

The Persistent Effects of Entry and Exit

Aubhik Khan

The Ohio State University

Tatsuro Senga

Queen Mary, University of London, RIETI and ESCoE

Julia K. Thomas

The Ohio State University and NBER

February 2018

ABSTRACT

We develop a model with endogenous entry and exit in an economy subject to aggregate total factor productivity shocks that are non-stationary. Firms exhibit a life-cycle consistent with data and our model economy reproduces both their size and age distribution. In this setting, persistent shocks to aggregate total factor productivity growth rates endogenously drive long term reductions in business formation. The economic consequences of this persistent decline in entry grows over time.

In our model, individual firms vary in both the permanent and transitory components of their total factor productivity and in their capital stock. Capital adjustment is subject to one period time-to-build and involves both convex and nonconvex costs. Our dynamic stochastic general equilibrium model involves an aggregate state that includes a distribution of firms over total factor productivity and capital. Changes in this distribution, following aggregate shocks to the common component of TFP, drive persistent fluctuations in aggregate economic activity. We show that equilibrium movements in firms' stochastic discount factors, following persistent shocks to TFP growth, imply long-run declines in the value of entry. The resulting fall in the number of firms propagates a reduction in economic activity. This slows down the recovery.

We apply our model to understanding the last decade of economic activity in the U.S. This period began with a large recession followed by a period of slow economic growth. At the same time there was a persistent reduction in the new business formation. Our dynamic stochastic general equilibrium analysis is consistent with relatively small reductions in the level of total factor productivity, as seen in the last recession, and large reductions in GDP and Business Fixed Investment. Moreover, the recovery from such a large recession is slowed by a persistent reduction in firm entry.

Aubhik Khan: mail@aubhik-khan.net, Tatsuro Senga: tatsuro.senga@gmail.com, Julia Thomas: mail@juliathomas.net.

1 Introduction

Recent economic experience after the Great Recession, including the lack of recovery of output to its trend, declining business dynamism, and the productivity slowdown, provoked extensive concern among policymakers and researchers over the possibility of secular stagnation. Meanwhile, official growth forecasts have been downwardly revised repeatedly, creating uncertainty about the growth rate of the U.S. economy. How does such a growth-rate uncertainty affect firm dynamics at the micro-level? What is the extent to which the productivity slowdown in the aggregate came in response to this micro factor? A growing literature highlights the importance of uncertainty at the firm-level and its impact on firm dynamics and aggregate productivity (see Bloom (2009) and Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2016) for the seminal works). This proposal aims to extend the canonical heterogeneous firm framework to allow for shocks to the growth rate of aggregate total factor productivity (TFP), and firm entry and exit. Using this model to contemplate the possibility of secular stagnation, we assess the relevance of such growth-rate shocks in accounting for declining business dynamism and its impact on the productivity slowdown.

The last few years have seen a decline in business dynamism. For example, the birth rate of new establishments has been on a persistent decline after the Great Recession: falling from 12.3 percent in 2006 to 10.0 percent in 2014. This is a concern not only because the contribution of start-up firms to job creation is large, but also because it may lead to a long-lasting impact on the aggregate economy.¹ For example, Sedláček and Sterk (2016) show that start-up firms incorporated during recessions remain persistently smaller even at their later stage of growth trajectory.² Such cohort effects can have a long-run effects on aggregate output and employment; in fact, Gourio, Messer, and Siemer (2016) use state-level data in the U.S. and find that the decline of new firm formation leads to a significant and persistent drop in real GDP. As argued by Gourio et al. (2016), two questions should be answered: (1) what are the causes of the declining firm entry? and (2) what are the

¹See Haltiwanger, Jarmin, and Miranda (2013).

²Moreira (2017) finds such cohort effects on firm size using data taken from the Census Bureau's Longitudinal Business Database (LBD).

aggregate consequences? This paper studies the impact of uncertainty about future growth rates of the economy on firm entry and its aggregate implication like a persistent decline in GDP and productivity. A quantitative analysis of the effects of these must be consistent with rich features of firm dynamics as shown by the above studies, and importantly, our model proposed below is distinct in that it contains a unique non-Gaussian process of firm-level productivity, endogenous entry and exit, and capital adjustment costs, allowing us to capture all of these elements.

