Firm Dynamics and the Granular Hypothesis*

Vasco M. Carvalho† Basile Grassi‡

March, 21, 2014

[Preliminary. Link the the latest version]

Abstract

Building on the standard firm dynamics setup of Hopenhayn (1992), we develop a quantitative theory of aggregate fluctuations arising from idiosyncratic shocks to firm level productivity. This allows us to generalize the theoretical results in Gabaix (2011) to account for persistent micro-level shocks, optimal size decisions as well as endogenous firm entry and exit. We then use our model to provide a quantitative evaluation of Gabaix's "granular hypothesis" and find that it yields aggregate fluctuations of the same order of magnitude as a standard representative-firm real business cycle model. A calibration of our model to the US economy with a large number of firms leads to sizable aggregate fluctuations: the standard deviation of aggregate TFP (respectively output) is 0.8% (respectively 1.7%). We use this calibration to explore firms' comovement over the business cycle. The model predicts that the differential growth between large and small firms is pro-cyclical as it is in the data.

Keywords: Firm Dynamics; Granular Hypothesis; Firm Size Distribution; Aggregate Fluctuations

*The authors would like to acknowledge helpful comments from Axelle Arquié, Florin Bilbiie, Gian Luca Clementi, Antoine Coste, Jan Eeckhout, Xavier Gabaix, Wouter den Haan, Hugo Hopenhayn, Jean Imbs, Boyan Jovanovic, Morten Ravn, Edouard Schaal and Rob Shimer. The authors thank seminar participants at CREi, CREST-INSEE, Essex, LSE, NYU, PSE, UCL and the Conference in Honor of Lars Hansen at the University of Chicago. Part of this work was made during Basile Grassi's stay at the NYU Economics department and he is grateful for their hospitality. He also thanks the Fulbright commission for financial support.

†University of Cambridge and Centre de Recerca en Economia Internacional, Universitat Pompeu Fabra and Barcelona GSE. vcarvalho@crei.cat.

‡Paris School of Economics, Université Paris I Panthéon Sorbonne, and Centre de Recherche en Économie et STatistique. basile.grassi@gmail.com.

1 Introduction

When large firms represent a disproportionately large share of the economy, aggregate fluctuations may arise from idiosyncratic shocks to these firms. This hypothesis – to which we refer as the "granular hypothesis" as proposed in Gabaix (2011) - opens the possibility of doing away with aggregate shocks, instead tracing back the origins of aggregate fluctuations to micro shocks hitting a small number of large firms. While this hypothesis has proven influential in the literature, it remains largely an accounting result - concerning the aggregation of firm level variability in static environments – whose relation to the extant theory of firm dynamics is left unspecified. Partly as a result, the granular hypothesis remains a possibility result whose quantitative relevance has not yet been established.

In this paper we cast the granular hypothesis in a standard firm dynamics setting. Building on the setup of Hopenhayn (1992), we develop a quantitative theory of aggregate fluctuations arising only from idiosyncratic shocks to firm level productivity. This allows us to generalize the theoretical results in Gabaix (2011) to account for persistent micro-level shocks, optimal size decisions as well as endogenous firm entry and exit. Further, we provide a quantitative evaluation of this hypothesis and find that it yields aggregate fluctuations of the same order of magnitude as a standard representative-firm real business cycle model. A calibration of our model to the US economy with a large number of firms leads to sizable aggregate fluctuations: the standard deviation of aggregate TFP (respectively output) is 0.8% (respectively 1.7%).

Our setup follows Hopenhayn (1992) closely. Firms differ in their idiosyncratic productivity level, which here is assumed to follow a discrete Markovian process. Incumbents have access to a decreasing returns to scale technology using labor as the only input. They produce a unique good in a perfectly competitive market. They face an operating cost at each period, which in turn generates endogenous exit. The crucial difference relative to Hopenhayn (1992)- and much of the large literature that follows from it - is that we do not rely on the traditional "continuum of firms" assumption in order to characterize the law of motion for the firm size distribution. Instead, we characterize the law of motion for any finite number of firms and show that, generically, the distribution of firms across productivity levels is time-varying in a stochastic fashion. As is well known, this distribution is the aggregate state variable in this class of models.

An immediate implication of out findings is therefore that aggregate productivity, aggregate output and factor prices are themselves stochastic. In a nutshell, we show that the standard workhorse model in the firm dynamics literature – once the assumption regarding a continuum of firms is dropped – already contains in it a theory of the aggregate fluctuations.

We then show that whether standard firm dynamics models can deliver a quantitatively relevant theory of the business cycle depends on the rate at which aggregate fluctuations vanish as a function of the number of firms. We offer a characterization of this decay rate under parametric conditions on the firm size distribution. In particular, we show analytically that, as the number of firms increases, the rate at which aggregate volatility decays is slower than what a simple central limit argument would predict. This result relies on the fat-tailedness of the firm size distribution as in Gabaix (2011). Therefore, our results generalize the latter by taking into account the persistence of the idiosyncratic shock process, firm entry, exit and optimal size decisions. In particular, we show that a low persistence of the idiosyncratic productivity process increases the rate of convergence of the aggregate volatility.

As mentioned above, quantitatively we are able to generate aggregate fluctuations of the same magnitude as the standard real business cylce model but arising from idiosyncratic firm-level shocks only. A calibration of our model to the US economy with a large number of firms leads to sizable aggregate fluctuations: the standard deviation of aggregate TFP (respectively output) is 0.8% (respectively 1.7%). Our framework also allows us to study the macro and micro dynamic effects of a negative productivity shock on the largest firm. Such a shock is contractionary at the aggregate level. By construction, the decrease in the productivity of the largest firm induces a smaller aggregate productivity, since the latter is shown to be an average of the idiosyncratic productivities. However, this shock induces the largest firm to reduce its size and to cut its labor demand. The latter pushes the wage down. Other firms benefit from this reduction in their labor cost. On average, their productivity remains the same while their cost decreases, so they expand. Overall, a negative shock on the largest firm is contractionary at the aggregate level but expansionary for the typical firm in the economy.

The upshot of this theory is that the business cycle, in our setting, is led by the large firms dynamics. In particular, if a small number of large firms face higher productivity than usual, this will lead to an aggregate expansion. To the econometrician, it would thus seem as if large firms are more cyclically sensitive, in that they co-move more with the aggregate echoing the recent findings in Moscarini and Postel-Vinay (2012). The latter find that the correlation between the differential net growth rate of large versus small firms is negatively correlated with the unemployment rate. We find a similar result on a simulated representative panel drawn from our calibrated model. Though our results are consistent with Moscarini and Postel-Vinay (2012) findings, causality is reversed : in our setting, it is not that large firms are more cyclically sensitive. Rather it is the business cycle itself that reflects large firms dynamics.

The paper relates to two distinct literatures: an emerging literature on the micro-origins of aggregate fluctuations and the more established firm dynamics literature. Gabaix (2011) describes the "granular hypothesis" and shows the possibility result that we extend to our framework. Other papers studying the micro-origins of aggregate fluctuations are Acemoglu et *al.* (2012), di Giovanni and Levchenko (2012), Carvalho (2010) and Carvalho and Gabaix (2013)[1]. Relative to this literature, we contribute by grounding the granular hypothesis in a standard firm dynamics setting, extending the existing theoretical results to this setting and providing a first attempt at quantification.

This paper is also related to the firm dynamics literature: Hopenhayn (1992), Campbell (1998), Veracierto (2002), Khan and Thomas (2003, 2008), Bachman and Bayer (2009). Some papers have studied aggregate fluctuations in an entry/exit framework as Lee and Mukoyama (2008), Clementi and Palazzo (2010) and Bilbiie et *al.* (2012). However they restrict their analysis to common, aggregate, shocks. Relative to this literature, we show that its standard workhorse model – once the assumption regarding a continuum of firms is dropped and the firm size distribution is fat tailed – already contains in it a theory of the business cycle. We show this both theoretically and quantitatively.

The paper is organized as follows. Section 2 derives the model. Section 3 shows our theoretical results on the rate of decay of aggregate fluctuations. Section 4 explains the calibration, the origin of fat-tailed firm size distribution and exposes our first quantitative results. Section 5 displays the impulse response to a negative shock on the largest firm. Section 6 looks at the cyclicality of large versus small firm in our model. Finally, section 7 concludes.

[1]Some empirical evidence can be found in di Giovanni, Levchenko and Mejean (2012).

2 Model

We analyze a standard firm dynamics' setup (Hopenhayn, 1992) with a finite but possibly large number of firms. We show how to solve for and characterize the evolution of the firm size distribution without relying on the usual law of large numbers assumption. We prove that in this setting the firm size distribution does not converge to a stationary distribution, but instead fluctuates stochastically around it. As a result, we show that aggregate prices and quantities are not constant over time as the continuum assumption in Hopenhayn (1992) - repeatedly invoked by the subsequent literature - does not apply. Because of this crucial property, we are able to show that the standard firm dynamics setup is able to generate aggregate fluctuations with only idiosyncratic, firm-level shocks.

In this section we first describe the economic environment. As is standard in this class of models, this involves specifying a firm-level productivity process, the incumbents' problem and the entrants' problem. We then show how to characterize the law of motion of the productivity distribution.

2.1 Model Setup

The setup follows Hopenhayn (1992) closely. Firms differ in their productivity level, which here is assumed to follow a discrete Markovian process. Incumbents have access to a decreasing returns to scale technology using labor as the only input. They produce a unique good in a perfectly competitive market. They face an operating cost at each period, which in turn generates endogenous exit. There is a large but finite number of potential entrants that differ in their productivity. To operate next period, potential entrants have to pay an entry cost. The economy is closed in a partial equilibrium fashion by specifying a labor supply function that increases with the wage.

Productivity Process

We assume a finite but potentially large number of idiosyncratic productivity levels. The productivity space is thus described by a S-uple $\Phi := \{\varphi_1, \ldots, \varphi_S\}$ such that $\varphi_1 < \ldots < \varphi_S$. A firm is in state (or productivity state) k when its idiosyncratic productivity is equal to φ_k. Each firm's (log) productivity level is assumed to follow a Markov chain with a transition matrix P. We denote $F(.|\varphi)$ as the conditional distribution of the next period's idiosyncratic productivity φ' given the current period's idiosyncratic productivity φ.[2]

Incumbents' Problem

The only aggregate state variable of this model is the distribution of firms on the set Φ. We denote this distribution by a $(S \times 1)$ vector μ_t giving, for each level of productivity, the number of firms at this productivity level at t. For the current setup description, we abstract from explicit time t notation, but will return to it when we characterize the law of motion of the aggregate state. Given an aggregate state μ, and an idiosyncratic productivity level φ, the incumbent solves the following static profit maximization problem:

$$\pi^*(\mu, \varphi) = \underset{n}{\text{Max}} \left\{ \exp(\varphi) n^{\alpha} - w(\mu) n - c_f \right\}$$

where n is the labor input, $w(\mu)$ is the wage for a given aggregate state μ, and c_f is the operating cost to be paid every period. It is easy to show that π^* is increasing in φ and decreasing in w for a given aggregate state μ. The output of a firm is then $y(\mu, \varphi) = \exp(\varphi)^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{w(\mu)} \right)^{\frac{\alpha}{1-\alpha}}$. In what follows, the size of a firm will refer to its output level if not otherwise specified.

The timing of decisions for incumbents is standard and described as follows. The incumbent first draws its idiosyncratic productivity φ at the beginning of the period, pays the operating cost c_f and then hires labor to produce. It then decides whether to exit at the end of the period or to continue as an incumbent the next period. We denote the present discounted value of being an incumbent for a given aggregate state μ and idiosyncratic productivity level φ by $V(\mu, \varphi)$, defined by the following Bellman equation:

$$V(\mu, \varphi) = \pi^*(\mu, \varphi) + \max \left\{ 0, \beta \int_{\mu' \in \Lambda} \sum_{\varphi' \in \Phi} V(\mu', \varphi') F(\varphi'|\varphi) \Gamma(d\mu'|\mu) \right\}$$

[2]That is, given a productivity level φ_s the distribution $F(.|\varphi_s)$ is given by the s^{th}-row vector of the matrix P. Throughout we assume that $F(.|\varphi)$ is decreasing in φ.

where β is the discount factor, $\Gamma(.|\mu)$ is the conditional distribution of μ', tomorrow's aggregate state and $F(.|\varphi)$ is the conditional distribution of tomorrow's idiosyncratic productivity given today's idiosyncratic productivity of the incumbent.

The second term on the right hand side of the value function above encodes an endogenous exit decision. As is standard in this framework, this decision is defined by a threshold level of idiosyncratic productivity given an aggregate state. Formally, since the instantaneous profit is increasing in the idiosyncratic productivity level and $F(.|\varphi)$ is decreasing in φ, there is a unique index $s^*(\mu)$, for each aggregate state μ, such that (i) for $\varphi \geq \varphi_{s^*(\mu)}$ the incumbent firm continues in operation next period and, conversely (ii) for $\varphi \leq \varphi_{s^*(\mu)-1}$ firm decides to exit next period.

After studying the incumbents' problem, we now turn to the problem of potential entrants.

Entrants' Problem

We formulate the entry decision as a discretized version of Clementi and Palazzo (2010). There is an exogenously given, constant and finite number of prospective entrants M. Each potential entrant has access to a signal about their potential productivity next period, should they decide to enter today. To do so they have to pay a sunk entry cost which, in turn, leads to an endogenous entry decision which is again characterized by a threshold level of initial signals.

Formally, entrants' signals are distributed according $G = (G^q)_{q \in [1 \ldots S]}$, a discrete distribution over Φ. There is a total of M potential entrants every period, so that the $M.G^q$ gives the number of potential entrants for each signal level φ_q. If a potential entrant decides to pay the entry cost c_e, then she will produce next period with a productivity level drawn from $F(.|\varphi_q)$. Given this we can define the value of a potential entrant with signal φ_q for a given the aggregate state μ as $V^e(\mu, \varphi_q)$:

$$V^e(\mu, \varphi_q) = \beta \int_{\mu' \in \Lambda} \sum_{\varphi' \in \Phi} V(\mu', \varphi') F(d\varphi'|\varphi_q) \Gamma(d\mu'|\mu)$$

A prospective entrant pay the entry cost and produce the next period if the above value is greater or equal to the entry cost c_e. As in the incumbent exit decision, this now induces a threshold level of signal, $e^*(\mu)$, for a given aggregate state mu such that (i)

for $\varphi_q \geq \varphi_{e^*(\mu)}$ the potential entrant starts operating next period and, conversely (ii) for $\varphi_q \leq \varphi_{e^*(\mu)-1}$ the potential entrant decides not to do so.

For simplicity, henceforth we assume that the entry cost is normalized to zero: $c_e = 0$ which in turn implies that $\varphi_{e^*(\mu)} = \varphi_{s^*(\mu)}$.

Labor Market and Aggregation

Following Clementi and Palazzo (2010), we assume that the supply of labor at a given wage w is given by $L^s(w) = L(M)w^\gamma$ with $\gamma > 0$. We assume that the parameter in the labor supply function is a function of M, the number of potential entrants. This assumption is necessary because in what follows we will be interested in characterizing the behavior of aggregate quantitites and prices as we let M increase. If total labor supply were to remain fixed across these economies while the total demand for labor market increases, the wage would increase mechanically. We therefore make this assumption to abstract from this effect of increasing M on the equilibrium wage.

