992)'s model of industry dynamics, amended to allow for investment in physical capital and for aggregate fluctuations. We assume that firms' productivity is the product of a common and of an idiosyncratic component, both driven by persistent stochastic processes and orthogonal to each other. Differently from Hopenhayn (1992), potential entrants are in finite mass and face different probability distributions over the first realization of the idiosyncratic shock.

We assume that the demand for firms' output and the supply of physical capital are infinitely elastic at the unit price, while the supply of labor services has finite elasticity. The wage rate fluctuates to ensure that the labor market clears. This is crucial, as it is often the case that the effects of shocks on endogenous variables are muted or reversed by the ensuing adjustment in prices.

When parameterized to match a set of empirical regularities on investment, entry, and exit, our framework replicates well–documented stylized facts about firm dynamics. To start with, the exit hazard rate declines with age. Employment growth is decreasing with size and age, both unconditionally and conditionally. The size distribution of firms is skewed to the right and the skewness of a cohort's size distribution declines with age. Furthermore, the entry rate is pro–cyclical, while the exit rate is counter–cyclical.

The mechanics of entry is straightforward. A positive shock to the common productivity component makes entry more appealing. This is the case because, with a labor supply elasticity calibrated to match the standard deviation of employment relative to output, the equilibrium response of the wage rate is not large enough to undo the impact of the

[1] This is the case for the possibility that the occasional synchronization in the timing of establishments' investment may influence aggregate dynamics when nonconvex capital adjustment costs lead establishments to adjust capital in a lumpy fashion. See Veracierto (2002) and Khan and Thomas (2003, 2008).

innovation in aggregate productivity.

Entrants are more plentiful, but of lower average idiosyncratic efficiency. This happens because, with a time–invariant cost of entry, firms with poorer prospects about their productivity find it worth entering. Aggregate output and TFP are lower than they would be in the absence of this selection effect. However, given the small output share of entering firms, the contemporaneous response of output is not very different from the one that obtains in a model that abstracts from entry and exit.

It is the evolution of the new entrants that causes a sizeable impact on aggregate dynamics. As the common productivity component declines towards its unconditional mean, there is a larger–than–average pool of young firms that increase in efficiency and size. The distribution of firms over idiosyncratic productivity improves. It follows that entry propagates the effects of aggregate productivity shocks on output and increases its unconditional variance.

For a version of our model without entry or exit to generate a data–conforming persistence of output, the first–order autocorrelation of aggregate productivity shocks must be as high as 0.775. In the benchmark scenario with entry and exit, it needs only be 0.685. As pointed out by Cogley and Nason (1995), many Real–Business–Cycle models have weak internal propagation mechanisms. In order to generate the persistence in aggregate time–series that we recover in the data, they must rely heavily on external sources of dynamics. Our work shows that allowing for firm heterogeneity and for entry and exit can sensibly reduce such reliance.

The propagation result depends on the pro–cyclicality of the entry rate, for which evidence abounds,[2] and on the dynamics of young firms. According to our theory, the relative importance of a cohort is minimal at birth and increases over time. Strong support for this prediction comes from Foster, Haltiwanger, and Krizan (2001), who document that the (positive) productivity gap between entrants and exiters grows larger over time.

In our setup idiosyncratic shocks only affect technical efficiency. In reality, however, plants also face random shifts in input supply and product demand schedules, which lead to changes in input and product prices, respectively.

It must be clear that the main implications of our theory do not hinge upon the origin of plant–level shocks. In particular, they will still hold in setups with horizontal product differentiation where idiosyncratic demand shocks play a major role, as long as such models admit reduced forms where innovations in demand boil down to shocks to Revenue Total Factor Productivity (TFPR) that resemble our own disturbances.

[2] See Campbell (1998) and Lee and Mukoyama (2012).

Exploiting data from the five Census of Manufacturing in the period 1977–1997, Foster, Haltiwanger, and Syverson (2008, 2012) conclude that in the case of ten seven–digit industries, the dynamics of TFPR was mostly the result of idiosyncratic demand shocks. Entrants displayed about the same technical efficiency as incumbents, but faced much lower demand schedules, and therefore charged lower prices. Over time, conditional on survival, demand shifted outwards, leading to higher prices and greater TFPR.

To rationalize these findings, Foster, Haltiwanger, and Syverson (2012) posit a stochastic process for idiosyncratic demand shocks which leads to a TFPR dynamics very similar to ours. Therefore we conjecture that our results for cross–sectional heterogeneity and aggregate dynamics would not change if we amended our model to allow for idiosyncratic random variation in demand.

A final caveat is that ours is not a theory of the firm. That is, we do not provide an explanation for why single–plant and multi–plant business entities coexist. While our calibration relies exclusively on plant–level data, throughout the paper we will alternatively refer to production units as firms or plants.

In recent years, a number of scholars have built upon Hopenhayn (1992) to produce novel theories of aggregate fluctuations with firm heterogeneity. We think for example at the business cycle theories of Veracierto (2002), Khan and Thomas (2003, 2008), and Bachman and Bayer (2013), the asset pricing model by Zhang (2005), and the work by Lee and Mukoyama (2012). With the exception of the latter, however, all of those papers abstract from entry and exit. Lee and Mukoyama (2012)'s framework differs from ours in key modeling assumptions. In particular, they do not model capital accumulation and let the free–entry condition pin down the wage rate.

Campbell (1998), one of the earliest treatments of entry and exit in a model with aggregate fluctuations, focuses on investment–specific technology shocks and makes a list of assumptions with the purpose of ensuring aggregation. In turn, this leads to an environment that has no implications for most features of firm dynamics. Cooley, Marimon, and Quadrini (2004) and Samaniego (2008) characterize the equilibria of stationary economies with entry and exit and study their responses to zero–measure aggregate productivity shocks.

A somewhat different strand of papers, among which Devereux, Head, and Lapham (1996), Chatterjee and Cooper (1993), and Bilbiie, Ghironi, and Melitz (2012), model entry and exit in general equilibrium models with monopolistic competition, but abstract completely from firm dynamics. Interestingly, also in these frameworks entry leads to a propagation of exogenous random disturbances. The reason, however, is very different.

Entry increases the diversity of the product space. Because of increasing returns, this encourages agents to work harder and accumulate more capital.

The remainder of the paper is organized as follows. The model is introduced in Section 2. In Section 3, we characterize firm dynamics in the stationary economy. The analysis of the scenario with aggregate fluctuations begins in Section 4, where we describe the impact of aggregate shocks on the entry and exit margins. In Section 5, we characterize the cyclical properties of entry and exit rates, as well as the relative size of entrants and exiters. We also gain insights into the mechanics of the model by describing the impulse responses to an aggregate productivity shock. In Section 6, we illustrate how allowing for entry and exit strengthen the model's internal propagation mechanism and we outline its implications for the cyclicality of the cross–sectoral distribution of productivity. Section 7 concludes.

2 Model

Time is discrete and is indexed by $t = 1, 2, ...$ The horizon is infinite. At time t, a positive mass of price–taking firms produce an homogenous good by means of the production function $y_t = z_t s_t (k_t^\alpha l_t^{1-\alpha})^\theta$, with $\alpha, \theta \in (0, 1)$. With k_t we denote physical capital, l_t is labor, and z_t and s_t are aggregate and idiosyncratic random disturbances, respectively.

The common component of productivity z_t is driven by the stochastic process

$$\log z_{t+1} = \rho_z \log z_t + \sigma_z \varepsilon_{z,t+1},$$

where $\varepsilon_{z,t} \sim N(0, 1)$ for all $t \geq 0$. The dynamics of the idiosyncratic component s_t is described by

$$\log s_{t+1} = \rho_s \log s_t + \sigma_s \varepsilon_{s,t+1},$$

with $\varepsilon_{s,t} \sim N(0, 1)$ for all $t \geq 0$. The conditional distribution will be denoted as $H(s_{t+1}|s_t)$.

Firms hire labor services on the spot market at the wage rate w_t and discount future profits by means of the time–invariant factor $\frac{1}{R}$, $R > 1$. Adjusting the capital stock by x bears a cost $g(x, k)$. Capital depreciates at the rate $\delta \in (0, 1)$.

We assume that the demand for the firm's output and the supply of capital are infinitely elastic and normalize their prices at 1. The supply of labor is given by the function $L_s(w) = w^\gamma$, with $\gamma > 0$.

Operating firms incur a cost $c_f > 0$, drawn from the common time-invariant distribution G. Firms that quit producing cannot re–enter the market at a later stage and recoup the undepreciated portion of their capital stock, net of the adjustment cost of driving it to 0. The timing is summarized in Figure 1.

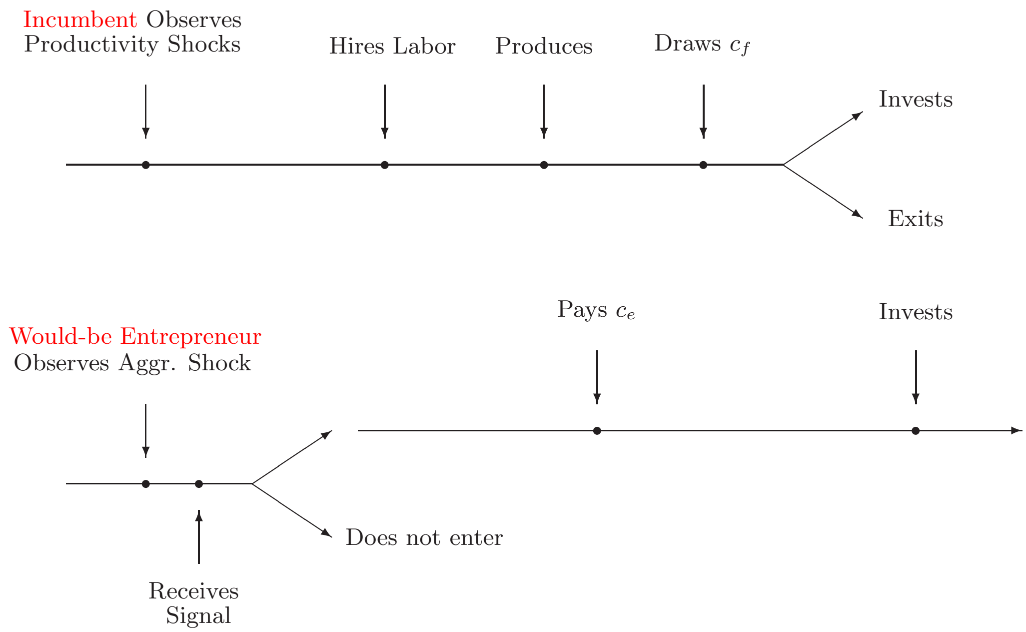

Figure 1: Timing in period t.

Every period there is a constant mass $M > 0$ of prospective entrants, each of which receives a signal q about their productivity, with $q \sim Q(q)$. Conditional on entry, the distribution of the idiosyncratic shock in the first period of operation is $H(s'|q)$, strictly decreasing in q. Entrepreneurs that decide to enter the industry pay an entry cost $c_e \geq 0$.