A defining feature of our model, on top of the rich heterogeneity of firms seen in the data, is the inclusion of growth-rate shocks to aggregate TFP. This is motivated by the unusually anemic recovery of the U.S. economy from the Great Recession trough. Over time, there has been increasing pessimism about future economic growth. One clear example is the IMF's outlook for real GDP growth in the U.S. which has been downwardly revised repeatedly, indicating persistent movements in expected growth rates and uncertainty surrounding them. Despite the dire need for a quantitative business cycle research examining this and its impact on firm dynamics and aggregate productivity dynamics, the canonical business cycle model is ill-suited to analyze this as it relies on shocks to the level of productivity, not the growth rate. While the focus is on long-run economic growth, another line of research has produced many endogenous growth models that highlight the role of firms and their innovative activity (see Aghion, Akcigit, and Howitt, 2014 for a recent survey). For example, Acemoglu, Akcigit, Bloom, and Kerr (2017) highlight the rich heterogeneity in innovation activity using micro data and build model of firm-level innovation and endogenous entry and exit to show the importance of selection in shaping aggregate productivity. To our best knowledge, ours is the first business cycle model that can be used to examine a mechanism through which changes in growth expectation influence business dynamism and macroeconomic series like business investment and aggregate productivity.

In the model, a negative shock to total factor productivity growth is propagated by changes in the distribution of firms. A fall in TFP implies a decline in the value of production. Firms expect a persistent fall in earnings. As a result, the onset of a recession is accompanied by a rise in exit and a very persistent reduction in entry. Our emphasis on matching the size and age distribution of employment across firms implies that the

aggregate effects of the resulting decline in the number of firms grows over time.

In the model, any cohort of entrants increases their share of total production gradually. Entrants begin small and, on average, experience trend growth in their individual total factor productivity. Initially low levels of firm-specific productivity, in presence of capital adjustment costs, implies slow growth in average firm size within any cohort of entrants. When in a recession there are several periods of entry below average, the economy fails to offset the reduction in the number of firms resulting from the increase in exit.

The entry and exit of firms adds an extensive margin absent in standard heterogeneous firm models focussing on capital adjustment such as Bloom (2009) or Khan and Thomas (2008). Given decreasing returns to scale at each individual firm, changes in the scale of production implies endogenous movements in aggregate total factor productivity. Fewer firms, given an aggregate level of resources, implies lower aggregate TFP. As the importance of a cohort of firms grows over time, with their productivity and capital, the effects of reduced entry are felt several periods after the onset of a recession. This leads to a very slow recovery. Thus a period of low entry propagates a recession, and measured TFP remains low when the exogenous component of aggregate TFP begins to recover.

A missing generation of firms created by growth rate uncertainty is a distinct aspect of our model, and we evaluate its quantitative importance in shaping aggregate productivity dynamics.³ The literature highlighted the relevance of uncertainty for the longer-run decisions such as investment and R&D. In fact, Barrero, Bloom and Wright (2017) find that investment and R&D are significantly more responsive to long-run uncertainty than hiring. Our results are consistent with their empirical finding; indeed, since creating new firms is so forward looking and uncertain, persistent fluctuations in expected growth rates play an extremely important role.⁴

Our inclusion of shocks to the growth rate of productivity is new in a heterogeneous firm framework.⁵ This is both motivated by the anemic recovery seen in the data, and

³See Siemer (2016) for a quantitative study how financial frictions plays a role in adversely affecting firm entry and its aggregate implications.

⁴See, for example, Garcia-Macia (2017) and Schmitz (2014) for studies on the long-run consequences of the Great Recessions through the lens of investment activities in productivity-enhancing technology.

⁵See Favilukis and Lin (2013) for a model with representative firm.

also aimed at addressing a few challenges in a quantitative macroeconomic research using a heterogeneous firm framework. First, the growing literature highlighted the importance of growth-rate shocks, instead of level shocks, in accounting for aggregate dynamics (see Nakamura, Sergeyev, Steinsson, 2017). Ours is the first to look at firm dynamics as a micro-origin that plays an crucial role in amplifying growth-rate shocks in driving macroeconomic fluctuations.⁶

2 Model

Our model has three types of economic agents: a representative household and a large number of perfectly competitive, heterogeneous firms, and a perfectly competitive representative investment goods producer.

Firms are heterogeneous in capital, k , and in their idiosyncratic TFP, ε , which is the product of a persistent random variable, ϵ , and a long-run productivity component, s , determined at entry, and x . Young firms initially operate with $x < 1$ and have a constant probability of it permanently rising to 1.

The aggregate state of the economy involves a time-varying distribution of firms over their idiosyncratic TFP and capital, μ , and the level of exogenous total factor productivity, $A = A_{-1}z$, where z is the trend shock and follows a Markov Chain, $\left\{z_i, \left\{\pi_{ij}^z\right\}_{j=1}^{N(z)}\right\}_{i=1}^{N(z)}$. Exploiting that value functions for households and firms are homogeneous of degree one in A_{-1} , we detrend the model by A_{-1} and thus the aggregate state of the economy becomes (z, μ) .