To find equilibrium wages, we now derive the aggregate labor demand in this economy. To do this, note that if Y_t is the aggregate output, i.e. the sum of all individual incumbents' output, then $Y_t = A_t(L_t^d)^\alpha$ where L_t^d is the aggregate labor demand, the sum of all period t incumbents' labor demand. A_t, the aggregate total factor productivity, is given by:

$$A_t = \left(\sum_{i=1}^{N_t} \exp(\varphi_t^i)^{\frac{1}{1-\alpha}} \right)^{1-\alpha}$$

where φ_t^i is the productivity level at date t of the i^{th} firm among the N_t operating firms at date t. This can be rewritten by aggregating all firms that have the same productivity level:

$$A_t = \left(\sum_{s=1}^{S} \mu_t^s \exp(\varphi_s)^{\frac{1}{1-\alpha}} \right)^{1-\alpha} = (B'.\mu_t)^{1-\alpha}$$

where B is the $(S \times 1)$ vector of parameters $(\exp(\varphi_1)^{\frac{1}{1-\alpha}}, \ldots, \exp(\varphi_S)^{\frac{1}{1-\alpha}})$. As discussed above, the distribution of firms μ_t across the discrete state space $\Phi = \{\varphi_1, \ldots, \varphi_S\}$ is a $(S \times 1)$ vector equal to $(\mu_t^1, \ldots, \mu_t^S)$ such that μ_t^s is equal to the number of operating firms in state s at date t. By the same argument, it is easy to show that aggregate labor demand is given by $L^d(w_t) = \left(\frac{\alpha A_t}{w_t}\right)^{\frac{1}{1-\alpha}}$. Note that the model behaves as a one factor model with aggregate TFP A_t.

The market clearing condition then equates labor supply and labor demand, i.e. $L^s(w_t) = L^d(w_t)$. Given date t productivity distribution μ_t, we can then solve for the equilibrium wage to get:

$$w_t = \left(\alpha^{\frac{1}{1-\alpha}} \frac{B'.\mu_t}{L(M)}\right)^{\frac{1-\alpha}{\gamma(1-\alpha)+1}} \tag{1}$$

From this expression, note that the wage is fully pinned down by the distribution μ_t. That is, given a current period distribution of firms across productivity levels we can solved for all equilibrium quantities and prices. We are left to understand how this distribution evolves over time, i.e. how to solve for for $\Gamma(.|\mu_t)$, which is addresses in the following section.

2.2 Aggregate State Dynamics and Uncertainty

In this section, we first show how to characterize the law of motion for the productivity distribution, the aggregate state in this economy. We will prove that, generically, the distribution of firms across productivity levels is time-varying in a stochastic fashion. An immediate implication of this result is that aggregate productivity A_t and aggregate prices are themselves stochastic as they are simply a function of this distribution. Additionally, we then show that the characterization of the stationary firm productivity distribution offered in Hopehayn obtains only as a limiting result of our setup, as the number of firms goes to infinity.

Law of Motion of the Productivity Distribution

In the standard setting with a continuum of firms, Hopenhayn (1992) shows that by appealing to a law of large numbers, the law of motion for the productivity distribution is in fact deterministic. In the current setting, with a finite number of incumbents, a similar argument cannot be made. We now show how to characterize the law of motion for the productivity distribution.

In order to build intuition for our general result below, we start by exploring a simple example where, for simplicity, we ignore entry and exit of firms. Assume there are only

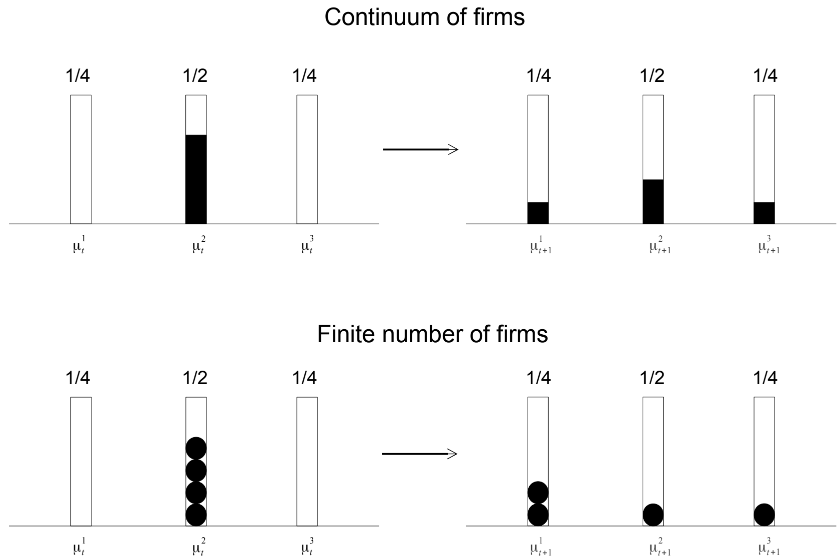

Figure 1: Why the vector μ_{t+1} follows a multinomial distribution. Top panel: continuum of firms case, unique possible outcome. Bottom panel: finite number of firms, one (of many) possible outcome.

three levels of productivity ($S = 3$) and four firms. At time period t these firms are distributed according to the bottom-left panel of figure 1, i.e. all four firms produce with the intermediate level of productivity. Further assume that these firms have an equal probability of $1/4$ of going up or down in the productivity ladder and that the probability of staying at the same intermediate level is $1/2$. That is, the second row of the transition matrix P in this simple example is given by $(1/4, 1/2, 1/4)$. First note that, if instead of four firms we had assumed a continuum of firms, the law of large number would hold such that at $t+1$ there would be exactly $1/4$ of the (mass of) firms at the highest level of productivity, $1/2$ would remain at the intermediate level and $1/4$ would transit to the lowest level of productivity (top panel of figure 1). This is not the case here, since the number of firms is finite. For instance, a distribution of firms such as the one presented in the bottom-right panel of figure 1 is possible with a positive probability. Of course, many other arrangements would be possible outcomes. Thus, in this example, the number of firms in each productivity bin at $t + 1$ follows a multinomial distribution with a number of trials of 4 and an event probability vector $(1/4, 1/2, 1/4)'$.

In this simple example, all firms are assumed to have the same productivity level at time t. It is easy however to extend this example to any initial arrangement of firms over producivity bins. This is because, for any initial number of firms *at a given* productivity level, the distribution of *these* firms across productivity levels next period follows a multinomial. Therefore, the *total* number of firms in each productivity level next period, is simply a sum of multinomials, i.e. the result of transitions from *all* initial productivity bins.

More generally, for S productivity levels, an (endogenous) finite number of incumbents, N_t making optimal employment and production decisions and and accounting for entry and exit decisions, the following theorem holds.

Theorem 1 *The number of firms at each productivity level at* $t+1$*, given by the* $(S\times 1)$ *vector* μ_{t+1}*, conditional on the current vector* μ_t*, follows a sum of multinomial distributions and can be expressed as:*

$$\mu_{t+1} = m(\mu_t) + \epsilon_{t+1} \tag{2}$$

where ϵ_{t+1} *is a random vector with mean zero and a variance-covariance matrix* $\Sigma(\mu_t)$ *and*

$$m(\mu_t) = (P_t^*)'(\mu_t + MG)$$
$$\Sigma(\mu_t) = \sum_{s=s^*(\mu_t)}^{S} (MG^s + \mu_t^s).W_s$$

where P_t^* *is the transition matrix* P *with the first* $(s^*(\mu_t) - 1)$ *rows replaced by zeros.* $W_s = diag(P_{s,.}) - P_{s,.}P_{s,.}^{'}$ *where* $P_{s,.}$ *denotes the* s*-row of the transition matrix* P*.*

Proof. The distribution of firms μ_t across the discrete state space $\Phi = \{\varphi_1, \ldots, \varphi_S\}$ is a $(S\times 1)$ vector equal to $(\mu_t^1, \ldots, \mu_t^S)$ such that μ_t^s is equal to the number of operating firms in state s at date t. The next period's distribution of firms across discrete state space $\Phi = \{\varphi_1, \ldots, \varphi_S\}$ is the sum of the evolution of incumbents and successful entrants.

In what follows, we define two conditional distributions. First, the distribution of incumbent firms at date $t+1$ conditional on the fact that incumbents were in state s at date t is denoted as $f_{t+1}^{.,s}$. This $(S\times 1)$ vector is such that for each state k in $\{1, \ldots, S\}$, $f_{t+1}^{k,s}$, the k^{th} element of $f_{t+1}^{.,s}$ gives the number of incumbents in state k at $t+1$ which were in state φ_s at t.

Similarly, let us define $g_{t+1}^{\cdot,s}$ the distribution of successful entrants at date $t+1$ given that they received the signal φ_s at date t. This $(S \times 1)$ vector is such that for each state k in $\{1, \dots, S\}$, $g_{t+1}^{k,s}$, the k^{th} element of $g_{t+1}^{\cdot,s}$ gives the number of entrants in state k at $t+1$ which received a signal φ_s at t.

The period $t+1$ firms distribution is the sum of all these conditional distributions and thus the vector μ_{t+1} satisfies:

$$\mu_{t+1} = \sum_{s=s^*(\mu_t)}^{S} f_{t+1}^{\cdot,s} + \sum_{s=s^*(\mu_t)}^{S} g_{t+1}^{\cdot,s} \tag{3}$$

Note that $f_{t+1}^{\cdot,s}$ and $g_{t+1}^{\cdot,s}$ are now multivariate random vectors implying that μ_{t+1} also is a random vector.

At date $t+1$ for $s \geq s^*(\mu_t)$, $f_{t+1}^{\cdot,s}$ follows a multinomial distribution with two parameters: the integer μ_t^s and the $(S \times 1)$ vector $P'_{s,.}$ where $P_{s,.}$ is the s^{th} row vector of the matrix P. For $s \leq s^*(\mu_t) - 1$, $f_{t+1}^{\cdot,s}$ is equal to zero. Similarly, at date $t+1$ for $s \geq s^*(\mu_t)$, $g_{t+1}^{\cdot,s}$ follows a multinomial distribution with two parameters: the integer $M.G^q$ and the $(S \times 1)$ vector $P'_{q,.}$. For $s \leq s^*(\mu_t) - 1$, $g_{t+1}^{\cdot,s}$ is equal to zero.

Recall that the mean and variance-covariance matrix of a multinomial distribution, $M(m,h)$, is respectively the $(S \times 1)$ vector mh and the $(S \times S)$ matrix $H = diag(h) - h'h$. So let us define $W_s = diag(P_{s,.}) - P_{s,.}P'_{s,.}$. From the right hand side of equation 3, using the fact that the $f_{t+1}^{\cdot,s}$ and $g_{t+1}^{\cdot,s}$ follow multinomials, μ_{t+1} has a mean $m(\mu_t)$ and a variance-covariance matrix $\Sigma(\mu_t)$ where

$$m(\mu_t) := \sum_{s=s^*(\mu_t)}^{S} \Big[\mu_t^s.P_{s,.} + MG^s.P_{s,.}\Big] = (P_t^*)'(\mu_t + MG)$$

$$\Sigma(\mu_t) := \sum_{s=s^*(\mu_t)}^{S} (MG^s + \mu_t^s).W_s$$

where P_t^* is the transition matrix P with the first $(s^*(\mu_t) - 1)$ rows replaced by zeros.

Equation 3 can be rewritten in a simple way as the sum of its mean and a zero-mean shock:

$$\mu_{t+1} = m(\mu_t) + \epsilon_{t+1}$$

where

$$\epsilon_{t+1} = \sum_{s=s^*(\mu_t)}^{S} \left[f_{t+1}^{\cdot,s} - \mu_t^s.P_{s,.} \right] + \sum_{s=s^*(\mu_t)}^{S} \left[g_{t+1}^{\cdot,s} - MG^s.P_{s,.} \right]$$

i.e ϵ_{t+1} is the demeaned version of μ_{t+1}. This gives us the result stated in the theorem.

□

After taking into account the dynamics of incumbent firms and entry/exit decisions, the law of motion of the aggregate state -the firm distribution over productivity levels- is remarkably simple: tomorrow's firm distribution is an affine function of today's distribution up to a stochastic term that reshuffles firms across productivity levels.

It is easy to understand this characterization by recalling our simple example economy above without entry and exit. Given the state transition probabilities in this simple case we should, for example, observe that *on average* the number of firms remaining at the intermediate level of productivity is twice that of those transiting to the highest level of productivity. This is precisely what the affine part of equation 2 captures: the $m(\mu_t)$ term reflects these typical transitions, which are a function of the P matrix only. However, with a finite number of firms, in any given period there will be stochastic deviations from these typical transitions as we discuss above. In the theorem, this is reflected in the "reshuffling shock" term, ϵ_{t+1}, in the law of motion 2. How important this "reshuffling shock" is for the evolution of the firm distribution is dictated by the variance-covariance matrix $\Sigma(\mu_t)$ which, in turn, is a function of the transition matrix P, the current firm distribution μ_t and, in the general case with entry and exit, the signal distribution available to potential entrants.

Limit Economies

The above characterization - in particular, Equation 2 - is instructive of the differences of the current setup relative to a standard Hopenhayn economy. The latter corresponds to the case where all of the relevant firm dynamics are encapsulated by the affine term $m(\mu_t)$. Appealing to a law of large numbers argument with a finite number of firms is therefore equivalent to imposing the variance-covariance matrix to be zero so that the aggregate state μ_t becomes non-stochastic. That is, the "reshuffling shock", ϵ_{t+1}, would be absent. In this section we show that this assumption is justified as the number of firms goes to infinity. That is, we show that the characterization of the stationary firm productivity distribution offered in Hopehayn obtains as a limiting result of our setup, as the number of firms goes to infinity. In order to show this, we first make an assumption 1 about the labor supply function.

Assumption 1 *The labor supply, for a given wage, is a linear function of the number of potential entrants:* $L(M) = \underline{L}M$.

This assumption ensures that the equilibrium wage is independent of the number of potential entrants. To see this note that given the labor market equilibrium condition (equation 1), under assumption 1 the equilibrium wage is now a function of $\hat{\mu}_t := \frac{\mu_t}{M}$, the normalized productivity distribution across productivity levels. Given that M is a parameter of the model, we can therefore use μ_t or $\hat{\mu}_t$ interchangeably as the agggregate state variable. We find it convenient to work with the latter in what follows.

In the following limit characterization we will be taking the number of potential entrants M to infinity. Notice that this implies that the number of incumbents, N_t, also goes to infinity, since the latter scales with the former. To see this note that letting e be the $(S \times 1)$ unit vector, then $N_t = e'.\mu_t = Me'.\hat{\mu}_t$. Indeed one can see from equation 2 that μ_t scales with M and so does N_t.

We can now state the following corollary of theorem 1 which shows that the Hopenhayn framework is nested in ours as the number of potential entrants goes to infinity.