At all $t \geq 0$, the distribution of operating firms over the two dimensions of heterogeneity is denoted by $\Gamma_t(k, s)$. Finally, let $\lambda_t \in \Lambda$ denote the vector of aggregate state variables and $J(\lambda_{t+1}|\lambda_t)$ its transition operator. In Section 4, we will show that $\lambda_t = \{\Gamma_t, z_t\}$.

2.1 The incumbent's optimization program

Given the aggregate state λ, capital in place k, and idiosyncratic shock s, the employment choice is the solution to the following static problem:[3]

$$\pi(\lambda, k, s) = \max_{l} \; sz[k^{\alpha} l^{1-\alpha}]^{\theta} - wl.$$

Upon exit, a firm obtains a value equal to the undepreciated portion of its capital k, net of the adjustment cost it incurs in order to dismantle it, i.e. $V_x(k) = k(1-\delta) - g\left[-k(1-\delta), k\right]$.

Then, the start–of–period value of an incumbent firm is dictated by the function $V(\lambda, k, s)$ which solves the following functional equation:

$$V(\lambda, k, s) = \pi(\lambda, k, s) + \int \max\left\{V_x(k), \widetilde{V}(\lambda, k, s) - c_f\right\} dG(c_f)$$

[3]We drop time indexes in the remainder of the paper, except in parts where such choice may jeopardize clarity of exposition.

$$\widetilde{V}(\lambda, k, s) = \max_{x} \; -x - g(x,k) + \frac{1}{R}\int_{\Lambda}\int_{\Re} V(\lambda', k', s')dH(s'|s)dJ(\lambda'|\lambda)$$
$$\text{s.t. } k' = k(1-\delta) + x$$

2.2 Entry

For an aggregate state λ, the value of a prospective entrant that obtains a signal q is

$$V_e(\lambda, q) = \max_{k'} -k' + \frac{1}{R}\int V(\lambda', k', s')dH(s'|q)dJ(\lambda'|\lambda).$$

She will invest and start operating if and only if $V_e(\lambda, q) \geq c_e$.

2.3 Recursive Competitive Equilibrium

For given Γ_0, a recursive competitive equilibrium consists of (i) value functions $V(\lambda, k, s)$, $\widetilde{V}(\lambda, k, s)$ and $V_e(\lambda, q)$, (ii) policy functions $x(\lambda, k, s)$, $l(\lambda, k, s)$, $k'(\lambda, q)$, and (iii) bounded sequences of wages $\{w_t\}_{t=0}^{\infty}$, incumbents' measures $\{\Gamma_t\}_{t=1}^{\infty}$, and entrants'measures $\{\mathcal{E}_t\}_{t=0}^{\infty}$ such that, for all $t \geq 0$,

1. $V(\lambda, k, s)$, $\widetilde{V}(\lambda, k, s)$, $x(\lambda, k, s)$, and $l(\lambda, k, s)$ solve the incumbent's problem;

2. $V_e(\lambda, q)$ and $k'(\lambda, q)$ solve the entrant's problem;

3. The labor market clears: $\int l(\lambda_t, k, s)d\Gamma_t(k, s) = L^s(w_t) \; \forall \; t \geq 0$,

4. For all Borel sets $\mathcal{S} \times \mathcal{K} \subset \Re^+ \times \Re^+$ and $\forall \; t \geq 0$,

$$\mathcal{E}_{t+1}(\mathcal{S} \times \mathcal{K}) = M \int_{\mathcal{S}}\int_{\mathcal{B}_e(\mathcal{K}, \lambda_t)} dQ(q)dH(s'|q),$$

   where $\mathcal{B}_e(\mathcal{K}, \lambda_t) = \{q \text{ s.t. } k'(\lambda_t, q) \in \mathcal{K} \text{ and } V_e(\lambda_t, q) \geq c_e\}$;

5. For all Borel sets $\mathcal{S} \times \mathcal{K} \subset \Re^+ \times \Re^+$ and $\forall \; t \geq 0$,

$$\Gamma_{t+1}(\mathcal{S} \times \mathcal{K}) = \int_{\mathcal{S}}\int\int_{\mathcal{B}(\mathcal{K}, \lambda_t, c_f)} d\Gamma_t(k, s)dG(c_f)dH(s'|s) + \mathcal{E}_{t+1}(\mathcal{S} \times \mathcal{K}),$$

   where $\mathcal{B}(\mathcal{K}, \lambda_t, c_f) = \Big\{(k, s) \text{ s.t. } \widetilde{V}(\lambda_t, k, s) - c_f \geq V_x(k) \text{ and } k(1-\delta) + x(\lambda_t, k, s) \in \mathcal{K}\Big\}$.

3 The Stationary Case

We begin by analyzing the case without aggregate shocks, i.e. $\sigma_z = 0$. In this scenario, our economy converges to a stationary equilibrium in which all aggregate variables are constant. First, we introduce our assumptions about functional forms. Then we illustrate the mechanics of entry, investment, and exit. After calibrating the model, we detail the model's implications for several features of firm dynamics, among which the relation between firm growth and survival, size, and age.

3.1 Functional Forms

Investment adjustment costs are the sum of a fixed portion and of a convex portion:

$$g(x,k) = \chi(x)c_0 k + c_1 \left(\frac{x}{k}\right)^2 k, \quad c_0, c_1 \geq 0,$$

where $\chi(x) = 0$ for $x = 0$ and $\chi(x) = 1$ otherwise. Notice that the fixed portion is scaled by the level of capital in place and is paid if and only if gross investment is different from zero.

The distribution of signals for the entrants is Pareto. See the left panel of Figure 2. We posit that $q \geq \underline{q} \geq 0$ and that $Q(q) = (\underline{q}/q)^{\xi}$, $\xi > 1$. The realization of the idiosyncratic shock in the first period of operation follows the process $\log(s) = \rho_s \log(q) + \sigma_s \eta$, where $\eta \sim N(0,1)$.

Finally, we assume that the distribution of the operating cost G is log–normal with parameters μ_{c_f} and σ_{c_f}.

3.2 Entry, Investment, and Exit

Since an incumbent's value $V(k,s)$ is weakly increasing in the idiosyncratic productivity shock s and the conditional distribution $H(s'|q)$ is decreasing in the signal q, the value of entering $V_e(q)$ is a strictly increasing function of the signal. In turn, this means that there will be a threshold for q, call it q^*, such that prospective entrants will enter if and only if they receive a draw greater than or equal to q^*.

The mass of actual entrants with productivity less than or equal to any $\bar{s} \geq 0$ will be $M \int^{\bar{s}} \int_{q^*} dQ(q) dH(s|q)$. See the right panel of Figure 2.

Let $k'(q)$ denote the optimal entrants' capital choice conditional on having received a signal q. For every $\bar{k} \geq k'(q^*)$ and for all $\bar{s} \geq 0$, the portion of entrants characterized by pairs (s,k) such that $s \leq \bar{s}$ and $k \leq \bar{k}$ will be $M \int^{\bar{s}} \int_{q^*}^{q^{-1}(\bar{k})} dQ(q) dH(s|q)$, where q^{-1} is the inverse of the mapping between signal and capital choice.

Our treatment of the entry problem is different from that in Hopenhayn (1992). There, prospective entrants are identical. Selection upon entry is due to the fact that firms that paid the entry cost start operating only if their first productivity draw is greater than the exit threshold. In our framework, prospective entrants are heterogeneous. Some

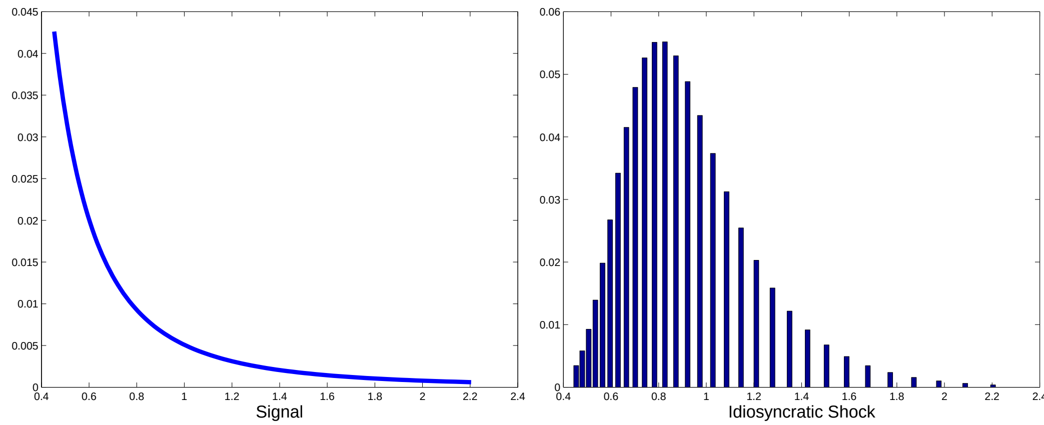

Figure 2: Signal's distribution (left) and productivity distribution of actual entrants.

obtain a greater signal than others and therefore face better short–term prospects. As a consequence, they start with greater capital and are likely to hire more employees.

Investment is modeled as in the standard neoclassical framework with non-convex adjustment costs. See for example Khan and Thomas (2008). The fixed cost of investment gives rise to policy functions of the S–s type, which in turn are responsible for the periods of inaction that are characteristic of all firm– and plant–level datasets.

Firms exit whenever they draw an operating cost c_f such that the value of exiting $V_x(k)$ is greater than the value of continuing operations, i.e. $\widetilde{V}(k,s) - c_f$. Figure 3 depicts the probability of survival as a function of the state variables. For given capital, the probability of exiting is decreasing in the idiosyncratic productivity. This is not surprising, given that the value of exiting does not depend on s, while the continuation payoff is obviously greater, the larger is productivity. Survival chances also improve with capital, for given level of productivity. Both the value of continuing and exiting are increasing in capital, but simple inspection reveals that the marginal impact of raising capital is greater on the value of staying.

Differently from Hopenhayn (1992), where firms exit with certainty when productivity falls below a certain threshold and continue with certainty otherwise, in our framework all firms survive with positive probability and survival is smoothly increasing in both capital and productivity. If we let the distribution of the operating cost collapse to a singleton, the exit policy would be characterized by a decreasing schedule, call it $\underline{s}(k)$, such that a firm equipped with capital k will exit if and only if its productivity is lower than $\underline{s}(k)$.