2.1 Households

A representative household values consumption, owns all firms and supplies labor. The household receives aggregate dividends from production firms and has access to a full set of Arrow Securities. We assume Epstein-Zin preferences for this household. The house-

⁶See Kung and Schmid (2014).

hold's expected lifetime value is $W = \left((1 - \beta) C^{1-\sigma} + \beta \left(E(W')^{1-\gamma} \right)^{\frac{1-\sigma}{1-\gamma}} \right)^{\frac{1}{1-\sigma}}$, where γ is the coefficient of risk aversion and $\frac{1}{\sigma}$ is the intertemporal elasticity of substitution. The detrended problem for the representative household can be written as follows:

$$h(a, z_i, \mu) = \max_{c, \{a'_j\}_{j=1}^{N(z)}} \left((1 - \beta) c^{1-\sigma} + \beta \left(\sum_{j=1}^{N(z)} \pi_{ij} \left(z_i h(a'_j, z_j, \mu') \right)^{1-\gamma} \right)^{\frac{1-\sigma}{1-\gamma}} \right)^{\frac{1}{1-\sigma}} \quad (1)$$

subject to

$$c + \sum_{j=1}^{N(z)} q_j(z_i, \mu) z_i a'_j \leq w(z_i, \mu) n(z_i, \mu) + \pi(z_i, \mu) \quad (2)$$

$$c \geq 0, a'_j \geq 0 \text{ and } \mu' = \Gamma(z_i, \mu),$$

where we used $W(a, z, A_{-1}, \mu) = h(a, z, \mu) A_{-1}$ and $c A_{-1} = C$. Our parametrization will have that $\gamma > \sigma$, so that the household has a preference for early resolution of uncertainty and thus would not want to face uncertainty about long-run growth rates.

In recursive equilibrium, detrended household consumption will be a function of the aggregate state, $c = c(z, \mu)$. Furthermore, equilibrium in the markets for contingent claims involves $a'_j = 0$ for each $j = 1, \dots, N(z)$. As a result, the equilibrium price of Arrow Securities is given by

$$q_j(z_i, \mu) = \frac{z_i^{1-\sigma} h(0, z_j, \mu')^{\sigma-\gamma}}{\left(\left(\sum_{j=1}^{N(z)} \pi_{ij} \left(h(0, z_j, \mu') \right)^{1-\gamma} \right)^{\frac{1}{1-\gamma}} \right)^{\sigma-\gamma}} \beta \frac{\pi_{ij} c(z_j, \mu')^{-\sigma}}{c(z, \mu)^{-\sigma}}. \quad (3)$$

An important component of our dynamic stochastic equilibrium model is to have firms discount their future value using (3).

2.2 Production

Production involves two sets of agents: heterogeneous goods producers and the capital producer.

Goods producers: Goods producers produce a homogeneous output, y , using capital, k , and labor, n , using a decreasing returns to scale technology, $y = (z\varepsilon)^{1-\alpha} k^\alpha n^\nu$, where z is

the trend shock to the aggregate TFP, A , and follows a Markov Chain, $\left\{z_i, \left\{\pi_{ij}^z\right\}_{j=1}^{N(z)}\right\}_{i=1}^{N(z)}$ and ε is idiosyncratic TFP. We denote π_{ij}^z the probability of $z' = z_m$ conditional on $z = z_l$. We also denote π_{mn}^ε the probability of $\varepsilon' = \varepsilon_n$ conditional on $\varepsilon = \varepsilon_m$.⁷

The timing of firm decisions is as follows. Each firm at the beginning of each period is identified by its capital stock carried from the previous period, k , and its current firm-level TFP, ε . It observes the aggregate state of the economy (z, μ) and hires labor from households at real wage $\omega(z, \mu)$ and production takes place and the firm pays its wage bill and operating cost. After production, it faces a random fixed cost of continuing into the following period, $\xi_c \sim \vartheta_c(\xi_c; s(\varepsilon))$. Conditional on continuing, the firm chooses its future capital, k' . Capital adjustment is subject to both convex, $\phi(k', k)$, and fixed, $\xi_k \sim \vartheta_k(\xi_k; s(\varepsilon))$, costs. Both these distribution vary by firms long-run idiosyncratic productivity component and the function $s(\varepsilon) = s$ where $\varepsilon = \varepsilon s$ and s is this long-run level of productivity. A firm avoids the convex cost when it pays its current random ξ_k . Their capital depreciates at rate $\delta \in (0, 1)$.