Corollary 1 *Let us define* $\hat{\mu}_t := \frac{\mu_t}{M}$ *for any* t. *Under assumption 1, as* M *goes to infinity we have:*

$$\hat{\mu}_{t+1} = (P_t^{\infty})'(\hat{\mu}_t + G) \tag{4}$$

where P_t^{∞} *is the transition matrix* P *where the first* $s^{\infty}(\mu_t) - 1$ *rows are replaced by zero, and where* $s^{\infty}(\mu_t)$ *is the threshold of the exit/entry rule when there is no aggregate uncertainty i.e the variance-covariance of the* ϵ_{t+1} *is zero.*

Proof. See appendix B.1. □

Therefore, as the number of potential entrants, M, goes to infinity, the law of motion for the distribution of firms across productivity levels is deterministic and its evolution is given by equation 4. An immediate consequence of this is that under appropriate conditions on the transition matrix P, a stationary distibution exists and is given by:

$$\hat{\mu} = (I - P^{\infty\prime})^{-1} P^{\infty\prime} G \tag{5}$$

the matrix P^{∞} as the matrix P where the first $(s^{\infty}(\hat{\mu}) - 1)$ rows are replaced by zeros. This is the analogue in the current setting to Hopenhayn's (1992) stationary distribution.

Taking stock, we have derived a law of motion for any finite number of firms and shown that, generically, the distribution of firms across productivity levels is time-varying in a stochastic fashion. An immediate implication of this is that aggregate productivity A_t is itself stochastic as it is simply a function of this distribution. The latter corollary implies that only in the continuum case does the distribution converge to a stationary distribution and hence in this limiting case there are no aggregate fluctuations. As such, whether the model considered above can be taken as a quantitatively relevant theory of the business cycle, depends on the rate at which aggregate fluctuations vanish as a function of the number of firms. We characterize this decay rate under some conditions in the next section.

3 Aggregate Fluctuations

In this section, we explore a tractable parameterization of the stationary distribution and characterize analytically the rate of decay of aggregate fluctuations. We first show some preliminary results and define a fat-tail distribution in our context. Then we turn to our theoretical result concerning the rate of decay of aggregate volatility. Finally, we also briefly study the implications of firm level disturbances on the persistence of aggregate variable.

3.1 Preliminary Results and Definitions

We first show how the tail of the productivity distribution is related to its counterpart in the case of infinite number of firms. We then specialize our assumption on the class of firm-level productivity processes. We then explain why the firm-size distribution is simply a relabelling of the productivity distribution. Finally we define what we mean by a fat-tail distribution.

We start by stating a lemma that shows that, for a large enough number of firms, the right tail of the firm productivity distribution at each date t is well approximated by the right tail behavior of the stationary distribution $\hat{\mu}$. Formally, we show that the tail of firm productivity distribution at any point in time, $\hat{\mu}_t$ converges in probability to the tail of the stationary productivity distribution, $\hat{\mu}$.

Lemma 1 *Assuming that the economy starts at date $t = 0$ with the stationary firm productivity distribution $\hat{\mu}$. Then at each period t countercumulative of the firm productivity distribution $\hat{\mu}_t$ converges in probability to the countercumulative stationary distribution $\hat{\mu}$ as M, the number of potential entrants, goes to infinity.*

Proof See appendix B.2. □

The proof first defines the deviation of μ_t from μ making use of equation 2 derived in the previous section. It then defines the associated (discrete) counter-cumulative distribution function (CCDF). Finally, it shows that the CCDF of the productivity distribution at any given date t converges in probability to the countercumulative stationary productivity distribution. Additionally, the proof also shows that for a given history of "reshuffling shocks" and a fixed M, the distance between the countercumulative productivity distribution at any given date t and the countercumulative stationary distribution is governed by the firm-level transition probability matrix P^*. Heuristically, this is stating that for any finite number of firms, the productivity distribution hovers around the stationary distribution and how far it can be from the latter is a function of persistence of the stochastic productivity process.

For the remainder of this section, we restrict the class of transition probability matrices P to be that given by the Rouwenhorst (1995) method (see Kopecky and Suen, 2010 for a formal discussion). This is a method to discretize an AR(1) process, such that the markov chain is entirely characterized by three parameters, ρ, the first order autocorrelation of the corresponding AR(1), σ_e, its conditional variance and φ_S, the upper bound of the discretized state space. We follow Kopecky and Suen, (2010) in calling such transition matrices Rouwenhorst matrices. Specifically in the proofs of the theorems below we make use of the fact that, for this class of transition probability matrices, the first-order autocorrelation and conditional variance associated to the markov chain P are exactly ρ and σ_e.[3].

Assumption 2 *The transition probability matrix P is a Rouwenhorst matrix.*

Thus far we have been making use of the aggregate state variable, μ_t, the firm-level productivity distribution. As typical in this class of models, the firm-size distribution maps one to one to μ_t. To see this note that the output of a firm with productivity level φ_s at time t is given by: $y_t^s = \exp(\frac{\varphi_s}{1-\alpha})\left(\frac{\alpha}{w_t}\right)^{\frac{\alpha}{1-\alpha}}$. Therefore, at time t, the number of firms of size y_t^s is given by μ_t^s so that the firm size distribution is given by μ_t.

We also define the latent size of a potential entrant with signal φ_s, as the output level that would have obtained if an operating firm had productivity equal to φ_s. By the same argument as above, the number of such firms would be given by G^s. Therefore the latent size distribution is given by the vector $M.G$.

In what follows, following up on Gabaix (2011), we restrict our attention to power law distributions which we now define formally following Acemoglu *et al.* (2012).

Definition 1 *A sequence of $(S \times 1)$-vector $(E_M)_{M\in\mathbb{N}}$ has a power law tail if there exist a constant $\beta > 1$, a slowly varying function $L(.)$ satisfying $\lim_{t\to\infty} L(t)t^{\delta} = \infty$ and $\lim_{t\to\infty} L(t)t^{-\delta} = 0$ for all $\delta > 0$, and a sequence of M positive numbers $c_M = O(1)$ such that, for all $M \in \mathbb{N}$ and all $k < S$, we have*

$$\mathbb{P}(k) = c_M k^{-\beta} L(k)$$

where $\mathbb{P}(k) = \sum_{s=k+1}^{S} E_{s,M}$ is the counter-cumulative function.

In our context, when we say that μ_t has a power law tail with parameter ξ, we are applying the definition 1 with $E_M = M.\hat{\mu}_t$ and $\beta = \xi$. Simililarly we say that the latent size distribution of potential entrant has a power law tail with parameter ζ.

[3]For completeness we describe this class of transition matrices in detail in the appendix A

3.2 Rate of Decay of Aggregate Volatility

We are now ready to state theorem 2 describing the rate of decay of the aggregate volatility, under the assumption of power law distribution.

Theorem 2 *Let the firm size distribution have a power law tail with parameter ξ and the latent size distribution of potential entrants has a power law tail with parameter ζ. Under assumption 2, if $\xi/(2\rho) < 1$ and $\zeta/(2\rho) < 1$ then*

$$\sigma_t(\frac{\Delta Y_{t+1}}{Y_t}) = \Omega\left(\sigma\left[\frac{u}{(N_t)^{2-2\rho/\xi}}\left(\frac{N_t^I}{N_t}\right)^{2\rho/\xi} + \frac{w}{(N_t)^{2-2\rho/\zeta}}\left(\frac{N_t^E}{N_t}\right)^{2\rho/\zeta}\right]^{1/2}\right) \tag{6}$$

where $\Omega()$ means that, given two sequences of positive real numbers $\{a_M\}_{M\in\mathbb{N}}$ and $\{b_M\}_{M\in\mathbb{N}}$, $a_n = \Omega(b_n)$ if $\liminf_{M\to\infty} a_M/b_M > 0$ and where u and w are non degenerate random variables; σ is the standard deviation of $\exp(e_i^t/(1-\alpha))$ and N_t, N_t^I, N_t^E are respectively the number of incumbents at t, surviving time t incumbents and actual entrants in period t.

Proof See appendix B.3. □

Theorem 2 characterizes the conditional volatility of aggregate output growth in this economy, i.e. the volatility of the growth rate of output conditional on the time t state. As announced before, we are assuming that both the stationary incumbent size distribution and the stationary latent size distribution of potential entrants are power law distributed. As a result of Lemma 1 we know that, as M goes to infinity, the time t firm size distribution is also power law distributed. The assumptions on ξ and ζ ensure that these distributions are sufficiently fat tailed.

The Ω notation means that aggregate growth volatility scales with N_t, the number of time t incumbents, at rate that is no faster than the rate of the variable in the right hand side parenthesis. The latter reflects the separate contribution of (i) surviving time t incumbents (i.e. the first term in parenthesis) and that of (ii) time t entrants that start producing at $t+1$ (i.e. the second term in parenthesis).

The key conclusion of the theorem is therefore that, for $\xi/(2\rho) < 1$ and $\zeta/(2\rho) < 1$, the aggregate volatility scales at a slower rate than $1/\sqrt{(N_t)}$. Recall that the latter would be the rate of decay of aggregate volatitlity implied by a standard diversification argument relying on standard central limit theorems. Such a theorem does not apply in a world of fat tailed distributions. In this way, the theorem generalizes Gabaix (2011) to an environment where firms are making endogenous decisions about their optimal scale (along with exit and entry decisions) while being faced with potentially persistent idiosyncratic productivity shocks. More generally, the theorem implies that the large extant literature that builds on the Hopenhayn (1992) in order to understand firm-level dynamics and the firm size distribution potentially contains in it a theory of aggregate fluctuations.

To gain an understanding of how slow this convergence to the continuum case is, we find it useful to abstract from the contribution of entry an exit. This can be justified by the fact that at any time period, both in quantitative versions of the model (as we show below) and in the data, entry/exit dynamics are of second order in accounting for aggregate fluctuations. Thus, focusing on the first term alone for $N_t^I \approx N_t$:

$$\sigma_t(\frac{\Delta Y_{t+1}}{Y_t}) = \Omega\left(\sigma \frac{u^{1/2}}{(N_t)^{1-\rho/\xi}}\right)$$

This approximation is useful to isolate the role played by mean-reversion of the firm-level productivity in shaping the rate of decay of aggregate volatility. As the expression makes clear, the higher is the persistence of the firm-level productivity process, the slower is the rate of decay of aggregate volatility. To understand this, recall our previous discussion on the role of productivity transition matrix P^* in determining how distant can the μ_t distribution can be from its stationary counterpart μ. Given asssumption 2, this role is now pinned down by a single persistence parameter ρ. Therefore, higher persistence, i.e. a higher ρ, implies that any time t deviation from the stationary distribution is likely to persist at time $t+1$. Intuitively, high persistence implies that time t large firms are reverting to the average size in a slower fashion. So in the high persistence case, a given large firm would be larger on average next period relative to the low persistence case. Its impact on the aggregate is then higher, i.e. its contribution to the aggregate volatility is higher than in the low persistence case.[4]

[4]The persistence parameter ρ actually plays a dual role: (i) as we have just discussed, it influences the out of stationary equilibrium dynamics; (ii) it also influences directly the tail behavior of the stationary distribution as equation 5 shows.

3.3 Aggregate Persistence (Preliminary)

We now turn our attention to the persistence of aggregate output growth and how it scales with the number of firms. The theorem below characterizes the conditional expectation of aggregate growth rate as a function of model parameters. Recall that, for a generic AR(1) process with persistence parameter ρ the conditional expectation of aggregate growth rate is given by $\rho - 1$, so that the latter simply reflects the first-order autocorrelation of the underlying process.

Theorem 3 *Let the firm size distribution have a power law tail with parameter ξ and the latent size distribution of potential entrants has a power law tail with parameter ζ. Under assumption 2 and for $1 < \xi < 2$ and $0 < \zeta'/\rho < 1$ then:*

$$\frac{\Delta Y_{t+1}}{Y_t} \approx \left(1 - \frac{\alpha}{\gamma(1-\alpha)+1}\right)\frac{\Delta T_{t+1}}{T_t}$$

$$\mathbb{E}_t\left[\frac{\Delta T_{t+1}}{T_t}\right] \sim \frac{N_t^I}{N_t}\frac{k\bar{Z}^{\rho}{}_t - \bar{Z}_t}{\bar{I}_t} - \frac{N_t^X}{N_t}\frac{\bar{X}_t}{\bar{I}_t} - \frac{v/\bar{I}_t}{N_t^{1-1/\xi}}\left(\frac{N_t^I}{N_t}\right)^{1/\xi} + \frac{u/\bar{I}_t}{N_t^{1-\rho/\zeta'}}\left(\frac{N_t^E}{N_t}\right)^{\rho/\zeta'}$$

N_t, N_t^I, N_t^E and N_t^X are respectively the number of incumbents at t, surviving time t incumbents, entrants in period t and exiters in period t. $\bar{Z}_t, \bar{Z}_t^{\rho}, \bar{X}_t, \bar{I}_t$ are time dependent variables proportional to, respectively, the average size of surviving time t incumbents, the average size at $t+1$ of surviving time t incumbents, the average size of exiters at t and the average size of all time t incumbents. v and u are non-degenerate random variables and k is a constant.

Proof See appendix B.4. □

The theorem expresses the conditional expectation of aggregate growth rate as the sum of the contribution of entry, exit and surviving incumbent dynamics. To understand how the conditional expectation of aggregate growth rate scales with the number of firms we again find it convenient to focus on a particular case where we abstract from entry and exit dynamics[5]:

$$\mathbb{E}_t\left[\frac{\Delta T_{t+1}}{T_t}\right] \sim \rho - 1 - \frac{v/\bar{Z}_t}{(N_t)^{1-1/\xi}}$$

[5] Formally, we assume $N_t^I \approx N_t$, $N_t^E \approx N_t^X \approx 0$ and $k \approx 1$

As expected, the conditional expectation of aggregate growth rate of aggregate output converges to $\rho - 1$ in the limit. Under the usual central limit theorem arguments this convergence should take place at rate $1/\sqrt{N}$. The above equation states that, under a fat tailed firm size distribution, this rate of convergence is slower and governed by the tail of the firm size distribution.

4 Simulations (Preliminary)

The model is solved using the algorithm described in section C.1. A discussion of the source of the fat-tailedness of the firm size distribution follows. Finally, we compute some business cycle statistics.

4.1 Calibration

To calibrate the model to the US economy, we first set the value of some deep parameters and the functional form associated with the distribution of potential entrants and the labor supply function. The span of control parameter α is set at 0.8. This value is chosen to be on the lower end of recent estimates, such as Basu and Fernald (1997) and Lee (2005). The discount factor β is set at 0.95 so that the implied annual gross interest rate is 4%, a value in line with the business cycle literature. The entry cost is normalized to be zero to keep the mathematics as simple as possible. Finally, the potential entrant distribution is such that the latent size distribution of potential entrant is a generalized pareto distribution with tail parameter ζ, scale parameter λ and zero as location parameter. As we discussed above, we assume that the labor supply scaling parameter is a linear function of the number of potential entrants: $L(M) = \underline{\mathrm{L}}M$. The labor supply elasticity parameter, γ, is chosen to be 2.

We jointly calibrate the operating cost c_f, the first order autocorrelation ρ and variance σ_e of the productivity process[6], the constant $\underline{\mathrm{L}}$, the parameter of the latent size distribution of potential entrants ζ, λ, and, the upper bound of the idiosyncratic producivity space φ_S. The parameter are chosen such that some statistics of the stationary equilibrium (the one when M goes to infinity) match their empirical counterpart. These

[6]The productivity process is a discretization of a AR(1) process with autocorrelation ρ and variance of shocks σ_e^2. To do so, we follow the Tauchen (1986) method.