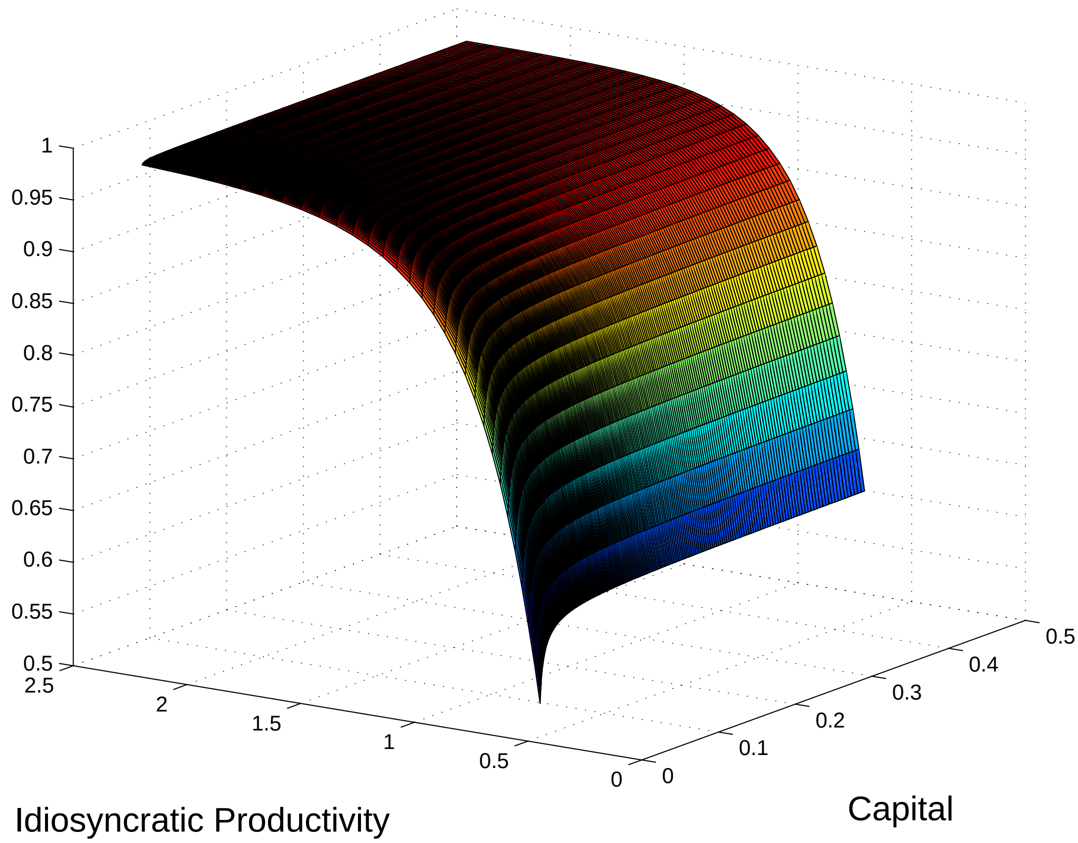

Figure 3: Conditional probability of survival.

3.3 Calibration

Our parameter values are listed in Table 1. One period is assumed to be one year. Consistent with most macroeconomic studies, we assume that $R = 1.04$, $\delta = 0.1$, and $\alpha = 0.3$. We set θ, which governs returns to scale, equal to 0.8. This value is on the lower end of the range of estimates recovered by Basu and Fernald (1997) using industry–level data. Using plant–level data, Lee (2005) finds that returns to scale in manufacturing vary from 0.828 to 0.91, depending on the estimator.

The remaining parameters are chosen in such a way that a number of statistics computed using a panel of simulated data are close to their empirical counterparts. Unless indicated otherwise, the simulated data is drawn from the stationary distribution. We list the simulated moments and their empirical counterparts in Table 2.

Because of non–linearities, it is not possible to match parameters to moments. However, the mechanics of the model clearly indicates which are the key parameters for each set of moments. What follows is a summary description of the algorithm that assigns values to parameters.

First, notice that there are uncountably many pairs (M, γ) which yield stationary equilibria identical to each other except for the volume of operating firms (and entrants). All the statistics of interest for our study are the same across all such allocations. To see why this is the case, start from a given stationary equilibrium and consider raising the labor supply elasticity γ. The original equilibrium wage will now elicit a greater supply of labor. Then, simply find the new, greater value for M, such that the aggregate demand for labor in stationary equilibrium equals supply at the original wage.

The property just described will not be true in the general model with $\sigma_z > 0$. We decide to borrow the value of γ from the calibration of that model, to be illustrated in Section 4. Given this choice, we set M in such a way that the equilibrium wage is 3.0. This value is clearly arbitrary. However – it will be clear at the end of this section – it involves no loss of generality.

Description	Symbol	Value
Capital share	α	0.3
Span of control	θ	0.8
Depreciation rate	δ	0.1
Interest rate	R	1.04
Labor supply elasticity	γ	2.0
Mass of potential entrants	M	1,766.29
Persistence idiosync. shock	ρ_s	0.55
Variance idiosync. shock	σ_s	0.22
Operating cost – mean parameter	μ_{c_f}	-5.63872
Operating cost – var parameter	σ_{c_f}	0.90277
Fixed cost of investment	c_0	0.00011
Variable cost of investment	c_1	0.03141
Pareto exponent	ξ	2.69
Entry cost	c_e	0.005347

Table 1: Parameter Values.

The parameters of the process driving the idiosyncratic shock, along with those governing the adjustment costs, were chosen to match the mean and standard deviation of the investment rate, the autocorrelation of investment, and the rate of inaction, i.e. the fraction of periods in which the investment rate is less than 1%. The targets are the moments computed by Cooper and Haltiwanger (2006) using a balanced panel from the LRD from 1972 to 1988. Because of selection in entry and exit, such statistics are likely to be biased estimates of the population moments. To ensure consistency between the simulated statistics and their empirical counterparts, we follow Cao (2007) and compute our moments on balanced panels extracted from the simulated series.

A simpler version of the neoclassical investment model with lognormal disturbances – one without investment adjustment costs – has the interesting properties that (i) the mean investment rate is a simple non linear function of of the parameters ρ_s and σ_s and that (ii) the standard deviation of the investment rate is a simple non–linear function of the mean. It follows that in that framework, mean and standard deviation do not identify the pair $\{\rho_s, \sigma_s\}$. While these properties do not hold exact in our model, inspection reveals that similar restrictions exist. We proceed to set $\sigma_s = 0.22$, a value consistent with available estimates of the standard deviation of the innovation to idiosyncratic productivity shocks,[4] and set ρ_s to minimize a weighted average of the distances of mean and standard deviation of investment rate from their targets.

The remaining parameters are set so that the model gets close to match the average entry rate and the relative size of entrants and exiters with respect to survivors, respectively. The definitions of these ratios are those of Dunne, Roberts, and Samuelson (1988) and the targets are the statistics obtained by Lee and Mukoyama (2012) using the LRD. The mean of the operating cost has the largest effect on the exit rate, which in stationary equilibrium must equal the entry rate. The variance of the operating cost pins down the relative size of exiters.

Finally, the pair $\{\xi, c_e\}$ determines the relative size of entrants. For simplicity, we set c_e equal to the mean operating cost, and then pick the value of the Pareto exponent to hit the target.

Statistic	Model	Data
Mean investment rate	0.153	0.122
Std. Dev. investment rate	0.325	0.337
Investment autocorrelation	0.059	0.058
Inaction rate	0.067	0.081
Entry rate	0.062	0.062
Entrants' relative size	0.58	0.60
Exiters' relative size	0.47	0.49

Table 2: Calibration Targets.

We claimed above that our arbitrary choice of M implies no loss of generality. How would the stationary equilibrium look like if, say, we chose a greater value? The wage rate would be higher and therefore all firm sizes would decline. Exit would be more appealing,

[4]See Haltiwanger (2011) and Castro, Clementi, and Lee (2013).

which in turn would imply higher exit (and entry) rate. However, firms dynamics will be exactly the same, as the investment process is invariant to the wage rate. It follows that by revising our choices for the parameters governing entry and exit, we can generate an economy that differs from ours only in its scale. All the moments of interest will be the same, and so will be the model's implications for entry, growth, and survival.

3.4 Firm Dynamics

In this section, we describe the model's implications for firm growth and survival, and compare them with the empirical evidence. Unless otherwise noted, size is measured by employment. The appropriate empirical counterpart of our notion of productivity is revenue total factor productivity (TFPR), as defined in Foster, Haltiwanger, and Syverson (2008).

The left panel of Figure 4 illustrates the unconditional relation between exit hazard rate and age. Consistent with Dunne, Roberts, and Samuelson (1989) and all other studies we are aware of, the exit hazard rate decreases with age. This is the case because on average entrants are less productive than incumbents. As a cohort ages – see the right panel – the survivors' productivity and value increase, leading to lower exit rates.

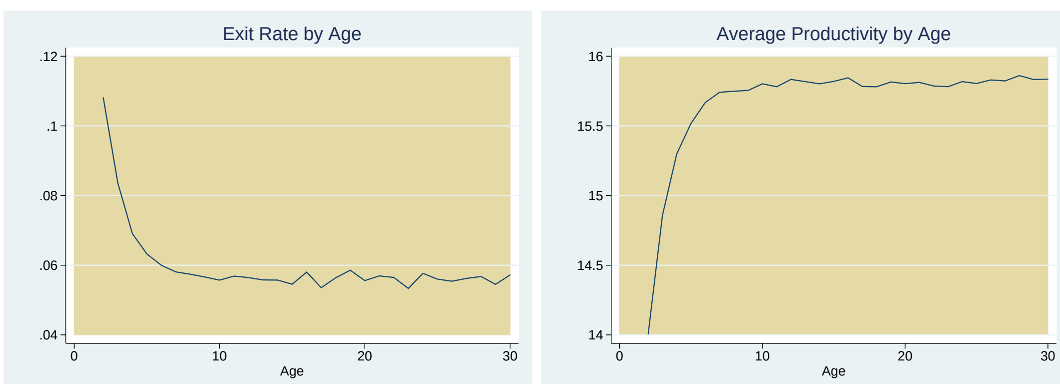

Figure 4: The exit hazard rate

A similar mechanism is at work in Hopenhayn (1992). In his framework, however, there exists a size threshold such that the exit rate is 100% for smaller firms and identically zero for larger firms. This feature is at odds with the evidence.[5] In our model, firms with the same employment are characterized by different combinations of (k, s) and therefore have different continuation values. Those with relatively low capital and relatively high productivity are less likely to exit.

[5]See Dunne, Roberts, and Samuelson (1989) and Lee and Mukoyama (2012).

Dunne, Roberts, and Samuelson (1989) also found that in the US manufacturing sector, establishment growth is unconditionally negatively correlated with both age and size – a finding that was confirmed for a variety of sectors and countries.[6] Evans (1987) and Hall (1987) found evidence that firm growth declines with size even when we condition on age, and viceversa.

Hopenhayn (1992) is consistent with these facts, with the exception of the conditional correlation between growth and age. In his model, idiosyncratic productivity is a sufficient statistics for firm size and growth. Conditional on age, smaller firms grow faster because the stochastic process is mean–reverting. However, firms of the same size behave identically, regardless of their age. The model generates the right unconditional relation between age and growth, simply because age and size are positively correlated in the stationary distribution. When controlling for size, age is uncorrelated with growth.