Statistic	Model	Data	References
Entry Rate	0.02	0.10	BDS 2011
Entrants' relative size	0.29	0.29	BDS 2011
Maximum size	1.1×10^3	1.1×10^3	Larger Category in BDS is 1000+
Tail index of Firm size dist.	1.21	1.03	Gabaix (2011)

Table 1: Targets for the calibration procedure

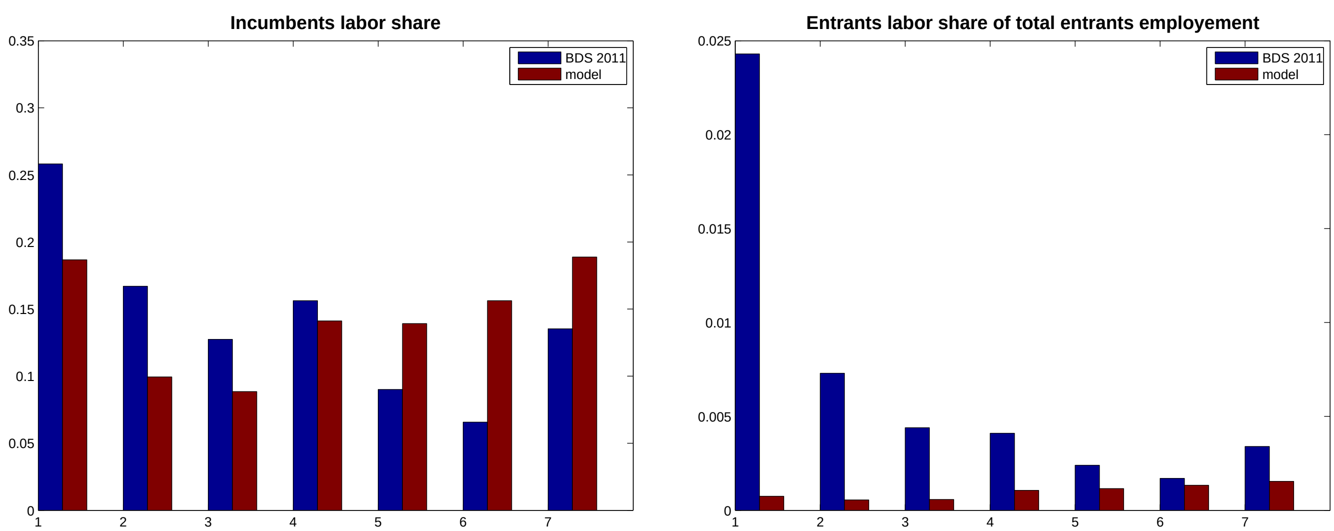

Figure 2: Incumbent and entrants labor share by size clase (Sources: BDS and Model)

statistics are the entry rate, the entrant relative size, the maximum size, the incumbent and entrant size distribution, and finally the tail index of the firm size distribution. Except form the latter, these statistics are computed using the Business Dynamics Statistcs (BDS) in 2011. The target for the tail index of the firm size distribution is the one estimated by Gabaix (2011).

Table 1 and figure 2 summarize the result of this calibration procedure. The corresponding moments in the model are computed for the stationary deterministic equilibrium. All the calibrated parameters are described in table 2.

The model is able to match most moments, except the entry rate, which is too low in our baseline calibration. As the right panel of figure 2 shows the entry rate is too low in our baseline calibration because we are not able to match the number of small entrants in our model. This can be address by assuming a more sophisticated distribution for entrants. The first order autocorrelation and the conditional standard deviation of the productivity process are in line with previous estimates as in Lee and Mukoyama (2008) or in Castro et al (2013).

4.2 The Tail of the Firm Size Distribution

The calibration strategy is such that the model matches the fat tail of firm size distribution documented in Gabaix (2011). However the reason why the model is able to do so is different from the usual explanation provided in the firm dynamics literature. In this section, we show numerically that the fat tail comes from the high persistence and the fat tail of the potential entrants distribution rather than Gibrat's law or imitation.

Power law distribution can come from proportional random growth also called Gibrat's law, as described in Gabaix (2009). The literature initiated by Luttmer (2007, 2010, 2012) build models that can generate fat-tailed distribution along a balanced growth path based on technology diffusion and spillover of technology by entrants. The flow of ideas literature (Lucas 2009, Lucas and Moll 2011 and Alvarez et al. 2008) also experiments growth and fat-tailed distribution along a balanced growth path.

In our framework, the fat-tailed distribution comes from a high persistence and large entrants. The persistence of the productivity process has an effect on the tail because with high persistence, firms benefit longer from a good shock and remain larger for a longer time than in a low persistence case. At the stationary equilibrium, there are

Parameters	Value	Description
ρ	0.98	Autocorrelation of firms level shocks
σ_e	0.22	Std of idio shocks
S	300	Number of productivity levels
$[\varphi_1, .., \varphi_S]$	$[-5.13, .., 5.13]$	Productivity grid
γ	2	Labor Elasticity
α	0.8	Production function
c_f	0.0073	Operating cost
c_e	0	Entry cost
β	0.95	Discount rate
M	3×10^6	Number of potential entrants
$\underline{\mathrm{L}}$	7.7×10^{-5}	Parameters of the labor supply function
G	Generalized Pareto(ζ,λ,0)	Entrant's distr. of $\exp(\varphi)$
ζ	0.76	Tail parameter of the distr. G
λ	0.70	Scale parameter of the distr. G
$\zeta(1-\alpha)$	0.152	Tail parameter of the entrants size distr.

Table 2: Baseline calibration

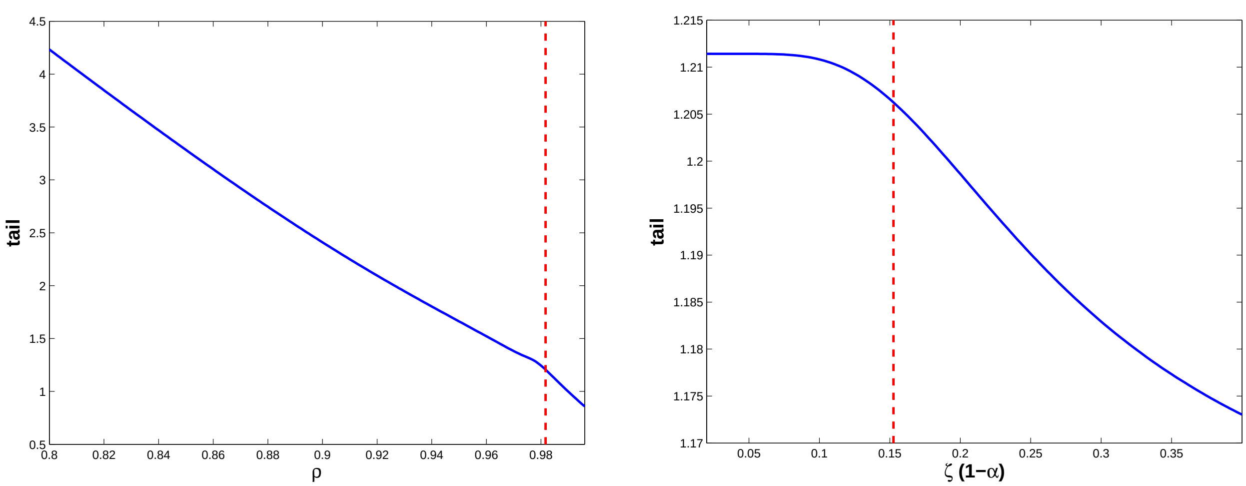

Figure 3: The tail of the stationary firm size distribution as a function of ρ (left) and $\zeta(1-\alpha)$, the tail of the potential entrant distribution G (right)

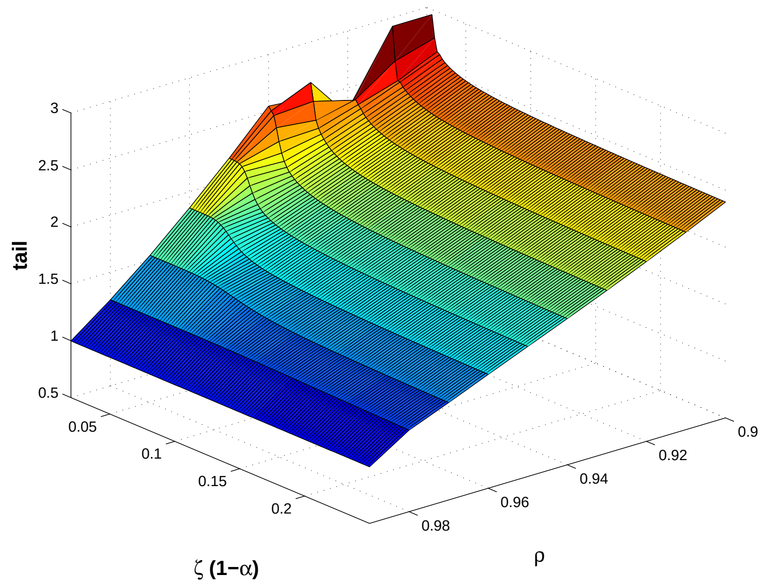

Figure 4: The tail of the stationary firm size distribution as a function of both ρ and $\zeta(1-\alpha)$, the tail of the potential entrant distribution G

more large firms and thus the tail is fatter. This can be seen in the left panel of figure 3: as the persistence ρ increases, the tail parameter decreases which means that the tail becomes fatter.

The fatter the distribution (i.e the smaller the tail index) of potential entrants, the fatter the stationary firm size distribution as one can see on the right panel of figure 3. This is pretty intuitive, since bigger entrants also implies bigger firms and thus a fatter tail of the firm size distribution.

A high persistence and a fat-tailed distribution are both necessary to get fat-tailed firm size distribution at the stationary equilibrium. This can be seen in figure 4 which plots the tail of the stationary firm size distribution as a function of both the persistence of the productivity process and the tail of the potential entrant distribution. The tail of the firm size distribution gets close to one only for high persistence and fat tail of the potential entrants distribution.

	Model				Data			
	$\sigma(x)$	$\frac{\sigma(x)}{\sigma(Y)}$	$\rho(x_t, Y_t)$	$\rho(x_{t+1}, x_t)$	$\sigma(x)$	$\frac{\sigma(x)}{\sigma(Y)}$	$\rho(x_t, Y_t)$	$\rho(x_{t+1}, x_t)$
Output	1.7	1.0	1.0	0.68	2.2	1.0	-	0.63
Hours	1.1	0.67	1.0	-	1.8	0.83	0.85	-
Agg. Pdty.	0.8	0.46	1.0	-	na	na	na	na

Table 3: Business Cycle Statistics

NOTE: These statistics are computed for the baseline calibration (cf. Table 2) for an economy simulated during 1,000 periods.

4.3 Business Cycle Statistics

After solving the model using the algorithm described in section C.1, we compute the business cycle statistics. We simulate time series for output, hours and aggregate TFP using the law of motion (2) of the productivity distribution. These statistics are presented in table 3.

First, the standard deviation of output is 1.7%, which is in line which represent 77% of actual output volatility. To assess the performance of the model in producing aggregate fluctuations without any aggregate shock, a better statistic is the volatility of aggregate TFP. For the baseline calibration, the standard deviation of aggregate productivity is 0.8%, which is non-negligible. We are able to generate about the same aggregate fluctuations as in the standard representative firm real business cycle (RBC) model with only idiosyncratic shocks and no aggregate shocks. The table 3 shows that a model with only idiosyncratic shocks can account for a large share of aggregate fluctuations.

The firm level dynamics of the baseline model is different from the one of a standard RBC model. In the latter all firms, or the representative firm, face similar productivity and have the same dynamics. In the baseline model, it is the aggregation of idiosyncratic shocks that gives rise to the aggregate fluctuations and thus each firm has its own dynamic.

The standard deviation of hours is 67% of the one of output whereas the same number is 83% in the data. Our model fails to match this number as it is the case in standard RBC model. It is the case because our model, without any aggregate shocks, evolves as a one-factor model. One could choose a higher value of the labor supply elasticity to match the relative volatility of hours but we choose to stay closer as possible to the RBC literature.

In the same simulated path, we also compute the first-order autocorrelation of output which is, for the baseline calibration, 0.68 whereas the corresponding number in the data is 0.63. Our model, calibrated to match firm level dynamics is able to explain some key statistics of output dynamics such as the output volatility and the first-order autocorrelation.

5 A Shock to the Largest Firm

In this section, we will describe the micro and macro consequences of a shock to the largest firm in the economy. First, we will describe the methodology and then the impulse response to such shock.

5.1 Methodology

We assume that the largest firm suffers a one standard deviation negative shock on its productivity. The standard deviation is σ_e and its calibrated value is 0.22. The initial "reshuffling" shock on the firm size distribution ϵ_0 is only a vector of zero where at the highest level we subtract one and add one to the level of productivity corresponding to a one standard deviation on the left. Figure 5 shows on the left panel the initial distribution (blue) and the shocked distribution (dashed red). The difference between the two distributions can hardly be seen on the graph. The right panel of figure 5 gives a closer view at the right end of the difference between the shock and the initial distribution. In this figure, one can see that a mass of firms has been moved from the highest level to the left by a one standard deviation. This figure displays the initial shock ϵ_0.

From the structure of the model, computing the impulse response is straightforward. As one can see in equation 2, the transition between date t and date $t+1$ firm size distribution is a linear operator. So after computing the initial "reshuffling" shock ϵ_0, we do not need to simulate a large number of paths and to take the average. Instead, we assume ϵ_t to be zero for $t \geq 1$ and thanks to the linearity of transition described by equation 2 the result is exactly the same.

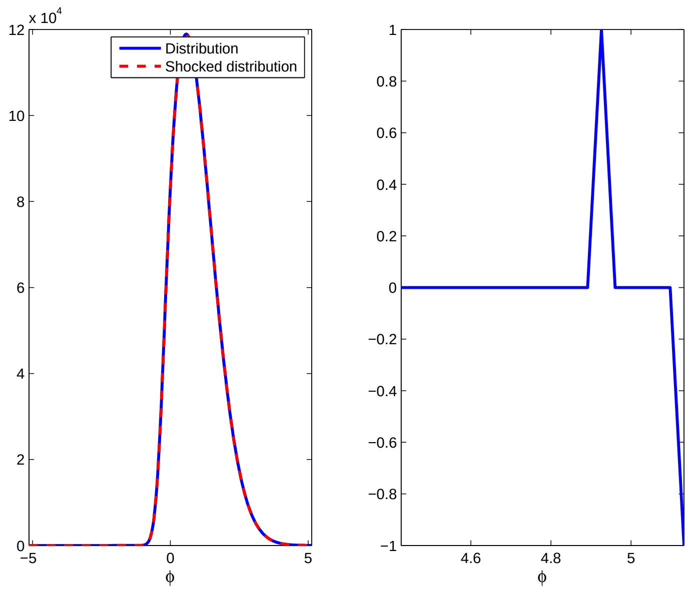

Figure 5: The shock and initial distribution (left panel) and the initial shock (right panel).

5.2 Impulse Responses

The top left panel of figure 6 shows the responses of aggregate productivity and aggregate output to this shock. Since our model behaves as a one-factor model, the dynamics of output follows closely the one of aggregate productivity. After this negative shock to the largest firm, aggregate productivity decreases by 0.05% which is non negligible. This decrease in aggregate productivity has a non-negligible effect on the aggregate output, which decreases by about 0.12% on impact. This decrease in output in turn decreases the aggregate labor demand and thus pushes the equilibrium wage down as can be seen in the top right panel of figure 6.

The intuition is as follows. Since the largest firm becomes less productive, its optimal size becomes smaller. Since there are no frictions in adjustment of labor, the largest firm shrinks by cutting its labor demand. Aggregate productivity, which is simply the weighted average of the productivity of each firm, decreases because one firm, the largest, becomes less productive. Due to the decrease in the wage, the second largest firm increases its size in response to the negative shock on the largest firm. Since the wage that this second largest firm is now facing is lower and since its productivity remains the same, its optimal size is now bigger. The output of the second largest firms increases as one can see on the bottom left panel of figure 6. Other firms benefit from this as well through a slacker labor market. This latter effect reduces the impact on aggregate output, since other firms produce more. In this exercice, if we keep the wage constant the output decrease is of 0.27%, whereas when the labor market mechanism is active, output decreases only by 0.12%. However, this effect is not strong enough to mitigate the decrease in aggregate output because labor is reallocated towards less productive firms. The overall effect of a negative shock on the largest firm is contractionary at the aggregate level.

In this setting firms that are subject to exit are the smallest firms. Because of the decrease in their cost, the wage, small firms can suffer bigger negative shocks while keeping their profitability positive. There are fewer exiters, so on impact the number of exiters decreases. In the meantime, entry becomes more profitable as the wage is smaller. The number of entrants becomes larger. However, there are more small entrants. Some potential entrants start producing whereas it would have not been the case if the wage had not decreased. The bottom right panel of figure 6 describes the dynamics of the number of incumbents, exiters and entrants. As the wage returns to its stationary value, the profitability of these entrants and the small firms that have survived, thanks to this decrease in the wage, shrinks. The number of small firms that are subject to exit is thus larger. Overall the number of exiters exibits a hump-shaped response to this negative shock on the largest firm. The number of exiters hump-shaped pattern implies an hump-shaped response of the number of incumbents.

Overall, we find that such a negative event on the largest firm productivity is contractionary at the macro level and expansionary at the micro level.

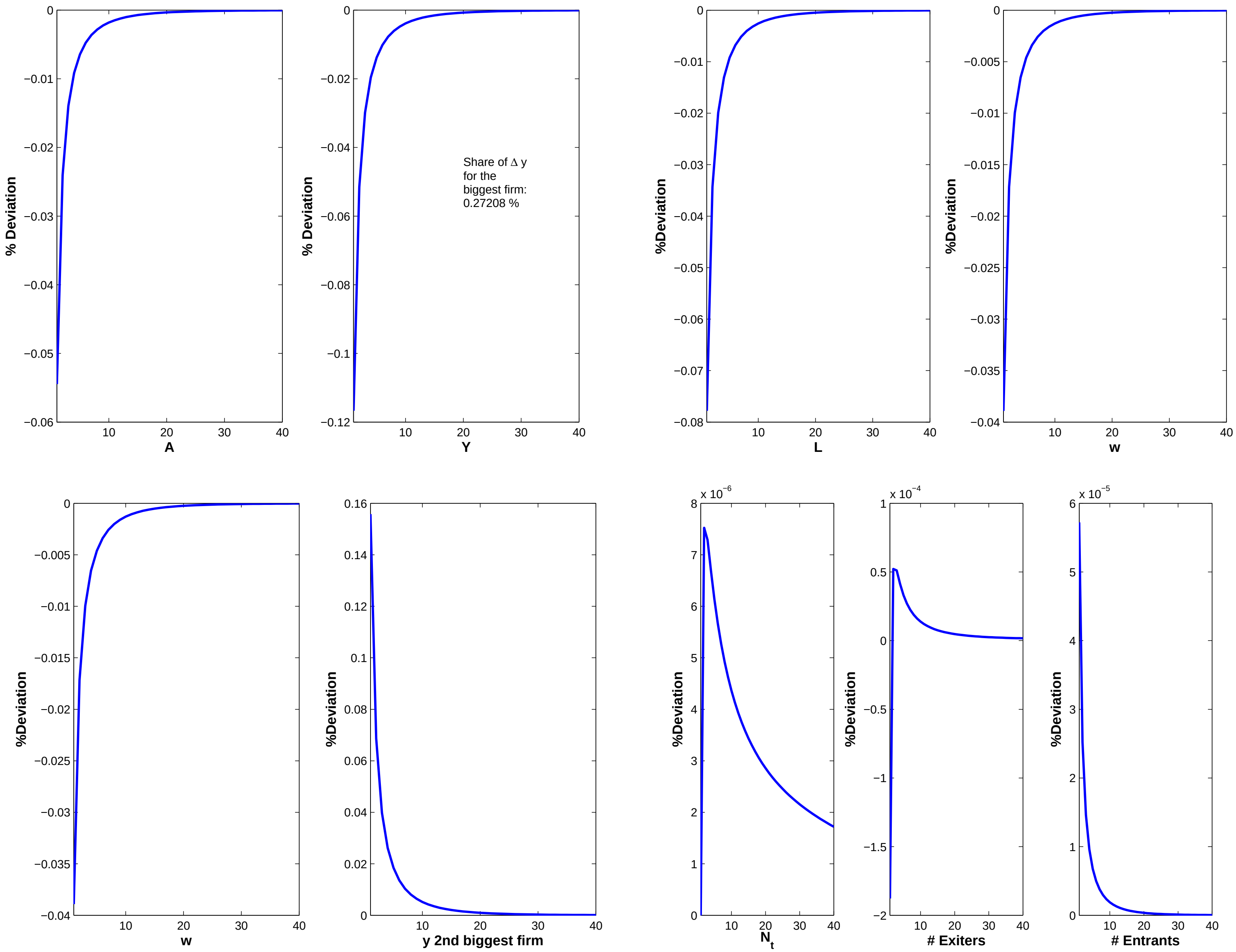

Figure 6: Impulse response to a one standard deviation negative productivity shock on the largest firm.

6 Large vs. Small Firms Over the Business Cycle

In this section, we study the co-movement of large and small firms over the business cycle. In particular, following a recent literature (see Moscarini and Postel-Vinay, 2012 and Chari, Christiano and Kehoe, 2013 for example), we analyse the correlation of the differential growth rate of large versus small firms with aggregate conditions. We first review some salient features of this comovement in the data and then ask whether our model can shed light on the observed correlations.

Moscarini and Postel-Vinay (2012) provide a first set of facts by looking at the cyclical behavior of labor markets. They focus on the differential behavior of net job creation - across the firm size distribution - over the business cycle. In the baseline case considered by Moscarini and Postel-Vinay (2012), small firms are defined as firms with less than 50 employees and large firms with more than 1000 employees. The differential net growth rate between large and small firms is then defined as the difference between the net job creation rate of large firms and that of of small firms, as given by Business Dynamics Statistics data of the U.S. Census[7]. Moscarini and Postel-Vinay (2012) find that the correlation between this differential growth rate and the deviation of unemployment from an HP trend is negative, significant and equal to -0.52. That is during good times, when unemployment is low, the differential growth rate is higher, i.e. large firms create more employment than small firms. Conversely, in recessions, relative to small firms, large firms destroy proportionally more jobs. We collect this statistic in Table 4 below, where $\Delta\widehat{g}_{t-1,t}$ gives the differential growth rate between large and small firms. For completeness, we also breakdown this number into the growth rate of large, $\widehat{g}_{t-1,t,LARGE}$ and small firms, $\widehat{g}_{t-1,t,SMALL}$.

A second set of correlations can be obtained by inspecting the behavior of large vs. small firms' sales over the business cycle, where the latter is defined by some measure of aggregate output. In an influential early paper, Gertler and Gilchrist (1994) concluded that small firms fall by more than large firms' sales in response to a contractionary monetary policy shock as identified by Romer-Romer dates. Chari, Christiano and Kehoe (2013) have recently revisited this evidence to conclude that, once the original data is extended to the last two decades, the original Gertler and Gilchrist's (1994) results

[7] Moscarini and Postel-Vinay (2012) also consider other data sources and size-cutoff definitions and show that their results are robust to these variations. Here we focus on their baseline results obtain with BDS data

	$\Delta\widehat{g}_{t-1,t}$	$\widehat{g}_{t-1,t,LARGE}$	$\widehat{g}_{t-1,t,SMALL}$
Large vs Small Employment (data)	-0.52 (0.00)	-0.45 (0.01)	-0.02 (0.93)
Baseline calibration (simulated data) Correlation with $-L_t$	-0.34 (0.00)	-0.31 (0.00)	0.19 (0.00)
Large vs Small Sales (data)	0.25 (0.01)	0.17 (0.01)	-0.04 (0.72)
Baseline calibration (simulated data)	0.29 (0.00)	0.34 (0.00)	-0.25 (0.00)

Table 4: Correlation of the differential growth rate between large and small firms and the cycle indicator (p-value in parenthesis) .

cannot be maintained. More importantly for our purposes here, Chari, Christiano and Kehoe (2013) also report that during NBER recessions - rather than Romer-Romer dates - large firms' sales fall by more than small firms' sales.

This suggests that the correlation between the differential growth rate of sales of large versus small firms and aggregate output is positive, i.e. that large firms contract their sales by relatively more than small sales in bad times and expand relatively more in good times. In order to ascertain whether this is the case, we follow Gertler and Gilchrist (1994) and Chari, Christiano and Kehoe (2013) in sourcing data from the Quarterly Financial Report (QFR) of the U.S. Census. This quarterly dataset reports nominal sales by asset size category. In what follows we report results for the period 1988Q1-2012Q4. We focus on the manufacturing sector and deflate the nominal quarterly sales value by the corresponding quarter's Producer Price Index for the manufacturing sector. Following the previous literature, we define small firms as those accounting for 30 percent of total sales. In particular, we define the sales of small firms as a weighted average of cumulative sales growth in the two asset size categories which straddle the 30th percentile of sales in a given quarter[8]. As a measure of aggregate business cycle conditions we use HP-filtered real aggregate output in the manufacturing sector. Figure 7 displays the results. We start by visually confirming the results in Chari, Christiano and Kehoe (2013) for our sample period: each of the last three NBER recessions is

[8]The weights are chosen so that the weighted average of cumulative sales in the two asset size categories average 30 percent of sales in a given quarter. This procedure is necessary as asset classes in the QFR are defined in nominal terms. As a result firms drift from low to nominal asset categories to high categories over time. See Chari, Christiano and Kehoe (2013) for further discussion and details on the construction of small vs. large firms sales growth

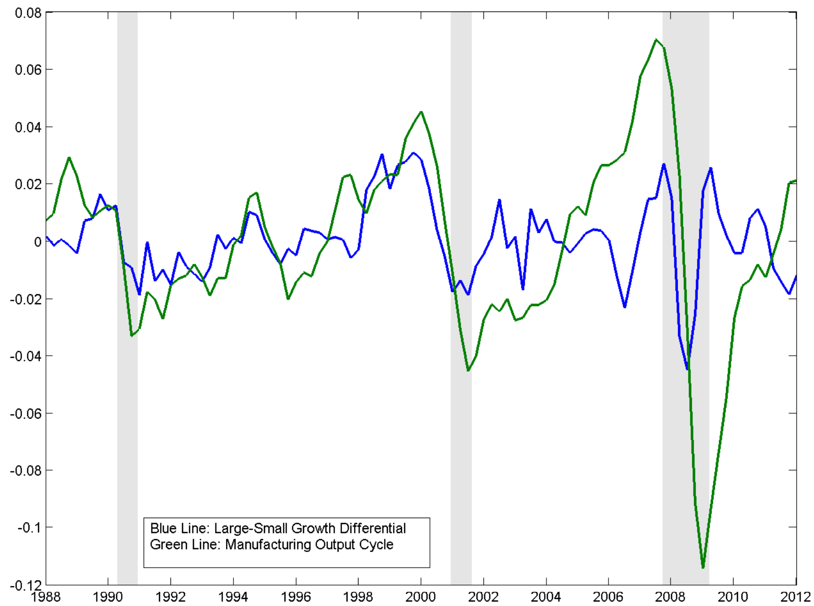

Figure 7: Growth rate of sales of large versus small firms (blue line) Deviation of aggregate manufacturing output from HP trend (green line). (Source: QFR, authors computation)

characterized by a sharp decline of the growth rate of large firms relative to small firms. Going beyond these results, we find that over the last quarter century, the large-small firm growth differential is positively and significantly correlated with the aggregate business cycle with a correlation of 0.25. In particular, while large firms' growth is positively correlated with the business cycle (statistically significant correlation of 0.17), small firms' growth is essentially uncorrelated with aggregate business cycle conditions (insignificant point estimate of -0.04). We again collect these correlations in Table 4 above.

We now ask whether the firm dynamics' setup introduced above can help shed light on the differential behavior of large versus small firms over the business cycle. To compute the corresponding statistics in our baseline model, we first simulate a path for the productivity distribution following equations 2 and 3. We then perform the same computations on our simulated data as those reported above for the employment and sales data. For the employment-based correlations, we apply Moscarini and Postel-

Vinay's (2012) definitions to compute growth rate of employment for large and small firms. That is, just like in the data, we track the simulated employment growth rates of firms with 1000 employes and more (large firms) and that of small firms with 50 employees or less. The closest counterpart to aggregate unemployment in our setting is (minus) aggregate employment ($-L_t$). For the output-based correlations, we track the simulated output growth rate of firms whose cumulated sales account for 70 percent of sales (i.e. large firms as per the Gertler and Gilchrist (1994) criterion) and that of firms that account for 30 percent of sales (i.e. small firms are defined by the same criterion as in the data).

We collect the simulation results in Table 4 and Figure 8. As in the data, we find that the large vs. small firms' growth rate differential is procyclical as it is in the data. This is true both for employment (0.34 and statistically significant) and output (0.28 and significant). In the model, during "good" times, when aggregate employment and output is high, large firms create more employment and increase production by more than small firms. In recessions, the converse happens. As in the data, this behavior is being led by a strong procyclical behavior of large firms[9]. The upshot of the model is that business cycle is led by large firms dynamics. If a small number of large firms face a higher productivity than usual, the economy expands. To the econometrician, it would thus seem as if large firms are more cyclically sensitive, i.e. they co-move more with the aggregate. In our setting, it is not that large firms are more cyclically sensitive. Rather it is the business cycle itself that reflects large firm dynamics.