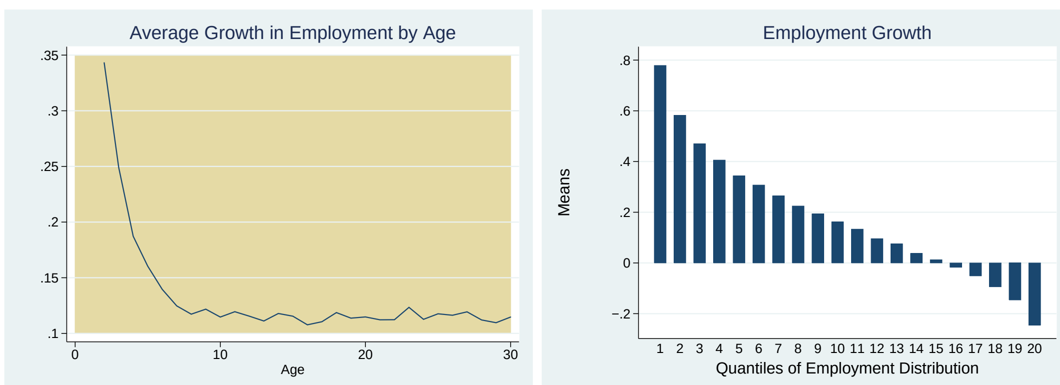

Figure 5: Unconditional Relationship between Growth, Age, and Size.

Our version of the model is consistent with all the facts about growth listed above. Figure 5 illustrates the unconditional correlations. Recall that the state variables are productivity and capital. Conditional on age, employment growth declines with size because larger firms tend to have higher productivity levels. Given that productivity is mean–reverting, their growth rates will be lower.

Now consider all the firms with the same employment. Since adjustment costs prevent the instantaneous adjustment of capital to the first–best size implied by productivity, some firms will be characterized by a relatively low capital and high shock, and others by a relatively high capital and low shock. The former will grow faster, because investment and employment are catching up with the optimal size dictated by productivity. The latter will shrink, as the scale of production is adjusted to the new, lower level of productivity.

[6]See Coad (2009) for a survey of the literature.

The conditional negative association between age and size arises because, on average, firms with relatively high k and low s (shrinking) will be older than firms with low k and high s (growing). For shrinking firms, productivity must have declined. For this to happen, they must have had the time to grow in the first place. On average, they will be older than those that share the same size, but are growing instead.

The model is consistent with the evidence on firm growth even when we measure size with capital rather than employment. Conditional on age, capital is negatively correlated with growth for the same reason that employment is. It is still the case that larger firms have higher productivity on average. Another mechanism contributes to generating the right conditional correlation between growth and size. Because of investment adjustment costs, same–productivity firms have different capital stocks. It turns out that on average the larger ones have seen their productivity decline, while the smaller ones have seen their efficiency rise. The former are in the process of shrinking, while the latter are growing.

We just argued that firms with the same capital will have different productivity levels. For given capital, firms with higher shocks are growing, while firms with lower shocks are shrinking. Once again, the negative conditional correlation between growth and age follows from the observation that, on average, firms with higher shocks are younger.

It is worth emphasizing that, no matter the definition of size, the conditional relation between age and growth is driven by relatively young firms. Age matters for growth even when conditioning on size, because it is (conditionally) negatively associated with productivity. To our knowledge, only two other papers present models that are consistent with this fact. The mechanism at work in D'Erasmo (2009) is similar to ours. Cooley and Quadrini (2001) obtain the result in a version of Hopenhayn (1992)'s model with financial frictions and exogenous exit.

The left panel of Figure 6 shows the firm size distribution that obtains in stationary equilibrium. Noticeably, it displays positive skewness. The right panel illustrates the evolution of a cohort size size distribution over time. Skewness declines as the cohort ages. Both of these features are consistent with the evidence gathered by Cabral and Mata (2003) from a comprehensive data set of Portuguese manufacturing firms.

4 Aggregate Fluctuations – Mechanics

We now move to the scenario with aggregate fluctuations. In order to formulate their choices, firms need to forecast next period's wage. The labor market clearing condition

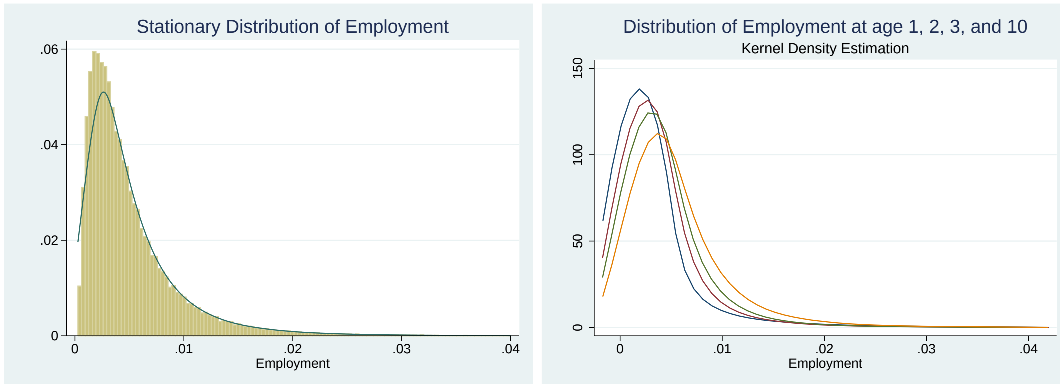

Figure 6: Evolution of a Cohort's Size Distribution

implies that the equilibrium wage at time t satisfies the following restriction:

$$\log w_t = \frac{\log[(1-\alpha)\theta z_t]}{1+\gamma[1-(1-\alpha)\theta]} + \frac{1-(1-\alpha)\theta}{1+\gamma[1-(1-\alpha)\theta]}\Omega_t, \tag{1}$$

with $\Omega_t = \log\left[\int \left(sk^{\alpha\theta}\right)^{\frac{1}{1-(1-\alpha)\theta}} d\Gamma_t(k,s)\right]$. The log–wage is an affine function of the logarithm of aggregate productivity and of a moment of the distribution.

Unfortunately, the dynamics of Ω_t depends on the evolution of Γ_t. It follows that the vector of state variables λ_t consists of the distribution Γ_t and the aggregate shock z_t. Faced with the formidable task of approximating an infinitely–dimensional object, we follow Krusell and Smith (1998) and conjecture that Ω_{t+1} is an affine function of Ω_t and $\log z_{t+1}$. Then, (1) implies the following law of motion for the equilibrium wage is

$$\log w_{t+1} = \beta_0 + \beta_1 \log w_t + \beta_2 \log z_{t+1} + \beta_3 \log z_t + \varepsilon_{t+1}. \tag{2}$$

When computing the numerical approximation of the equilibrium allocation, we will impose that firms form expectations about the evolution of the wage assuming that (2) holds true. This means that the aggregate state variables reduce to the pair (w_t, z_t). The parameters $\{\beta_0, \beta_1, \beta_2, \beta_3\}$ will be set equal to the values that maximize the accuracy of the prediction rule. The definition of accuracy and its assessment are discussed in Section 4.2. The algorithm is described in detail in Appendix A.

4.1 Calibration

With respect to the stationary case, we need to calibrate three more parameters. These are ρ_z and σ_z, which shape the dynamics of aggregate productivity, and the labor supply elasticity γ. We set them in order to match the standard deviation and auto–correlation of output growth, as well as the standard deviation of employment growth.

The targets for the first two are standard deviation and autocorrelation of the growth in non–farm private value added from 1947 to 2008, from the US Bureau of Economic Analysis. The third target is the standard deviation of employment growth, also in the non–farm private sector and for the same period, from the Bureau of Labor Statistics.

Description	Symbol	Value
Labor supply elasticity	γ	2.0
Persist. aggregate shock	ρ_z	0.685
Std. Dev. aggregate shock	σ_z	0.0163

Table 3: Parameter Values.

Our parameter values are listed in Table 3. Our choice for γ is close to recent estimates of the aggregate labor supply elasticity. For example, Rogerson and Wallenius (2009) conclude that for micro elasticities ranging from 0.05 to 1.25, the corresponding macro elasticities are in the range of 2.25 to 3.0. Fiorito and Zanella (2012) estimate a micro elasticity of 0.1 and macro elasticities in the range 1.1-1.7. At the quarterly frequency, our values for ρ_z and σ_z imply an autocorrelation of about 0.91 and a standard deviation of about 0.008. In comparison, Cooley and Prescott (1995) set the two parameters at 0.95 and 0.007, respectively.

Statistic	Model	Data
Standard deviation output growth	0.032	0.032
Autocorrelation output growth	0.069	0.063
Std. dev. employment growth (rel. to output growth)	0.656	0.667

Table 4: Additional Calibration Targets.

4.2 The Forecasting Rule

The forecasting rule for the equilibrium wage turns out to be

$$\log(w_{t+1}) = 0.38385 + 0.65105\log(w_t) + 0.53075\log(z_{t+1}) - 0.21508\log(z_t) + \varepsilon_{t+1}.$$

The wage is persistent and mean–reverting. A positive aggregate shock increases the demand for labor from both incumbents and entrants. This is why β_2, the coefficient of $\log(z_{t+1})$, is estimated to be positive. The coefficient of $\log(z_t)$ – β_3 – is negative because the larger the aggregate shock in the previous period, the smaller is going to be the expected increment in aggregate productivity, and therefore the lower the wage increase.

In the literature with heterogeneous agents and aggregate risk it has become standard to evaluate the accuracy of the forecasting rule by assessing the R^2 of the regression, which in our case is 0.9971. However, as pointed out by Den Haan (2010), this choice is questionable on at least three grounds. To start with, the R^2 considers predictions made conditional on wages generated by the true law of motion. In this sense, it only assesses the accuracy of one–period ahead forecasts. Second, the R^2 is an average. In the numerical methods literature, it is standard to report maximum errors instead. Last, but not least, the R^2 scales the error term by the variance of the dependent variable. The problem here is that it is often not clear what the appropriate scaling is. The root mean squared error (0.00103 in our case) does not suffer from the latter shortcoming, but is affected by the first two.

Here we follow Den Haan (2010)' suggestion to assess the accuracy of our forecasting rule by calculating the maximum discrepancy (in absolute value) between the actual wage and the wage generated by the rule *without* updating. That is, we compute the maximum pointwise difference between the sequence of actual market–clearing wages and that generated by our rule, when next period's predicted wage is conditional on last period's prediction for the current wage rather than the market clearing wage. The value of that statistics over 24,500 simulations is 0.765% of the market clearing wage.

The frequency distribution of percentage forecasting errors is illustrated in the left panel in Figure 7. The absolute value of the forecasting error is greater than 0.5% of the clearing wage in only 1.3% of the observations. The right panel is a scatter plot of equilibrium wages and their respective forecasts. It shows that the points are aligned along the 45^o line and that forecasting errors are small with respect to the variation in equilibrium wages. More diagnostics is reported in the Appendix.