7 Conclusion

We build a quantitative firm dynamics model in which we cast "the granular hypothesis". It features a finite number of firms and we do not rely on any law of large numbers assumption. Our model is able to generate aggregate fluctuations from idiosyncratic shocks only. We show analytically that aggregate fluctuations do not die out as the number of firms increases. A calibrated version of our framework to the US economy implies sizable aggregate volatility. We also look at the macro and micro effect of a

[9] Counterfactually however, note that in our baseline calibration, small firms shrink during good times which is at odds with the empirical evidence above

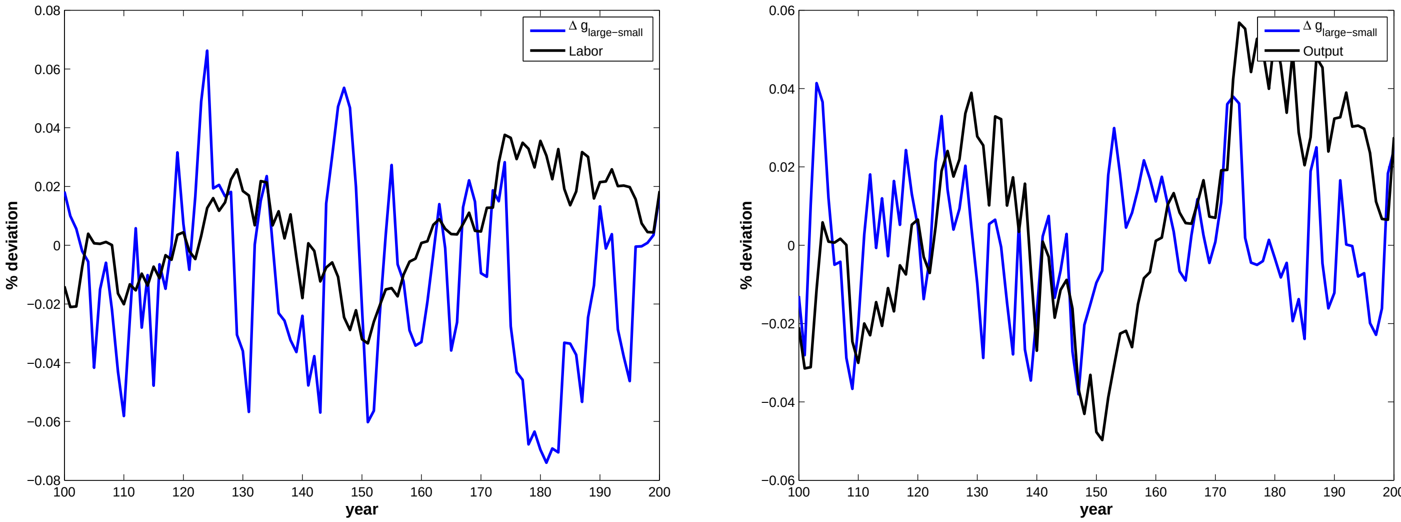

Figure 8: Left: Growth rate of employment of large versus small firms and aggregate employment. Right:Growth rate of output of large versus small firms and aggregate output. (Simulated data)

negative shock on the largest firm. Finally, we show that our model can reproduce the relative cyclicality of large firms compared to small firms'.

In our framework, a firm does not internalize its own effect on the aggregate wage. In other models of firms dynamics, the assumption that the law of large' numbers holds justifies thinking about firms as price takers. However, we have shown that in a standard firm dynamics framework large firms actually have an effect on the aggregate and thus the price taker assumption should be taken carefully. Our work is a first step toward this objective.

References

[1] Daron Acemoglu, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016, 09 2012.

[2] Fernando E. Alvarez, Francisco J. Buera, and Jr. Robert E. Lucas. Models of idea flows. NBER Working Papers 14135, National Bureau of Economic Research, Inc, June 2008.

[3] Ruediger Bachmann and Christian Bayer. Firm-specific productivity risk over the business cycle: Facts and aggregate implications. CESifo Working Paper Series 2844, CESifo Group Munich, 2009.

[4] Susanto Basu and John G Fernald. Returns to scale in u.s. production: Estimates and implications. *Journal of Political Economy*, 105(2):249–83, April 1997.

[5] Florin O. Bilbiie, Fabio Ghironi, and Marc J. Melitz. Endogenous entry, product variety, and business cycles. *Journal of Political Economy*, 120(2):304 – 345, 2012.

[6] Jeffrey Campbell. Entry, exit, embodied technology, and business cycles. *Review of Economic Dynamics*, 1(2):371–408, April 1998.

[7] Vasco M Carvalho. Aggregate fluctuations and the network structure of intersectoral trade. mimeo, 2010.

[8] Vasco M. Carvalho and Xavier Gabaix. The great diversification and its undoing. *American Economic Review*, Forthcoming, 2013.

[9] Rui Castro, Gian Luca Clementi, and Yoonsoo Lee. Cross-sectoral variation in firm-level idiosyncratic risk. *Journal of Industrial Economics*, Forthcoming, 2013.

[10] Varadarajan V. Chari, Lawrence J. Christiano, and Patrick J. Kehoe. The gertler-gilchrist evidence on small and large firm sales. mimeo, July 2013.

[11] Gian Luca Clementi and Dino Palazzo. Entry, exit, firm dynamics, and aggregate fluctuations. Working Paper Series 27 10, The Rimini Centre for Economic Analysis, January 2010.

[12] Wouter J. Den Haan. Assessing the accuracy of the aggregate law of motion in models with heterogeneous agents. *Journal of Economic Dynamics and Control*, 34(1):79–99, January 2010.

[13] Julian di Giovanni and Andrei A. Levchenko. Country size, international trade, and aggregate fluctuations in granular economies. *Journal of Political Economy*, 120(6):1083 – 1132, 2012.

[14] Julian di Giovanni, Andrei A. Levchenko, and Isabelle Mejean. Firms, destinations, and aggregate fluctuations. CEPR Discussion Papers 9168, C.E.P.R. Discussion Papers, October 2012.

[15] R. Durrett. *Probability: Theory and Examples*. Number v. 3 in Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010.

[16] Xavier Gabaix. Power laws in economics and finance. *Annual Review of Economics*, 1(1):255–294, 05 2009.

[17] Xavier Gabaix. The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772, 05 2011.

[18] Mark Gertler and Simon Gilchrist. Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms. *The Quarterly Journal of Economics*, 109(2):309–40, May 1994.

[19] Hugo A Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60(5):1127–50, September 1992.

[20] Aubhik Khan and Julia K. Thomas. Nonconvex factor adjustments in equilibrium business cycle models: do nonlinearities matter? *Journal of Monetary Economics*, 50(2):331–360, March 2003.

[21] Aubhik Khan and Julia K. Thomas. Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, 76(2):395–436, 03 2008.

[22] Karen Kopecky and Richard Suen. Finite state markov-chain approximations to highly persistent processes. *Review of Economic Dynamics*, 13(3):701–714, July 2010.

[23] Per Krusell and Anthony A. Smith. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896, October 1998.

[24] Yoonsoo Lee. The importance of reallocations in cyclical productivity and returns to scale: evidence from plant-level data. Working Paper 0509, Federal Reserve Bank of Cleveland, 2005.

[25] Yoonsoo Lee and Toshihiko Mukoyama. Entry, exit and plant-level dynamics over the business cycle. Working Paper 0718, Federal Reserve Bank of Cleveland, 2008.

[26] Robert E. Lucas. Ideas and growth. *Economica*, 76(301):1–19, 02 2009.

[27] Robert E. Lucas and Benjamin Moll. Knowledge growth and the allocation of time. NBER Working Papers 17495, National Bureau of Economic Research, Inc, October 2011.

[28] Erzo G. J. Luttmer. Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics*, 122(3):1103–1144, 08 2007.

[29] Erzo G.J. Luttmer. Models of growth and firm heterogeneity. *Annual Review of Economics*, 2(1):547–576, 09 2010.

[30] Erzo G.J. Luttmer. Technology diffusion and growth. *Journal of Economic Theory*, 147(2):602–622, 2012.

[31] Giuseppe Moscarini and Fabien Postel-Vinay. The contribution of large and small employers to job creation in times of high and low unemployment. *American Economic Review*, 102(6):2509–39, October 2012.

[32] K.G. Rouwenhorst. Asset pricing implications of equilibrium business cycle models. In T.F. Cooley, editor, *Frontiers of Business Cycle Research*, pages 294–330. Princeton University Press, 1995.

[33] George Tauchen. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, 20(2):177–181, 1986.

[34] Marcelo L. Veracierto. Plant-level irreversible investment and equilibrium business cycles. *American Economic Review*, 92(1):181–197, March 2002.

A Discretization of AR(1) by the Rouwenhorst Method

Consider the AR(1) process $\widetilde{\varphi}_{t+1} = \rho\widetilde{\varphi}_t + e_t$ where $e_t^i \rightsquigarrow \mathcal{N}(0, \sigma_e)$. We approximate this process by a discrete process $\{\varphi_t\}$ over the evenly distributed state space $\{\varphi_1, \dots, \varphi_S\}$. For $p, q \in (0, 1)$ and each S, we can defined recursively the matrix:

$$
\begin{aligned}
P_2 &= \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} \text{ if } S = 2 \\
P_S &= p \begin{pmatrix} P_{ns-1} & 0 \\ 0' & 0 \end{pmatrix} + (1-p) \begin{pmatrix} 0 & P_{ns-1} \\ 0 & 0 \end{pmatrix} + (1-q) \begin{pmatrix} 0 & 0 \\ P_{ns-1} & 0 \end{pmatrix} \\
&\quad + q \begin{pmatrix} 0 & 0 \\ 0 & P_{ns-1} \end{pmatrix} \text{ if } S > 2
\end{aligned}
$$

and normalizing all but the first and last rows in order that they sum to one. These matrix is a transition matrix for a discrete Markovian process. Rouwenhorst (1995) and Kopecky and Suen (2010) show that for $p = q = \frac{1+\rho}{2}$ and $\varphi_S = \sqrt{S-1}\frac{\sigma_e}{\sqrt{1-\rho^2}}$ we have:

$$
\begin{aligned}
\mathbb{E}(\varphi_{t+1} | \varphi_t = \varphi_s) &= \rho\varphi_s \\
\mathbb{V}ar(\varphi_{t+1} | \varphi_t = \varphi_s) &= \sigma_e^2
\end{aligned}
$$

Thus the Markovian process associated with the transition matrix P_S have a first order autocorrelation and a conditional variance equal to ρ and σ_e^2 respectively. Kopecky and Suen (2010) also shows that this discretization of an AR(1) is preferable for highly persistent process.

B Proofs

B.1 Proof of the Corollary 1

Let us divide both side of equation 2 by M, the number of potentials entrants, it yields:

$$
\hat{\mu}_{t+1} = (P_t^*)'(\hat{\mu}_t + G) + \frac{\epsilon_{t+1}}{M} \tag{7}
$$

Let us first shows that $\mathbb{V}ar\left[\frac{\epsilon_{t+1}}{M}\right] \to 0$ as M goes to infinity:

$$\mathbb{V}ar\left[\frac{\epsilon_{t+1}}{M}\right] = \frac{1}{M^2}\Sigma_t(\mu_t)] = \sum_{s=s^*(\mu_t)}^{S} \frac{MG^s + \mu_t^s}{M^2} W_s = \sum_{s=s^*(\mu_t)}^{S} \frac{G^s + \hat{\mu}_t^s}{M} W_s$$

$$\left|\mathbb{V}ar\left[\frac{\epsilon_{t+1}}{M}\right]\right| \leq \frac{1}{M} \sum_{s=s^*(\mu_t)}^{S} |(G^s + \hat{\mu}_t^s) W_s| \leq \frac{1}{M} \sum_{s=1}^{S} |(G^s + \hat{\mu}_t^s) W_s| \to 0$$

as M goes to infinity.

Let us call $E^M(\varepsilon, \hat{\mu})$ the pdf of the random vector $\frac{\epsilon}{M}$ for a given normalized productivity distribution $\hat{\mu}$ and a parameter M, the number of potential entrants. We have just shown that as M goes to ∞

$$\frac{\epsilon}{M} \longrightarrow^{L^2} 0$$

this implies that the convergence is also true in distribution:

$$E^M(., \hat{\mu}) \longrightarrow^{\mathcal{D}} \delta_0$$

where δ_0 is the dirac distribution with point mass zero.

Let us denote, for a given parameter M, the value of being an incumbent with a productivity level φ, given an aggregate state $\hat{\mu}$ to be $V^M(\hat{\mu}, \varphi)$. The incumbent problem writes, for a given M:

$$V^M(\hat{\mu}, \varphi) = \pi^*(\hat{\mu}, \varphi) + \max\left\{0, \beta \int_{\varepsilon'} \sum_{\varphi' \in \Phi} V^M\Big((P^*)'(\hat{\mu} + G) + \varepsilon, \varphi'\Big) F(\varphi'|\varphi) E^M(d\varepsilon, \hat{\mu})\right\} \tag{8}$$

P^* is the transition matrix P where the first $s^*(\hat{\mu}) - 1$ rows are replaced by zero, and where $s^*(\hat{\mu})$ is the threshold of the exit/entry rule when the number of potential entrants is M.

Let us defines the value function $V^\infty(\hat{\mu}, \varphi)$ as the solution of the incumbent problem when the number of firms is infinite:

$$V^\infty(\hat{\mu}, \varphi) = \pi^*(\hat{\mu}, \varphi) + \max\left\{0, \beta \sum_{\varphi' \in \Phi} V^\infty\Big((P^\infty)'(\hat{\mu} + G), \varphi'\Big) F(\varphi'|\varphi)\right\} \tag{9}$$

where P^∞ is the transition matrix P where the first $s^\infty(\hat{\mu})-1$ rows are replaced by zero, and where $s^\infty(\hat{\mu})$ is the threshold of the exit/entry rule when the number of potential entrants is infinite i.e. when[10]

$$\beta \sum_{\varphi' \in \Phi} V^\infty\Big((P^\infty)'(\hat{\mu}+G), \varphi'\Big) F(\varphi'|\varphi) = 0$$

The convergence in distribution of $E^M(., \hat{\mu})$ implies that the right hand side of equation 8 converges to the right hand side of equation 9 for a given $\hat{\mu}$ and φ. Assuming that $V^M(\hat{\mu}, \varphi) \to V^\infty(\hat{\mu}, \varphi)$, the convergence in distribution insure that $s^*(\hat{\mu}) \to s^\infty(\hat{\mu})$.

It is then clear that $(P^*)'(\hat{\mu}_t + G) \to (P^\infty)'(\hat{\mu}_t + G)$ and then for an infinite number of potential entrants the law of motion of the productivity distribution is

$$\hat{\mu}_{t+1} = (P_t^\infty)'(\hat{\mu}_t + G)$$

□

B.2 Proof of the Lemma 1

The law of motion of the firm size distribution as stated by equation 2 is $\mu_{t+1} = m(\mu_t) + \epsilon_t$ or

$$\mu_{t+1} = (P_t^*)'(\mu_t + MG) + \epsilon_t$$

The stationary distribution is defined as the solution of the following equation:

$$\mu = (P^*)'\mu + M(P^*)'G$$

Let us assume that $s_t^* = s^*$ and $\omega_t = \omega$ and thus that $P_t^* = P^*$ at each date. This assumption does not affect the results[11] and is just made for sake of simplicity of the proof.

Subtracting the two previous equations yields:

$$\mu_{t+1} - \mu = (P^*)'(\mu_t - \mu) + \epsilon_{t+1}$$

[10] Note that the following is not an equality, but $s^\infty(\hat{\mu})$ is such that the left hand side is positive when the $s^\infty(\hat{\mu})-1$ rows of P^∞ are zero and negative when $s^\infty(\hat{\mu})-2$ rows of P^∞ are zero.

[11] This assumption only affects the left-end of the firm size distribution and thus not the right tail.

Recall that at $t = 0$ the $\mu_0 = \mu$, let us iterate recursively on the above equation:

$$\begin{aligned}
\mu_{t+1} - \mu &= (P^*)'(\mu_t - \mu) + \epsilon_{t+1} \\
&= (P^{*'})^2(\mu_{t-1} - \mu) + P^{*'}\epsilon_t + \epsilon_{t+1} \\
&= \dots \\
&= (P^{*'})^{t+1}(\mu_0 - \mu) + \sum_{j=0}^{t}(P^{*'})^j \epsilon_{t+1-j} \\
\mu_{t+1} - \mu &= \sum_{j=0}^{t}(P^{*'})^j \epsilon_{t+1-j} \text{ since } \mu_0 = \mu \qquad (10)
\end{aligned}$$

We need to define the quantity that we interested in, the tail. The tail of the distribution is just the behavior of the counter cumulative mass function. By showing that two counter cumulative mass functions are similar, we show formally that the associated distribution have the same tail.

Let us define $\hat{\mu}_t = \frac{\mu_t}{M}$. The economy scale with M (assuming that the labor supply also increases[12]). When M goes to infinity so does μ_t^s for all s and t, but $\hat{\mu}_t^s$ does not. The counter cumulative mass function that we are interested in is thus:

$$\forall k \in [|1, S|] \qquad \left| \sum_{i=k+1}^{S} \hat{\mu}_{t+1}^i - \sum_{i=k+1}^{S} \hat{\mu}^i \right| = \left| \sum_{i=k+1}^{S} \frac{\mu_{t+1}^i}{M} - \sum_{i=k+1}^{S} \frac{\mu^i}{M} \right|$$

The left term in the left hand side equation is the counter cumulative mass function of $\hat{\mu}_{t+1}$, the right term is the same function for $\hat{\mu}$.

$$\begin{aligned}
\forall k \in [|1, S|] \qquad \left| \sum_{i=k+1}^{S} \hat{\mu}_{t+1}^i - \sum_{i=k+1}^{S} \hat{\mu}^i \right| &= \left| \sum_{i=k+1}^{S} \frac{\mu_{t+1}^i}{M} - \sum_{i=k+1}^{S} \frac{\mu^i}{M} \right| \\
&= \left| \sum_{i=k+1}^{S} \sum_{j=0}^{t} [\frac{(P^{*'})^j \epsilon_{t+1-j}}{M}]_i \right| \\
&= \left| \sum_{i=k+1}^{S} \sum_{j=0}^{t} \sum_{s=s^*}^{S} [\frac{(P^{*'})^j f_{t+1-j}^{\cdot,s}}{M} + \frac{(P^{*'})^j g_{t+1-j}^{\cdot,s}}{M} \right. \\
&\qquad \left. - \mathbb{E}\left[\frac{(P^{*'})^j f_{t+1-j}^{\cdot,s}}{M} + \frac{(P^{*'})^j g_{t+1-j}^{\cdot,s}}{M}\right]]_i \right|
\end{aligned}$$

[12]The wage is kept constant by increasing $\underline{L}$ as M increases: the ratio $M/\underline{L}$ is constant.

Soit $\varepsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\left|\sum_{i=k+1}^{S}\hat{\mu}^i_{t+1} - \sum_{i=k+1}^{S}\hat{\mu}^i\right| \geq \varepsilon\right)\\
&= \mathbb{P}\left(\left|\sum_{i=k+1}^{S}\sum_{j=0}^{t}\sum_{s=s^*}^{S}[\frac{(P^{*'})^j f^{\cdot,s}_{t+1-j}}{M} + \frac{(P^{*'})^j g^{\cdot,s}_{t+1-j}}{M} - \mathbb{E}\left[\frac{(P^{*'})^j f^{\cdot,s}_{t+1-j}}{M} + \frac{(P^{*'})^j g^{\cdot,s}_{t+1-j}}{M}\right]]_i\right| \geq \varepsilon\right)\\
&\leq \mathbb{P}\left(\sum_{j=0}^{t}\sum_{s=s^*}^{S}\left|\sum_{i=k+1}^{S}[\frac{(P^{*'})^j f^{\cdot,s}_{t+1-j}}{M} + \frac{(P^{*'})^j g^{\cdot,s}_{t+1-j}}{M} - \mathbb{E}\left[\frac{(P^{*'})^j f^{\cdot,s}_{t+1-j}}{M} + \frac{(P^{*'})^j g^{\cdot,s}_{t+1-j}}{M}\right]]_i\right| \geq \varepsilon\right)\\
&\leq \sum_{j=0}^{t}\sum_{s=s^*}^{S}\mathbb{P}\left(\left|\sum_{i=k+1}^{S}[\frac{(P^{*'})^j f^{\cdot,s}_{t+1-j}}{M} + \frac{(P^{*'})^j g^{\cdot,s}_{t+1-j}}{M} - \mathbb{E}\left[\frac{(P^{*'})^j f^{\cdot,s}_{t+1-j}}{M} + \frac{(P^{*'})^j g^{\cdot,s}_{t+1-j}}{M}\right]]_i\right| \geq \varepsilon\right)\\
&\leq \sum_{j=0}^{t}\sum_{s=s^*}^{S}\frac{1}{\varepsilon^2}\mathbb{V}ar\left[\sum_{i=k+1}^{S}[\frac{(P^{*'})^j f^{\cdot,s}_{t+1-j}}{M} + \frac{(P^{*'})^j g^{\cdot,s}_{t+1-j}}{M}]_i\right] \text{ by the Bienaymé-Chebyshev inequality}\\
&\leq \sum_{j=0}^{t}\sum_{s=s^*}^{S}\frac{1}{\varepsilon^2}\mathbb{V}ar\left[\sum_{i=k+1}^{S}[\frac{(P^{*'})^j f^{\cdot,s}_{t+1-j}}{M}]_i\right] + \mathbb{V}ar\left[\sum_{i=k+1}^{S}[\frac{(P^{*'})^j g^{\cdot,s}_{t+1-j}}{M}]_i\right] \qquad (11)\\
&\qquad \text{since the } f^{i,s}_t \text{ and } g^{i,s}_t \text{ are independent}
\end{aligned}$$

Recall that $f^{i,s}_{t+1-j} \rightsquigarrow Multi\Big(\mu^s_{t-j}, P_{s,\cdot}\Big)$ and $g^{i,s}_{t+1-j} \rightsquigarrow Multi\Big(MG^s, P_{s,\cdot}\Big)$.

It follows that:

$$\begin{aligned}
\mathbb{V}ar\left[\sum_{i=k+1}^{S}[\frac{(P^{*'})^j f^{\cdot,s}_{t+1-j}}{M}]_i\right] &= \mathbb{V}ar\left[\frac{(e_k)'(P^{*'})^j f^{\cdot,s}_{t+1-j}}{M}\right]\\
&= \mu^s_{t+1-j}\frac{(e_k)'(P^{*'})^j W_s (P^*)^j (e_k)}{M^2}\\
&= \hat{\mu}^s_{t+1-j}\frac{(e_k)'(P^{*'})^j W_s (P^*)^j (e_k)}{M} \longrightarrow 0
\end{aligned}$$

this last term goes to zero when M goes to infinity. Here, e_k is a $(S \times 1)$ vector with the first k elements equal to zero.

Since all the sums in equation 11 are finite, we can conclude that

$$\forall k \in [|1, S|], \forall \varepsilon > 0, \mathbb{P}\left(\left|\sum_{i=k+1}^{S}\hat{\mu}^i_{t+1} - \sum_{i=k+1}^{S}\hat{\mu}^i\right| \geq \varepsilon\right) \longrightarrow 0 \text{ as } M \to \infty$$

This just means that the counter cumulative mass function, the tail, of μ_{t+1} converge in probability toward the tail of μ as M goes to infinity. □

B.3 Proof of the Theorem 2

Let us first compute the aggregate output Y_t as a function of only T_t:

$$Y_t = \sum_{i=1}^{N_t} y_t^i = \sum_{i=1}^{N_t} \exp(\varphi_t^i)^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{w_t}\right)^{\frac{\alpha}{1-\alpha}} = \left(\frac{\alpha}{w_t}\right)^{\frac{\alpha}{1-\alpha}} T_t$$

Recall that:

$$w_t = \left(\alpha^{\frac{1}{1-\alpha}} T_t\right)^{\frac{1-\alpha}{\gamma(1-\alpha)+1}}$$

Substituting this expression of the wage in the latter equation and taking the growth rate yields:

$$Y_t = \alpha^{\frac{\alpha\gamma}{\gamma(1-\alpha)+1}} \left(T_t\right)^{1-\frac{\alpha}{\gamma(1-\alpha)+1}}$$

$$\frac{\Delta Y_{t+1}}{Y_t} \sim \left(1 - \frac{\alpha}{\gamma(1-\alpha)+1}\right) \frac{\Delta T_{t+1}}{T_t}$$

then we have that

$$\mathbb{V}ar_t\left[\frac{\Delta Y_{t+1}}{Y_t}\right] = \Omega\left(\mathbb{V}ar_t\left[\frac{\Delta T_{t+1}}{T_t}\right]\right) \tag{12}$$

Let us evaluate the variance of the growth rate of T_t. To do so, let us have a close look at this growth rate. Note that

$$T_{t+1} = \sum_{l \text{ successful incumbent at } t} \exp\left(\frac{\varphi_{t+1}^l}{1-\alpha}\right) + \sum_{e \text{ successful entrant at } t} \exp\left(\frac{\varphi_{t+1}^e}{1-\alpha}\right)$$

and

$$T_t = \sum_{l \text{ successful incumbent at } t} \exp\left(\frac{\varphi_t^l}{1-\alpha}\right) + \sum_{x \text{ exiters at } t} \exp\left(\frac{\varphi_t^x}{1-\alpha}\right)$$

Let us define $Z_t^l = \exp(\frac{\varphi_t^l}{1-\alpha})$ for l-successful incumbent at t, $E_{t+1}^e = \exp(\frac{\varphi_{t+1}^e}{1-\alpha})$ for e successful entrant at t and $X_t^x = \exp(\frac{\varphi_t^x}{1-\alpha})$ for x exiters at t.

The growth rate of T_t is:

$$\frac{\Delta T_{t+1}}{T_t} = \frac{1}{T_t}\left(\sum_l \Delta Z_{t+1}^l + \sum_e E_{t+1}^e - \sum_x X_t^x\right)$$

The variance of the Z_t^l follows from the productivity process. For a firm i, we have $\varphi_{t+1}^i = \rho\varphi_t^i + e_{t+1}^i$, where e_t^i is random variable with mean zero and variance σ_e^2. This is true since the first order autocorrelation and condition volatility is well defined and constant across the idiosyncratic state space. By taking the exponential and subtracting, this yields:

$$\exp(\frac{\varphi^i_{t+1}}{1-\alpha}) - \exp(\frac{\varphi^i_t}{1-\alpha}) = \exp(\frac{\varphi^i_t}{1-\alpha})^\rho \exp(\frac{e^i_{t+1}}{1-\alpha}) - \exp(\frac{\varphi^i_t}{1-\alpha})$$

Furthermore the e^i_t are i.i.d. across firms and time. Let us define $\sigma = \sqrt{\mathbb{V}ar\left(\exp(\frac{e^i_t}{1-\alpha})\right)}$. It follows that:

$$\mathbb{V}ar_t\left(\Delta Z^l_{t+1}\right) = (Z^l_t)^{2\rho}\sigma^2$$

and, by the same argument,

$$\mathbb{V}ar_t\left(E^e_{t+1}\right) = (E^e_t)^{2\rho}\sigma^2$$

where $E^e_t = \exp(\frac{\varphi^e_t}{1-\alpha})$ with φ^e_t the signal at t of the successful entrant e.

This leads to

$$\mathbb{V}ar_t\frac{\Delta T_{t+1}}{T_t} = \frac{\sigma^2}{(T_t)^2}\left(\sum_l (Z^l_t)^{2\rho} + \sum_e (E^e_t)^{2\rho}\right)$$

since the variance conditional on date t of X^x_t is equal to zero.

Denoting N^l_t, N^E_t and N^X_t the number of successful incumbents, successful entrants and exiters at date t respectively. According to the law of large number, we have:

$$(N^l)^{-1}\sum_l Z^l_t \to \mathbb{E}Z^l_t := \bar{Z}_t$$
$$(N^X)^{-1}\sum_x X^x_t \to \mathbb{E}X^x_t := \bar{X}_t$$

It is straightforward that

$$N^{-1}T_t \sim \frac{N^l_t}{N_t}\bar{Z}_t + \frac{N^X_t}{N_t}\bar{X}_t := \bar{I}_t$$

the average of incumbent size at date t (both successful and exiters).

The distribution of the random variable Z^l_t has a power law tail with parameters ξ, the tail parameter of firm size distribution (since only small firms exit). The distribution of the random variable E^e_t has a power law tail with parameters $\zeta' = \zeta(1-\alpha)$ the tail parameter of entrant size distribution (since only big entrants are successful).

Since $\xi/(2\rho) < 1$ and $\zeta'/(2\rho) < 1$ and using the Lévy theorem (Gabaix 2011 Durrett 2010), we have

$$(N_t^l)^{-2\rho/\xi} \sum_l (Z_t^l)^{2\rho} \to^d u$$
$$(N_t^E)^{-2\rho/\zeta'} \sum_e (E_t^e)^{2\rho} \to^d w$$

where u and w are standard Lévy distribution with parameters $\xi/2\rho$ and $\zeta'/2\rho$ respectively.

Computing the two above results yields

$$\begin{aligned}
\mathbb{V}ar_t \frac{\Delta T_{t+1}}{T_t} &\sim \sigma^2 N^{-2} \frac{\sum_l (Z_t^l)^{2\rho} + \sum_e (E_t^e)^{2\rho}}{(\bar{I}_t)^2} \\
&\sim \sigma^2 N^{-2} \frac{(N_t^l)^{2\rho/\xi}(N_t^l)^{-2\rho/\xi} \sum_l (Z_t^l)^{2\rho} + (N_t^E)^{2\rho/\zeta}(N_t^E)^{-2\rho/\zeta} \sum_e (E_t^e)^{2\rho}}{(\bar{I}_t)^2} \\
&\sim \sigma^2 N^{-2} \frac{(N_t^l)^{2\rho/\xi} u + (N_t^E)^{2\rho/\zeta} w}{(\bar{I}_t)^2} \\
&\sim \frac{\sigma^2}{(\bar{I}_t)^2} \left[\frac{u}{(N_t)^{2-2\rho/\xi}} \left(\frac{N_t^l}{N} \right)^{2\rho/\xi} + \frac{w}{(N_t)^{2-2\rho/\zeta}} \left(\frac{N_t^E}{N_t} \right)^{2\rho/\zeta} \right]
\end{aligned}$$

Finally, by combining equation 12 with the one above, we have

$$\sigma(\frac{\Delta Y_{t+1}}{Y_t}) = \Omega \left(\sigma \left[\frac{u}{(N_t)^{2-2\rho/\xi}} \left(\frac{N_t^I}{N_t} \right)^{2\rho/\xi} + \frac{w}{(N_t)^{2-2\rho/\zeta}} \left(\frac{N_t^E}{N_t} \right)^{2\rho/\zeta} \right]^{1/2} \right)$$

□

B.4 Proof of the Theorem 3

Let us show that asymptoticaly (when the number of firms goes to infinity), we have

$$XXX$$

The left hand side is a measure of auto-correlation, because if T_{t+1} follows an AR(1) with autocorrelation parameter ρ_T, then $\mathbb{E}_t \left[\frac{T_{t+1} - T_t}{T_t} \right] = \rho_T - 1$.