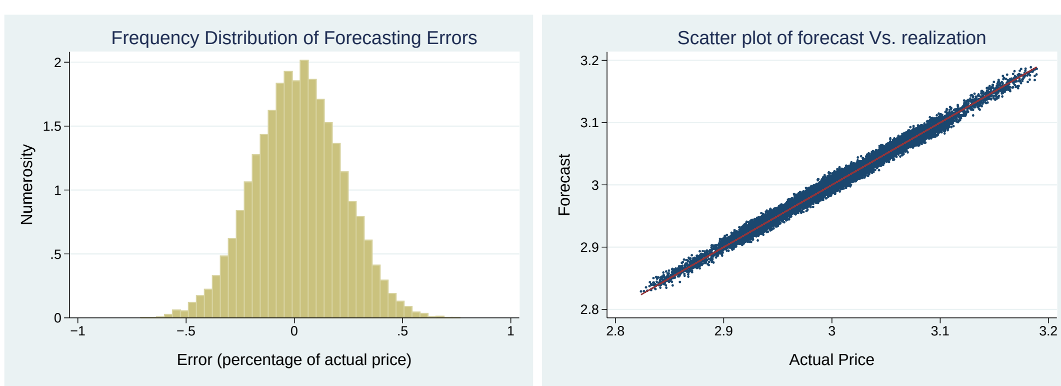

Figure 7: Accuracy of the Forecasting Rule.

4.3 Entry and Exit

Recall that in the stationary equilibrium analyzed in Section 3, the solution to the entry problem consists of a time–invariant threshold on the signal space and a policy function for investment. Here the threshold will be time-varying, and will depend on the wage and on the aggregate productivity realization.

The value of an incumbent is strictly increasing in aggregate productivity and strictly decreasing in the wage, while the cost of entry is constant. It follows that the entry threshold will be greater the lower is aggregate productivity and the greater is the wage. See the left panel in Figure 8 for an illustration.

Everything else equal, a rise in aggregate productivity (or a decline in the wage) will lead to an increase in the number of entrants. Given that the distribution of idiosyncratic shocks is stochastically increasing in the value of the signal, such a rise will also lead to a decline in entrants' average idiosyncratic efficiency.[7]

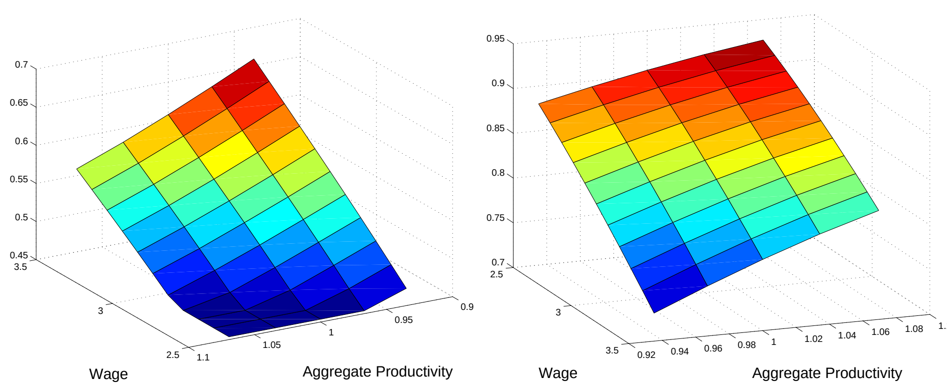

Figure 8: Left: Entry Threshold on the Signal Space. Right: Survival Probability.

The conditional probability of survival is also time–varying. Since the value recovered upon exit is not a function of aggregate variables, survival chances are greater the higher is aggregate productivity and the lower is the wage. The right panel of Figure 8 illustrates how the probability of continuing to operate changes with the two aggregate state variables, for a given pair of individual states $\{k, s\}$.

Everything equal, a positive shock to aggregate productivity leads to a decline in the number of exiting firms. Since the distribution of the idiosyncratic shock is invariant over time and across firms, the average idiosyncratic efficiency of exiters will also decline.

[7] For high levels of aggregate productivity and low wages the schedule on the left panel of figure 8 is flat. In that portion of the aggregate state space, the signal threshold coincides with its lower bound and all potential entrants find it optimal to start producing. It turns out, however, that this portion of the state space does not belong to the ergodic set.

The equilibrium dynamics of the entry and exit margins will obviously depend on the response of the wage to innovations in aggregate productivity. The wage and aggregate shock will be strongly positively correlated. The milder the response of the wage to positive aggregate productivity shocks, the higher the likelihood that periods of high aggregate TFP will be characterized by high entry volumes, high survival rates, and relatively low productivity of entrants and exiters.

5 Aggregate Fluctuations – Results

5.1 Cyclical Behavior of Entry and Exit

Table 5 reports the raw correlations of entry rate, exit rate, and the size of entrants and exiters (relative to incumbents) with industry output. Consistent with the evidence presented by Campbell (1998), the entry rate is pro–cyclical and the exit rate is counter–cyclical.

Entry Rate	Exit Rate	Entrants' Size	Exiters' Size
0.402	-0.779	-0.725	-0.892

Table 5: Correlations with industry output.

Interestingly, Campbell (1998) also provides evidence on the correlations between entry rate, exit rate, and future and lagged output growth. He finds that the correlation of entry with lagged output growth is greater than the contemporaneous correlation and that the exit rate is positively correlated with future output growth.

Our model is consistent with both features. The correlation of entry rate with contemporaneous output growth is 0.0039, while that with one–period lagged output growth is 0.8359. The reason is that the entry decision is taken contingent on the information available one period before the start of operations. The correlation coefficient between exit rate and one– and two–period ahead output growth are 0.13 and 0.26, respectively. Periods of low exit tend to be periods of high output. Given the mean–reverting nature of aggregate productivity, on average such periods are followed by times of low output growth.

Analyzing data from the LRD, Lee and Mukoyama (2012) find that selection at entry is quantitatively very important. Relative to incumbents, entering plants tend to be larger during recessions. This is the case, according to their evidence, because entrants' average idiosyncratic productivity is counter-cyclical. Our model shares these features of the data.

When the common productivity component is low, only firms with a relatively high level of idiosyncratic productivity find it worthwhile to enter.

The banking literature also found evidence in support of the claim that aggregate conditions have an impact on selection at entry. A number of papers, among which Cetorelli (2013), find that when credit market conditions are relatively favorable, entering firms are less productive on average.

In our model, the relative size of exiters is also higher during recessions. A drop in the common productivity component leads to a lower value of all incumbents. It follows that the marginal exiter will have a higher value of the idiosyncratic productivity component.

5.2 Impulse Responses

The objective of this section is to describe the impulse response functions in order to gain some more intuition about the model's dynamics. We initialize the system by assuming that the distribution of firms is equal to the point–wise mean of distributions on the ergodic set. The common productivity component is set at its mean value. At time $t = 1$, we impose that the exogenous aggregate productivity component rises by about 7% and we compute the evolution of the size distribution over the next 25 periods. We repeat this experiment for 3,000 times and depict the averages of selected variables in Figures 9 and 10.

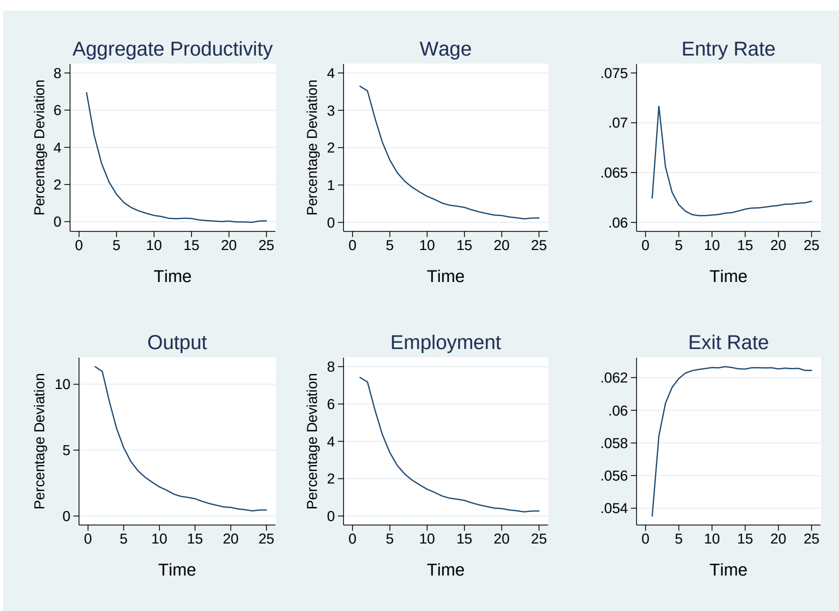

Figure 9: Response to a positive productivity shock.

The left–most panel on the top of Figure 9 reports the percentage deviation of the common productivity component from its unconditional mean. Not surprisingly, output, the wage rate and employment display similar dynamics.

Because of our timing assumptions, the industry composition does not depend upon contemporaneous shocks. Entry is governed by past aggregate conditions, and exit occurs after production has taken place. It follows that the contemporaneous response of output, employment, and wage to the positive productivity shock is entirely due to the expansion in hiring by incumbents.

The compositional effects of the aggregate shock appear in the next period. The exit rate declines, while the entry rate rises.[8] As a result, the number of operating firms also rises. It peaks in period 5, when the exit volume, which is rising back towards its unconditional mean, overcomes the declining entry volume.

Figure 10 illustrates the dynamics of selection in entry and exit. The average size and idiosyncratic productivity of entrants (relative to incumbents) declines and then converges back to its unconditional mean. This is the case because, on balance, the hike in aggregate productivity and the ensuing increase in the wage imply a rise in the value of incumbency. It follows that the average idiosyncratic productivity of the pool of entrants declines.

The average size and idiosyncratic productivity of exiters (relative to non-exiters) also declines. The improvement in aggregate factors implies an increase in the value of continuing operations. As a result, the marginal firms, i.e. those indifferent between exiting and staying, decline in quality. Since entry increases while exit declines, the cross–sectional mean of idiosyncratic productivity decreases.

Notice that the convergence of entry rate and exit rate to their respective unconditional means is not monotone. The exit rate overshoots its long–run value. The entry rate undershoots. Similarly, for all the statistics illustrated in Figure 10, except for the mass of operating firms. This happens because the wage decays at a lower pace than aggregate productivity.

A few periods after the positive shock hits, the aggregate component of productivity is back close to its unconditional mean. However, the wage is still relatively high. The reason is that the volume of entry is finite and firms are born relatively small. As the new entrants become more efficient, their labor demand increases, keeping the wage from falling faster. With a relatively high wage and relatively low aggregate productivity, the selection effect changes sign. Entry falls below its long–run value, while exit is higher than that.

[8] Following Dunne, Roberts, and Samuelson (1988), the period-t exit rate is the ratio of firms exiting between period t and $t+1$ and the number of operating firms in time t. This statistic drops in the period where the shocks hits, but the effect of exit on the distribution only appears in the next period.