We have shown in theorem 2's proof that the growth rate of T_t is:

$$\frac{\Delta T_{t+1}}{T_t} = \frac{1}{T_t}\left(\sum_l \Delta Z_{t+1}^l + \sum_e E_{t+1}^e - \sum_x X_t^x\right)$$

and that

$$\Delta Z_{t+1}^l = \exp(\frac{\varphi_t^l}{1-\alpha})^\rho \exp(\frac{e_t^l}{1-\alpha}) - \exp(\frac{\varphi_t^l}{1-\alpha})$$

It follows that

$$\mathbb{E}_t\left[\Delta Z_{t+1}^l\right] = \exp(\frac{\varphi_t^l}{1-\alpha})^\rho \exp(\frac{\sigma_e^2}{2(1-\alpha)^2}) - \exp(\frac{\varphi_t^l}{1-\alpha})$$

With the same reasonning, we also have:

$$\mathbb{E}_t\left[E_{t+1}^e\right] = (E_t^e)^\rho \exp(\frac{\sigma_e^2}{2(1-\alpha)^2})$$
$$\mathbb{E}_t\left[X_t^x\right] = X_t^x$$

At the first order for σ_e small enought, we have $\exp(\frac{\sigma_e^2}{2(1-\alpha)^2}) \sim 1$, thus

$$\mathbb{E}_t\left[\frac{\Delta T_{t+1}}{T_t}\right] = \frac{1}{T_t}\left(\sum_l \left(Z_t^l\right)^\rho - \sum_l Z_t^l + \sum_e \left(E_t^e\right)^\rho - \sum_x X_t^x\right) \tag{13}$$

Since Z_t^l has a Pareto tail with parameter ξ, it is clear that $\left(Z_t^l\right)^\rho$ has a Pareto tail with parameter ξ/ρ. Since we assumed that $1 < \xi < 2$ and $0 < \zeta'/\rho < 1$ the Lévy theorem (Gabaix 2011, Durrett 2010) says that:

$$(N_t^I)^{-1/\xi}\left[\sum_l (Z_t^l) - N_t^I \bar{Z}_t\right] \longrightarrow^d v$$

and

$$(N_t^E)^{-\rho/\zeta'}\left[\sum_e (E_t^e)^\rho\right] \longrightarrow^d u$$

The weak law of large number says that

$$(N_t^I)^{-1}\sum_l (Z_t^l)^\rho \longrightarrow^{\mathcal{P}} \bar{Z}^\rho{}_t$$
$$(N_t^X)^{-1}\sum_x X_t^x \longrightarrow^{\mathcal{P}} \bar{X}_t$$
$$(N_t)^{-1} T_t \longrightarrow^{\mathcal{P}} \bar{I}_t$$

where N_t^I, N_t^E, N_t^X, N_t are respectively the number of successful incumbents, of successful entrants, of exiters and of all incumbents at date t. For a firm level variable F_t^i at date t, $\bar{F}_t$ is the mean over the cross-section at date t.

Let us multiplying the bracket in the right hand side of equation 13 by $(N_t)^{-1}$:

$$\begin{aligned}
(N_t)^{-1}\mathbb{E}_t\left[\Delta T_{t+1}\right] &= \frac{N_t^I}{N_t}(N_t^I)^{-1}\sum_l (Z_t^l)^\rho - (N_t)^{-1}(N_t^I)^{1/\xi}(N_t^I)^{-1/\xi}\sum_l (Z_t^l) \\
&\ldots + (N_t)^{-1}(N_t^E)^{\rho/\zeta'}(N_t^E)^{-\rho/\zeta'}\sum_e (E_t^e)^\rho - \frac{N_t^X}{N_t}(N_t^X)^{-1}\sum_x X_t^x \\
&\sim \frac{N_t^I}{N_t}\bar{Z^\rho}_t - (N_t)^{-1}(N_t^I)^{1/\xi}\left[v + (N_t^I)^{1-1/\xi}\bar{Z}_t\right] \\
&\ldots + (N_t)^{-1}(N_t^E)^{\rho/\zeta'}u - \frac{N_t^X}{N_t}\bar{X}_t \text{ using the above results} \\
&\sim \frac{N_t^I}{N_t}\left(\bar{Z^\rho}_t - \bar{Z}_t\right) - \frac{N_t^X}{N_t}\bar{X}_t - \frac{v}{N_t^{1-1/\xi}}\left(\frac{N_t^I}{N_t}\right)^{1/\xi} + \frac{u}{N_t^{1-\rho/\zeta'}}\left(\frac{N_t^E}{N_t}\right)^{\rho/\zeta'}
\end{aligned}$$

Since $\xi > 1$, The weak law of large number also says that

$$(N_t)^{-1}T_t \longrightarrow^{\mathcal{P}} \bar{I}_t$$

Gathering the above computations together yields

$$\begin{aligned}
\mathbb{E}_t\left[\frac{\Delta T_{t+1}}{T_t}\right] &= \frac{(N_t)^{-1}\mathbb{E}_t\left[\Delta T_{t+1}\right]}{(N_t)^{-1}T_t} \\
&\sim \frac{N_t^I}{N_t}\frac{\bar{Z^\rho}_t - \bar{Z}_t}{\bar{I}_t} - \frac{N_t^X}{N_t}\frac{\bar{X}_t}{\bar{I}_t} - \frac{v/\bar{I}_t}{N_t^{1-1/\xi}}\left(\frac{N_t^I}{N_t}\right)^{1/\xi} + \frac{u/\bar{I}_t}{N_t^{1-\rho/\zeta'}}\left(\frac{N_t^E}{N_t}\right)^{\rho/\zeta'}
\end{aligned}$$

□

C Numerical Solution

C.1 Numerical Solution Algorithm

This section describes the algorithm used to solve numerically the full model.

The state variable of this model is only the distribution of productivity $\mu \in \mathbb{R}_+^S$. Since S is large, following the evolution of the distribution μ across time is not computationally feasible. To solve this model we use an algorithm similar to Krusell and Smith (1998), where we follow the evolution of the factor that matters for all the aggregate variables, namely T_t defined as:

$$T_t = \sum_{s=1}^{S} \mu_t^s \exp(\varphi_s)^{\frac{1}{1-\alpha}} = B'.\mu_t$$

where B is the $(S \times 1)$ vector $(\exp(\varphi_1)^{\frac{1}{1-\alpha}}, \ldots, \exp(\varphi_S)^{\frac{1}{1-\alpha}})$.

The true evolution of T_t is:

$$T_{t+1} = B'.m(\mu_t) + B'.\epsilon_{t+1}$$

or

$$T_{t+1} = B'.m(\mu_t) + \sqrt{B'.\Sigma(\mu_t).B}\varepsilon_{t+1}$$

where ε_{t+1} is drawn from a standard univariate normal distribution. At the first order, the process followed by $\log(T_t)$ is:

$$\log(T_{t+1}) = \log(B'.m(\mu_t)) + \frac{\sqrt{B'.\Sigma(\mu_t).B}}{B'.m(\mu_t)}.\varepsilon_{t+1}$$

Assuming the following approximations:

$$\begin{aligned}\log(B'.m(\mu_t)) &= \alpha_0 + \alpha_1 \log(T_t) + w_t \\ \frac{\sqrt{B'.\Sigma(\mu_t).B}}{B'.m(\mu_t)} &= \beta_0 + \beta_1 \log(T_t) + v_t\end{aligned}$$

leads to the following approximate law of motion of T_t:

$$\log(T_{t+1}) = \alpha_0 + \alpha_1 \log(T_t) + \beta_0.\epsilon_{t+1} + \beta_1.\epsilon_{t+1} \log(T_t) + u_t \tag{14}$$

where u_t,v_t and w_t are error terms.

This approximate law of motion is used to compute expectations as in Krusell and Smith (1998). The coefficients are updated using estimation of this equation for a simulated series. We iterate until convergence. The algorithm is formally described bellow:

1. Guess some parameters $\alpha_0^0, \alpha_1^0, \alpha_2^0, \beta_0^0$ and β_1^0.

2. Solve jointly for the value function of an individual firm and the exit rule for all (φ, T) using the approximation law of motion 14 to compute the expectation.

3. Simulate a series of $\{T_t, \epsilon_t\}_{t=0\ldots T}$ as follows:

    (a) Given a μ_0, compute $s^*(T_0)$ and $\omega(T_0)$ from the solution of step 2

    (b) Draw a ϵ_1 vector from its true distribution and use it to compute μ_1, and T_1, and using the law of motion of the productivity distribution described in equation (2).

    (c) Iterate from step 3a

4. Using the above simulated series, estimate the approximating rule 14.

5. Iterate from step 2 until convergence.

C.2 Accuracy of the Forecasting Rule

This section describes the numerical solution of the model by using the algorithm presented in section C.1 for the calibration in table 2. At the last stage of the algorithm, the approximate law of motion for the state T_t is estimated to be:

$$\log(T_{t+1}) = 1.8649 + 0.8383\log(T_t) + 1.4849\epsilon_{t+1} + u_t \tag{15}$$

The R^2 of the last step of the algorithm for this approximate law of motion is 0.9958. However, it has been shown in Den Haan (2010) that this is not enough to assert the quality of the approximate law of motion. Figure 9 reproduces both the simulated path of $\log T_t$ and the path of the approximate law of motion (15). Despite some minor differences, one can see that the two paths coincide.

In this appendix, we assess the accuracy of the forecastin rule used in our Krusell and Smith (1998) algorithm. As Den Haan (2010) points out, evaluating the accuracy of the forecasting rule by giving the R^2 of the regression in the last iteration of the algorithm is questionable. The root mean aquared error (in our case 0.0148) is better measure that the R^2 since the former is not scale by the variance of the dependent variable.

Following Den Haan (2010), we compute the maximum discrepancy (in absolute value) between the actual T and the one generated by the rule without updating. The value of this statistics is 0.1134% of the actual T. In left panel of figure 10, we report the frequency distribution of percentage forecasting errors. We also report the scatter plot

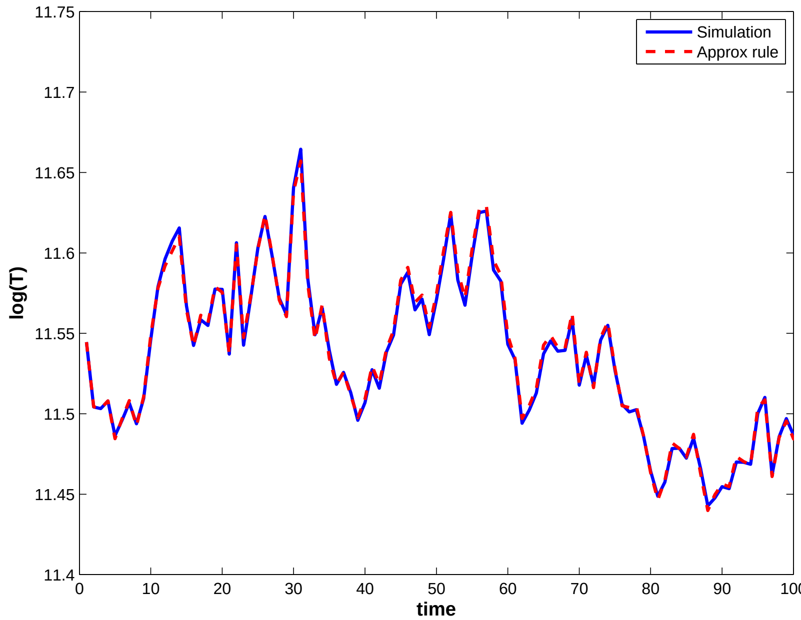

Figure 9: Simulated paths of the true and approximate evolution of T_t

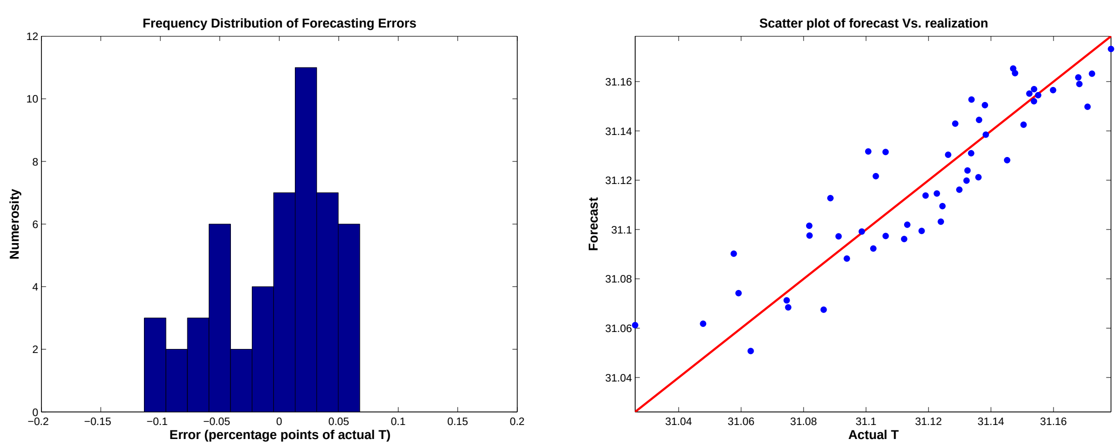

Figure 10: Accuracy of the forecasting rule. Left: frequency distribution. Right: Scatter plot of forecasting errors against actual realtization.

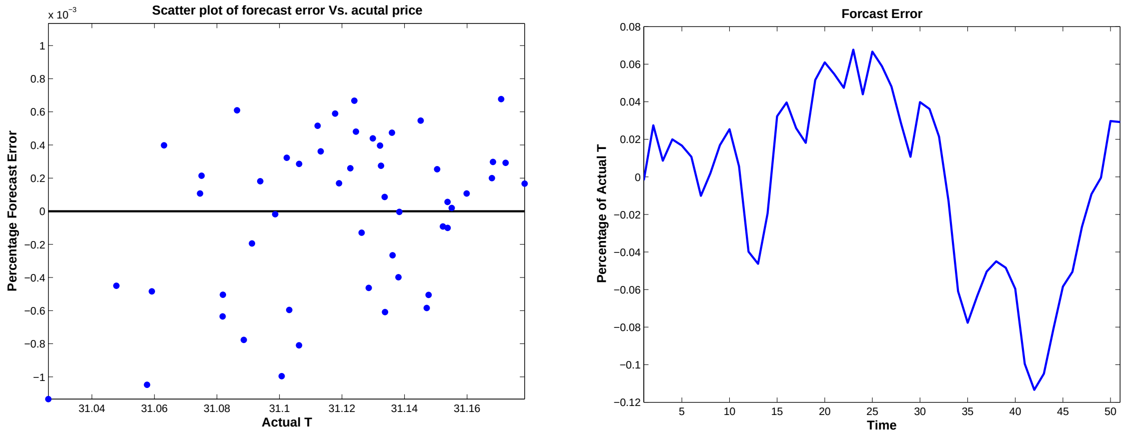

Figure 11: Accuracy of the forecasting rule. Left: frequency distribution. Right: Scatter plot of forecasting errors against actual realtization.

of the forecasting errors against the actual realtization. The points are along the 45 degree line.

The mean error of forecasting errors is -2.6449×10^{-05}, this indicates that forecastings errors are unbiased. They are also uncorrelated with the series of actual T -the correlation is -0.0208. In the figure 11, we display both the scatter plot of forecasting errors against actual realtization of T (left) and the time series of forecasting errors (right).