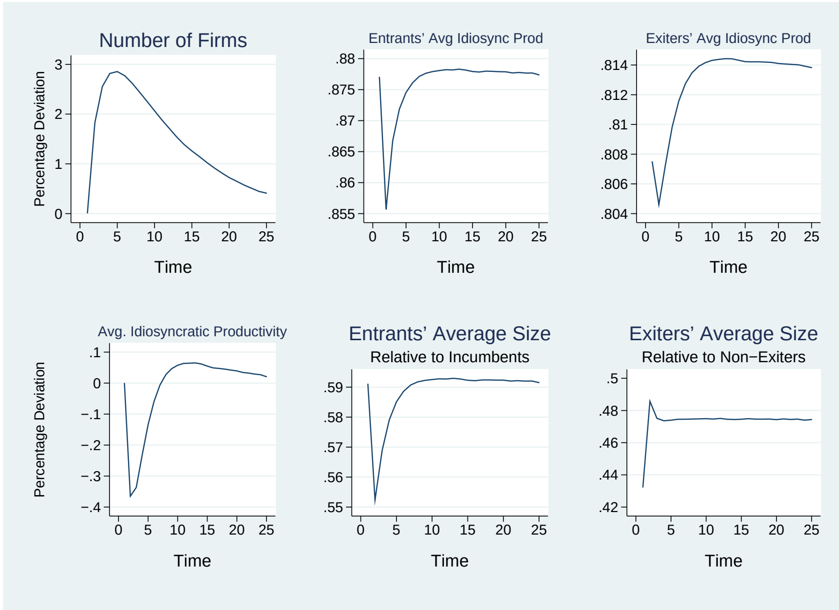

Figure 10: Response to a positive productivity shock.

6 The relevance of entry and exit for aggregate dynamics

In this section we show that allowing for entry and exit enhances the model's internal propagation mechanism. A corollary is that measuring the aggregate Solow residual as it is customary done in macroeconomics results in an upward bias in its persistence's estimate.

This is the outcome of two forces. One is the pro–cyclicality of the entry rate. The other is the fact that firms start out relatively unproductive, but quickly grow in size and efficiency. This dynamics is reflected in the contribution of net entry to aggregate productivity growth. As it is the case in the data, the contribution of new entrants is small in their first year, but grows over time.

Finally, we describe our model's implications for the cyclicality of selected cross–sectional moments.

6.1 Propagation

Think of the economy considered in Section 5, but abstract from entry and exit. At every point in time, there is a mass of firms whose technology is exactly as specified above. However, firms never exit. As our purpose is to compare such economy with our benchmark, assume that the number of operating firms is equal to the unconditional average number of incumbents that obtains along the benchmark's equilibrium path.

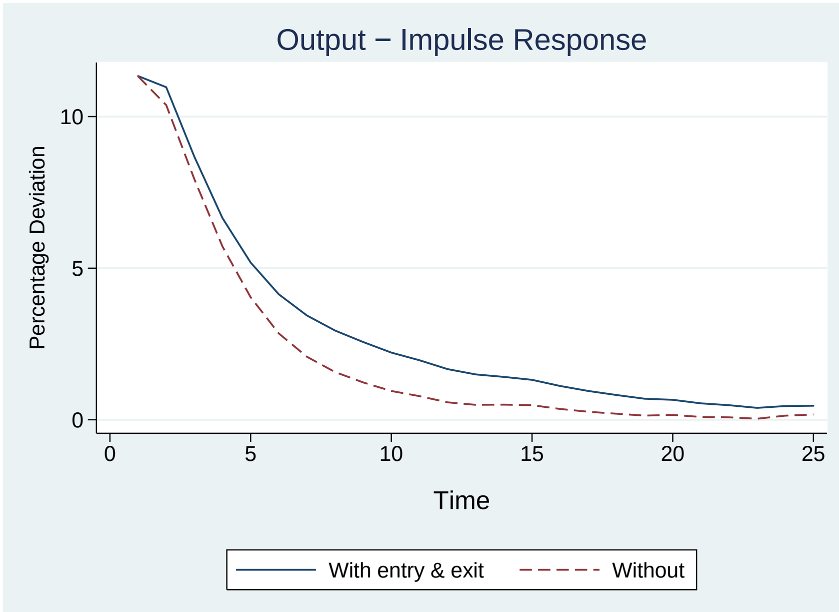

Figure 11: The effect of entry & exit on output dynamics.

Figure 11 depicts the impulse response of industry output to a positive aggregate shocks in our benchmark scenario and in the scenario without entry and exit, respectively.

We argued above that in the benchmark scenario the contemporaneous response of output is entirely due to incumbents. Therefore it is not surprising that the period–1 percentage deviation from the unconditional mean is the same across the two economies. In period 2, output is greater in the benchmark economy, due to the increase in net entry.

What we find particularly interesting is that the gap between the two time series keeps increasing beyond $t = 5$, the period when net entry becomes negative. This is due to the dynamics of firms born in the aftermath of the shock. Upon entry, these entities are very small and therefore account for a rather small fraction of total output. Over time, however, they grow in efficiency and size. This process takes place at the same time in which the aggregate productivity component regresses towards its unconditional mean. As a result, aggregate output falls at a slower pace.

The mean–reversion of output is slower when we allow for entry and exit. In other words, aggregate output is more persistent.

A confirmation that this mechanism is indeed at work comes from inspection of Figure 12, which illustrates the dynamics of the Solow residual in the two economies. The residual is computed by assuming an aggregate production function of the Cobb–Douglas form with capital share equal to 0.3. That is, we plot $\log(Y_t) - \alpha \log(K_t) - (1-\alpha)\log(L_t)$.

The Solow residual is uniformly higher in the benchmark scenario with entry and exit.

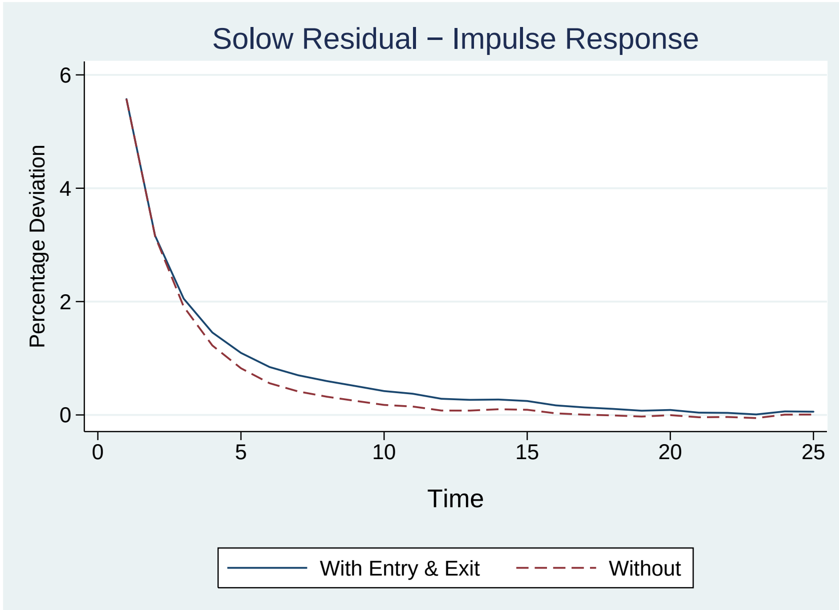

Figure 12: The effect of entry & exit on productivity dynamics.

The dynamics of the residual depends on the evolution of both z_t and the distribution of the idiosyncratic component s_t. In the benchmark economy, such distribution improves stochastically over time. In the case without entry and exit, instead, it is time–invariant.

This simple exercise hints that trying to infer information about the process generating factor–neutral exogenous technical change using a model without entry or exit will give the wrong answer. Such model will interpret changes in aggregate efficiency that results from the reallocation of output shares towards more efficient firms as changes in the exogenous aggregate component.

We also conducted the alternative experiment of setting the parameters in the economy without entry and exit in such a way that it generates the same values of the target moments as the benchmark economy. To generate the target autocorrelation for output growth, we had to set $\rho_z = 0.775$, much higher than the value of 0.685 assumed in Section 4.

6.2 Productivity Decomposition

Entry and exit enhance the model's propagation mechanism because entry is pro–cyclical and entrants' productivity grows faster than incumbents', drawing a reallocation of market share towards them. Strong evidence in support of the latter claim comes from the literature that exploits longitudinal establishment data to study the determinants of aggregate productivity growth.

After thoroughly reviewing that literature, Foster, Haltiwanger, and Krizan (2001) conclude that "*studies that focus on high–frequency variation tend to find a small contribution of net entry to aggregate productivity growth while studies over a longer horizon find a large role for net entry.*" They go on to add that "*Part of this is virtually by construction... Nevertheless, ... The gap between productivity of entering and exiting plants also increases in the horizon over which the changes are measured since a longer horizon yields greater differential from selection and learning effects.*" That is, conditional on survival entrants' productivity grows faster.

Not surprisingly, a productivity decomposition exercise carried out on simulated data generated by our model yielded results which are qualitatively consistent with the evidence illustrated by Foster, Haltiwanger, and Krizan (2001). On average, the contribution of net entry to productivity growth is positive, as entering firms tend to be more productive than the exiters they replace. Its magnitude is small when the interval between observations is one period (equivalent to one year), but it increases with the time between observations. In part, this is due to the mere fact that the output share accounted for by entrants is larger, the longer the horizon over which changes are measured. However, it is also due to the fact that entrants grow in size and productivity at a faster pace than incumbents.

Define total factor productivity as the weighted sum of firm–level Solow residuals, where the weights are the output shares. Let $\mathcal{C}_t$ denote the collection of plants active in both periods $t-k$ and t. The set $\mathcal{E}_t$ includes the plants that entered between the two dates and are still active at time t. In $\mathcal{X}_{t-k}$ are the firms that were active at time $t-k$, but exited before time t.

Following Haltiwanger (1997), the growth in TFP can be decomposed into five components, corresponding to the addenda in equation (3). They are known as (i) the within component, which measures the changes in productivity for continuing plants, (ii) the between–plant portion, which reflects the change in output shares across continuing plants, (iii) a covariance component, and finally (iv) entry and (v) exit components.[9]

$$\begin{aligned}\Delta \log(TFP_t) = &\sum_{i\in\mathcal{C}_t} \phi_{i,t-k}\Delta\log(TFP_{it}) + \sum_{i\in\mathcal{C}_t}[\log(TFP_{i,t-k}) - \log(TFP_{t-k})]\Delta\phi_{it} + \\ &\sum_{i\in\mathcal{C}_t}\Delta\log(TFP_{it})\Delta\phi_{it} + \sum_{i\in\mathcal{E}_t}[\log(TFP_{it}) - \log(TFP_{t-k})]\phi_{it} - \\ &\sum_{i\in\mathcal{X}_{t-k}}[\log(TFP_{i,t-k}) - \log(TFP_{t-k})]\phi_{i,t-k} \qquad (3)\end{aligned}$$

Table 6 reports the results that obtain when we set k equal to 1 and 5, respectively.

[9] With ϕ_{it} and TFP_{it} we denote firm i's output share and measured total factor productivity at time t, respectively. TFP_t is the weighted average total factor productivity across all firms active at time t.

In the last column, labeled Net Entry, we report the difference between the entry and exit components. Recall that in our model the unconditional mean of aggregate productivity growth is identically zero.

k	Within	Between	Covariance	Entry	Exit	Net Entry
1	-8.7778	-4.1623	12.4366	-0.2964	-0.8011	0.5047
5	-15.1437	-13.3884	27.5833	-0.7556	-1.7104	0.9548

Table 6: Productivity Decomposition (percentages).

The between and within components are necessarily negative, because of mean reversion in the process driving idiosyncratic productivity. Larger firms, which tend to be more productive, shrink on average. Smaller firms, on the contrary, tend to grow. The covariance component is positive, because firms that become more productive also increase in size.

On average, both the entry and exit contributions are negative. This reflects the fact that both entrants and exiters are less productive than average. However, entrants tend to be more productive than exiters. The contribution of net entry to productivity growth is positive regardless of the horizon.

What's most relevant for our analysis is that for $k = 5$ the contribution of net entry is about twice that for $k = 1$. In part, this is due to the fact that the share of output produced by entrants increases with k. However, this cannot be the whole story. The contribution of entry is roughly -0.3% for $k = 1$ and goes to -0.76% for $k = 5$. If entrants' productivity did not grow faster than average, the contribution of entry for the $k = 5$ horizon would be much smaller.

6.3 Cyclical Variation of Cross–sectional Moments

Starting with Eisfeldt and Rampini (2006), a number of papers have assessed the cyclical behavior of the cross–sectional distribution of productivity. Particular interest has been paid to the evolution of the second moment.

Exploiting the German dataset USTAN and data from the US Census of Manufacturing, respectively, Bachman and Bayer (2013) and Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2012) conclude that the standard deviation of idiosyncratic TFP growth is countercyclical. Kehrig (2011), also analyzing US Census data, reports that a measure of dispersion of the level of TFP is countercyclical as well.

In spite of our assumption that the stochastic process governing idiosyncratic productivity is constant across time and across firms, the equilibrium allocation of our model features non–trivial dynamics of the cross–sectional distribution of productivity. The reason, as it is by now clear, is that the net entry rate, as well as selection in entry and exit, are time–varying.

The impulse response functions discussed in section 5 hint that when industry output is high, the idiosyncratic productivity of both entrants and exiters is relatively low. Since the entry-rate is pro–cyclical while the exit–rate is counter–cyclical, it follows that the first moment of the idiosyncratic productivity distribution is counter-cyclical.

Another robust implication of the model is that skewness is pro–cyclical. In Section 3 we showed that, consistent with the evidence, the productivity distribution is positively skewed. During expansions, both entrants and exiters are less productive than average. Since the net entry rate is pro–cyclical, skewness increases.

This feature is illustrated in Figure 13, which plots the difference between the (average) fraction of firms associated with each level of idiosyncratic shock in expansion and in recession, respectively. When industry output is high, there are relatively more low–productivity firms and relatively less high–productivity firms. The right tail of the distribution is thinner.

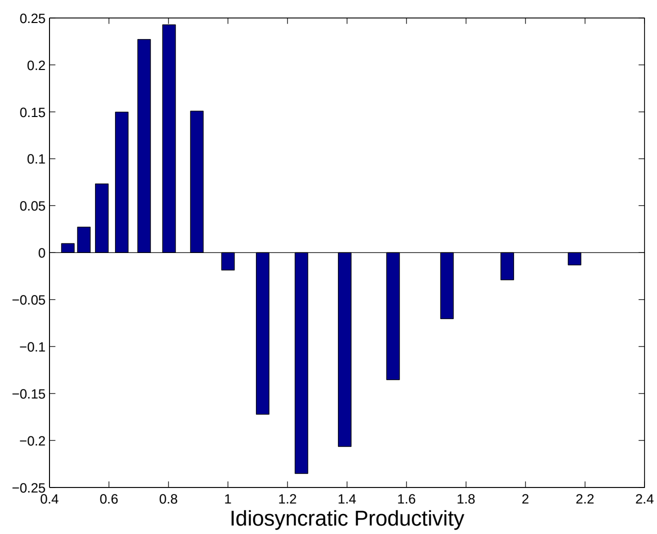

Figure 13: Change in the Cross-Sectional Distribution.

Unfortunately the prediction for the cyclicality of the second moment is not robust. Our simulations show that the sign of the correlation between second moment and industry output changes with relatively minor changes in parameter values.

7 Conclusion

This paper provides a framework to study the dynamics of the cross–section of firms and its implications for aggregate dynamics. When calibrated to match a set of moments of the investment process, our model delivers implications for firm dynamics and for the cyclicality of entry and exit that are consistent with the evidence.

The survival rate increases with size. The growth rate of employment is decreasing with size and age, both unconditionally and conditionally. The size distribution of firms is skewed to the right. When tracking the size distribution over the life a cohort, the skewness declines with age.

The entry rate is positively correlated with current and lagged output growth. The exit rate is negatively correlated with output growth and positively associated with future growth.

We show that allowing for entry and exit enhances the internal propagation mechanism of the model. This obtains because the entry rate is pro–cyclical and recent entrants grow faster than incumbents.

A positive shock to aggregate productivity leads to an in increase in entry. Consistent with the empirical evidence, the new entrants are smaller and less efficient than incumbents. The skewness of the distribution of firms over the idiosyncratic productivity component increases. As the exogenous component of aggregate productivity declines towards its unconditional mean, the new entrants that survive grow in productivity and size. That is, the distribution of idiosyncratic productivity improves.

For a version of our model without entry or exit to generate a data–conforming persistence of output, the first–order autocorrelation of aggregate productivity shocks must be 0.775. In the benchmark scenario with entry and exit, it needs only be 0.685.

A Numerical Approximation

Our algorithm consists of the following steps.

1. Guess values for the parameters of the wage forecasting rule $\hat{\beta} = \left\{\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3\right\}$;

2. Approximate the value function of the incumbent firm;

3. Simulate the economy for T periods, starting from an arbitrary initial condition (z_0, Γ_0);

4. Obtain a new guess for $\hat{\beta}$ by running regression (2) over the time–series $\{w_t, z_t\}_{t=S+1}^{T}$, where S is the number of observations to be scrapped because the dynamical system has not reached its ergodic set yet;

5. If the new guess for $\hat{\beta}$ is close to the previous one, stop. If not, go back to step 2.

A.1 Approximation of the Value Function

The incumbent's value function is approximated by value function iteration.

1. Start by defining grids for the state variables w, z, k, s. Denote them as Ψ_w, Ψ_z, Ψ_k, and Ψ_s, respectively. The wage grid is equally spaced and centered around the equilibrium wage of the stationary economy. The capital grid is constructed following the method suggested by McGrattan (1999). The grids and transition matrices for the two shocks are constructed following Tauchen (1986). For all pairs (s, s') such that $s, s' \in \Psi_s$, let $H(s'|s)$ denote the probability that next period's idiosyncratic shock equals s', conditional on today's being s. For all (z, z') such that $z, z' \in \Psi_z$, let also $G(z'|z)$ denote the probability that next period's aggregate shock equals z', conditional on today's being z.

2. For all triplets (w, z, z') on the grid, the forecasting rule yields a wage forecast for the next period (tildes denote elements not on the grid):

$$\log(\tilde{w}') = \hat{\beta}_0 + \hat{\beta}_1 \log(w) + \hat{\beta}_2 \log(z') + \hat{\beta}_3 \log(z).$$

   In general, $\tilde{w}'$ will not belong to the grid of wages. There will be contiguous grid points (w_i, w_{i+1}) such that $w_i \leq \tilde{w}' \leq w_{i+1}$. Now let $J(w_i|w, z, z') = 1 - \frac{\tilde{w}' - w_i}{w_{i+1} - w_i}$, $J(w_{i+1}|w, z, z') = \frac{\tilde{w}' - w_i}{w_{i+1} - w_i}$, and $J(w_j|w, z, z') = 0$ for all j such that $j \neq i$ and $j \neq i + 1$. This allows us to evaluate the value function for values of the wage which are off the grid by linear interpolation;

3. For all grid elements (w, z, k, s), guess values for the value function $V_0(w, z, k, s)$;

4. The revised guess of the value function, $V_1(w, z, k, s)$, is determined as follows:

$$\begin{aligned} V_1(w,z,k,s) =& \pi(w,z,k,s) + Pr[c_f > c_f^*(w,z,k,s)]V_x(k) \\ &+ Pr[c_f \le c_f^*(w,z,k,s)]\Big[\widetilde{V}(w,z,k,s) - E[c_f | c_f \le c_f^*(w,z,k,s)]\Big] \end{aligned}$$

subject to

$$\begin{aligned} \pi(w,z,k,s) &= \frac{1-(1-\alpha)\theta}{(1-\alpha)\theta} w^{-\frac{\theta(1-\alpha)}{1-\theta(1-\alpha)}} [(1-\alpha)\theta s z k^{\alpha\theta}]^{\frac{1}{1-\theta(1-\alpha)}}, \\ V_x(k) &= k(1-\delta) - g[-k(1-\delta), k], \\ \widetilde{V}(w,z,k,s) &= \max_{k' \in \Psi_k} \Big\{ -x - c_0 k\chi - c_1 \left(\frac{x}{k}\right)^2 k \\ &\quad + \frac{1}{R}\sum_j \sum_i \sum_n V_0(w_i, z_j, k', s_n) H(s_n|s) J(w_i|w, z, z_j) G(z_j|z) \Big\}, \\ x &= k' - k(1-\delta), \\ \chi &= 1 \text{ if } k' \neq k \text{ and } \chi = 0 \text{ otherwise}, \end{aligned}$$

where $c_f^*(w, z, k, s)$ is the value of the fixed operating cost that makes the firm indifferent between the choice of exiting and obtaining the undepreciated portion of its capital, net of the adjustment cost, or the choice of continuing and investing the optimal amount of capital.

5. Keep on iterating until $\sup \left| \frac{V_{t+1}(w,z,k,s) - V_t(w,z,k,s)}{V_t(w,z,k,s)} \right| < 10.0^{-6}$. Denote the latest value function as $V_\infty(w, z, k, s)$.

A.2 Entry

1. Define a grid for the signal. Denote it as Ψ_q. Let also $W(s_n|q)$ indicate the probability that the first draw of the idiosyncratic shock is s_n, conditional on the signal being q.

2. For all triplets (w, z, q) on the grid, compute the value of entering as

$$V_e(w,z,q) = \max_{k' \in \Psi_k} -k' + \frac{1}{R}\sum_j \sum_i \sum_n V_\infty(w_i, z_j, k', s_n) W(s_n|q) J(w_i|w, z, z_j) G(z_j|z).$$

3. For all grid points z, construct a bi-dimensional cubic spline interpolation of $V_e(w, z, q)$ over the dimensions (w, q). For all pairs $\tilde{w}, \tilde{q}$, denote the value of entering as $\tilde{V}_e(\tilde{w}, z, \tilde{q})$.

4. Define $\tilde{q}_e(\tilde{w}, z)$ as the value of the signal which makes prospective entrants indifferent between entering and not. That is, $\tilde{V}_e(\tilde{w}, z, \tilde{q}_e(\tilde{w}, z)) = c_e$.

A.3 Simulation

1. Given the current firm distribution Γ_t and aggregate shock z_t, compute the equilibrium wage $\tilde{w}_t$ by equating the labor supply equation $L^s(w) = w^\gamma$ to the labor demand equation

$$L_t^d(w) = \left(\frac{z_t\theta(1-\alpha)}{w}\right)\sum_m\sum_n\left[s_n k_m^{\alpha\theta}\right]^{\frac{1}{1-\theta(1-\alpha)}}\Gamma_t(s_n, k_m).$$

2. For all $z' \in \Psi_z$, compute the conditional wage forecast $\tilde{w}_{t+1}(z')$ as follows:

$$\log[\tilde{w}_{t+1}(z')] = \hat{\beta}_0 + \hat{\beta}_1\log(\tilde{w}_t) + \hat{\beta}_2\log(z') + \hat{\beta}_3\log(z_t).$$

For every z', there will be contiguous grid points (w_i, w_{i+1}) such that $w_i \le \tilde{w}_{t+1}(z') \le w_{i+1}$. Now let $J_{t+1}(w_i|z') = 1 - \frac{\tilde{w}_{t+1}(z')-w_i}{w_{i+1}-w_i}$, $J_{t+1}(w_i|z') = \frac{\tilde{w}_{t+1}(z')-w_i}{w_{i+1}-w_i}$, and $J_{t+1}(w_j|z') = 0$ for all j such that $j \neq i$ and $j \neq i+1$;

3. For all pairs (k, s) on the grid such that $\Gamma_t(k, s) > 0$, the optimal choice of capital $k'(\tilde{w}_t, z_t, k, s)$ is the solution to the following problem:

$$\begin{aligned}
\max_{k'\in\Psi_k}\ & \pi(\tilde{w}_t, z_t, k, s) - x - c_0 k\chi - c_1\left(\frac{x}{k}\right)^2 k \\
& + \frac{1}{R}\sum_j\sum_i\sum_n V_\infty(w_i, z_j, k', s_n)H(s_n|s)J_t(w_i|z_j)G(z_j|z_t), \\
& \text{s.t. } x = k' - k(1-\delta), \\
& \pi(\tilde{w}_t, z_t, k, s) = \frac{1-(1-\alpha)\theta}{(1-\alpha)\theta}\tilde{w}_t^{-\frac{\theta(1-\alpha)}{1-\theta(1-\alpha)}}\left[(1-\alpha)\theta s z_t k^{\alpha\theta}\right]^{\frac{1}{1-\theta(1-\alpha)}}, \\
& \chi = 1 \text{ if } k' \neq k \text{ and } \chi = 0 \text{ otherwise}.
\end{aligned}$$

4. There will be contiguous elements of the signal grid (q^*, q^{**}) such that $q^* \le \tilde{q}_e(\tilde{w}_t, z_t) \le q^{**}$.

   - For all $q \ge q^{**}$, the initial capital of entrants $k_e'(\tilde{w}_t, z_t, q)$ solves the following problem:

$$\max_{k_e'\in\Psi_k} -k_m + \frac{1}{R}\sum_j\sum_i\sum_n V_\infty(w_i, z_j, k_e', s_n)W(s_n|q)J_t(w_i|z_j)G(z_j|z_t).$$

   - We can easily compute the distribution of the idiosyncratic shock conditional on $\tilde{q}_e \equiv \tilde{q}_e(\tilde{w}_t, z_t)$, denoted as $\tilde{W}(s_n|\tilde{q}_e)$ and then compute the optimal capital $k_e'(\tilde{w}_t, z_t, \tilde{q}_e)$ as the solution to:

$$\max_{k_e'\in\Psi_k} -k_m + \frac{1}{R}\sum_j\sum_i\sum_n V_\infty(w_i, z_j, k_e', s_n)\tilde{W}(s_n|\tilde{q}_e)J_t(w_i|z_j)G(z_j|z_t).$$

5. Draw the aggregate productivity shock z_{t+1};

6. Determine the distribution at time $t+1$. For all (k, s) such that $V_\infty(\tilde{w}_{t+1}(z_{t+1}), z_{t+1}, k, s) = 0$, then $\Gamma_{t+1} = 0$. For all other pairs,

$$\Gamma_{t+1}(k, s) = \sum_m \sum_n \Gamma_t(k_m, s_n) H(s|s_n) \Upsilon_{m,n}(w_t, z_t, k) + \mathcal{E}_{t+1}(k, s),$$

where

$$\Upsilon_{m,n}(w_t, z_t, k) = \begin{cases} 1 & \text{if } k'(w_t, z_t, k_m, s_n) = k \\ 0 & \text{otherwise.} \end{cases}$$

and $\mathcal{E}_{t+1}(k, s) = M \sum_{i:q_i \geq q^{**}} H(s|q_i) Q(q_i) \Xi_{m,i}(w_t, z_t, k) + M \tilde{H}(s|q^e) \tilde{Q}(q^e) \tilde{\Xi}_{m,i}(w_t, z_t, k)$ where

$$\Xi_{m,i}(w_t, z_t) = \begin{cases} 1 & \text{if } k'(w_t, z_t, q_i) = k \\ 0 & \text{otherwise.} \end{cases}$$

A.4 More Accuracy Tests of the Forecasting Rule

Forecasting errors are essentially unbiased – the mean error is -0.00017% of the forecasting price – and uncorrelated with the price (the correlation coefficient between the two series is -0.0456) and the aggregate shock (-0.0045).

Figure 14: Accuracy of the Forecasting Rule.

In the left panel of Figure 14 is the scatter plot of the forecasting errors against the market clearing price. In the right panel is the time series of the forecasting error. The good news is that errors do not cumulate.

References

Bachman, R., and C. Bayer (2013): "Wait–and–see Business Cycles?," *Journal of Monetary Economics*, Forthcoming.

Basu, S., and J. Fernald (1997): "Returns to Scale in US Production: Estimates and Implications," *Journal of Political Economy*, 105(2), 249–283.

Bilbiie, F., F. Ghironi, and M. Melitz (2012): "Endogenous Entry, Product Variety, and Business Cycles," *Journal of Political Economy*, 120(2), 304–345.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. Terry (2012): "Really Uncertain Business Cycles," NBER Working Paper # 18245.

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

Campbell, J. (1998): "Entry, Exit, Embodied Technology, and Business Cycles," *Review of Economic Dynamics*, 1, 371–408.

Cao, S. (2007): "Three Essays on Capital Adjustment, Reallocation and Aggregate Productivity," PhD Dissertation, University of Texas at Austin.

Castro, R., G. L. Clementi, and Y. Lee (2013): "Cross-Sectoral Variation in The Volatility of Plant-Level Idiosyncratic Shocks," NBER Working Paper #17659.

Cetorelli, N. (2013): "Surviving Credit Market Competition," *Economic Inquiry*, Forthcoming.

Chatterjee, S., and R. Cooper (1993): "Entry and Exit, Product Variety and the Business Cycle," NBER Working Paper # 4562.

Coad, A. (2009): *The Growth of Firms. A Survey of Theories and Empirical Evidence.* Edward Elgar Publishing.

Cogley, T., and J. Nason (1995): "Output Dynamics in Real Business Cycle Models," *American Economic Review*, 85(3), 492–511.

Cooley, T., R. Marimon, and V. Quadrini (2004): "Aggregate Consequences of Limited Contract Enforceability," *Journal of Political Economy*, 112(4), 817–847.

Cooley, T., and E. Prescott (1995): "Economic Growth and Business Cycles," in *Frontiers of Business Cycle Research*, ed. by T. Cooley, pp. 1–38. Princeton University Press.

Cooley, T., and V. Quadrini (2001): "Financial Markets and Firm Dynamics," *American Economic Review*, 91(5), 1286–1310.

Cooper, R., and J. Haltiwanger (2006): "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, 73(3), 611–633.

Den Haan, W. J. (2010): "Assessing the Accuracy of the Aggregate Law of Motion in Models with Heterogeneous Agents," *Journal of Economic Dynamics and Control*, 34, 79–99.

D'Erasmo, P. (2009): "Investment and Firm Dynamics," Working Paper, University of Maryland.

Devereux, M., A. Head, and B. Lapham (1996): "Aggregate Fluctuations with Increasing Returns to Specialization and Scale," *Journal of Economic Dynamics and Control*, 20, 627–656.

Dunne, T., M. Roberts, and L. Samuelson (1988): "Patterns of Firm Entry and Exit in US Manufacturing Industries," *Rand Journal of Economics*, 19, 495–515.

——— (1989): "The Growth and Failure of US Manufacturing Plants," *Quarterly Journal of Economics*, 94, 671–698.

Eisfeldt, A., and A. Rampini (2006): "Capital Reallocation and Liquidity," *Journal of Monetary Economics*, 53, 369–399.

Evans, D. (1987): "Tests of Alternative Theories of Firm Growth," *Journal of Political Economy*, 95(4), 657–674.

Fiorito, R., and G. Zanella (2012): "The Anatomy of the Aggregate Labor Supply Elasticity," *Review of Economic Dynamics*, 15, 171–187.

Foster, L., J. Haltiwanger, and C. Krizan (2001): "Aggregate Productivity Growth: Lessons from Microeconomic Evidence," in *New Developments in Productivity Analysis*, ed. by C. Hulten, E. Dean, and M. Harper, pp. 303–363. University of Chicago Press.

Foster, L., J. Haltiwanger, and C. Syverson (2008): "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity of Profitability?," *American Economic Review*, 98(1), 394–425.

——— (2012): "The Slow Growth of New Plants: Learning about Demand?," NBER Working paper # 17853.

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

Haltiwanger, J. (1997): "Measuring and Analyzing Aggregate Fluctuations: The Importance of Building from Microeconomic Evidence," *Federal Reserve Bank of St. Louis Review*, 79, 55–77.

——— (2011): "Firm Dynamics and Productivity Growth," *European Investment Bank Papers*, 16(1), 116–136.

Hopenhayn, H. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60(5), 1127–1150.

Kehrig, M. (2011): "The Cyclicality of Productivity Dispersion," Working Paper, University of Texas at Austin.

Khan, A., and J. K. Thomas (2003): "Nonconvex Factor Adjustments in Equilibrium Business Cycle Models: Do Nonlinearities Matter?," *Journal of Monetary Economics*, 50, 331–360.

——— (2008): "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, 76(2), 395–436.

Krusell, P., and A. Smith (1998): "Income and wealth heterogeneity in the macroeconomy," *Journal of Political Economy*, 106, 867–896.

Lee, Y. (2005): "The Importance of Reallocations in Cyclical Productivity and Returns to Scale: Evidence from Plant-level Data," Working paper # 05–09, Federal Reserve Bank of Cleveland.

Lee, Y., and T. Mukoyama (2012): "Entry, Exit, and Plant–level Dynamics over the Business Cycle," Working paper, University of Virginia.

McGrattan, E. (1999): "Application of weighted residual methods to dynamic economic models," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon, and A. Scott. Oxford University Press.

Rogerson, R., and J. Wallenius (2009): "Micro and Macro Elasticities in a Life Cycle Model with Taxes," *Journal of Economic Theory*, 144, 2277–2292.

Samaniego, R. (2008): "Entry, Exit and Business Cycle in a General Equilibrium Model," *Review of Economic Dynamics*, 11, 529–541.

Tauchen, G. (1986): "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 20, 177–181.

Veracierto, M. (2002): "Plant–Level Irreversible Investment and Equilibrium Business Cycles," *American Economic Review*, 92(1), 181–197.

Zhang, L. (2005): "The Value Premium," *Journal of Finance*, 60(1), 67–103.