Capital producers: A competitive firm produces investment goods subject to diminishing marginal productivity of investment spending. It produces new capital goods at a total cost $\Phi(K', K)$ which it sells at price $p(z, \mu)$. This makes existing capital an input into the production of new investment goods and introduces a time-varying price of capital. After production, firms rent their capital to the investment firm at a price $d(z, \mu)$.

Let m be detrended capital, $m A_{-1} = K$. A competitive investment firm sells $z_i m' - (1 - \delta)m$ units of capital goods, $z_i m' = (1 - \delta)m + i$ at the competitive price $p(z_i, \mu)$. It rents capital from goods producers at the price $d(z_i, \mu)$. Deriving $\psi(z_i m', m)$ as the total cost of producing $z_i m' - (1 - \delta)m$ units of the capital good using the production function Φ . The firm solves the following

$$\max_{m', m} \left(p(z_i, \mu) (z_i m' - (1 - \delta)m) - \psi(z_i m', m) - d(z_i, \mu)m \right)$$

⁷Note that $\varepsilon = \varepsilon s$.

2.3 Firm values

Each firm that operates is constrained by a capacity to use labour, $n \leq n(\varepsilon, k, z)$. As explained below, in equilibrium, the real wage will imply that these capacity constraints do not bind and aggregate employment is procyclical in a model where the household does not value leisure. Let v_2 describe the continuation value of a firm, which requires paying ξ_c . An exiting firm avoids paying this fixed cost and receives the value of its non-depreciated capital. Given (ξ_c, ξ_k) , a firm's detrended problem is

$$v_1(\xi_c, \xi_k, \varepsilon_m, k, z_i, \mu) = \max_{n \leq n(\varepsilon, k, z)} \left((z_i \varepsilon_m)^{1-\alpha} k^\alpha n^\nu - w(z_i, \mu) n \right) + d(z_i, \mu) k + \max \left\{ v_2(\xi_k, \varepsilon_m, k, z_i, \mu) - w(z_i, \mu) \xi_c, p(z_i, \mu) (1 - \delta) k \right\}. \quad (4)$$

Taking expectations over ξ_c and ξ_k , let v be the start-of-period value,

$$v(\varepsilon_m, k, z_i, \mu) = \int_0^\infty \int_0^\infty v_1(\xi_c, \xi_k, \varepsilon_m, k, z_i, \mu) \vartheta_k(d\xi_k; s(\varepsilon_m)) \vartheta_c(d\xi_c; s(\varepsilon_m)). \quad (5)$$

The expected discounted future value of firm that chooses k' , when the future endogenous aggregate state is μ' , is defined using (5),

$$v_f(\varepsilon_m, k', z_i, \mu') \equiv \sum_{j=1}^{N(z)} \pi_{ij} \rho(z_j, z_i, \mu) \sum_{n=1}^{N(\varepsilon)} \pi_{mn}^\varepsilon v(\varepsilon_n, k', z_j, \mu'). \quad (6)$$

Above, $q_j(z_i, \mu) = \pi_{ij} \rho(z_j, z_i, \mu)$ with $q_j(z_i, \mu)$ given by (3).

Continuing firms choose between nonconvex and convex capital adjustment costs.

$$v_2(\xi_k, \varepsilon_m, k, z_i, \mu) = p(z_i, \mu) (1 - \delta) k + \max \left\{ \max_{k'} \left(-w(z_i, \mu) \xi_k - z_i p(z_i, \mu) k' + v_f(\varepsilon_m, k', z_i, \mu') \right), \max_{k'} \left(-p(z_i, \mu) z_i k' - \phi(z_i k', k) + v_f(\varepsilon_m, k', z_i, \mu') \right) \right\} \quad (7)$$

2.4 Firm entry

We assume a time-invariant stock of production locations, x , which may be active with firms producing, on inactive and available to potential entrants. Let $M(\mu) = \int \mu(d[\varepsilon \times k])$

represent the number of firms operating at the start of this period. Since there is no fixed cost to current production, all these firms hire labour and produce. The stock of potential entrants for entry at the start of next period is $\theta - M(\mu)$.

The number of firms of type s that may enter is $\omega(s)(\theta - M(\mu))$, $s \in \mathbf{S}$. Firms in each type, s , face a random fixed cost of entry which is drawn from a time-invariant distribution, $\xi_s \sim \vartheta_s(\xi_s)$.

$$v^e(s, k', z_i, \mu') \equiv \sum_{j=1}^{N(z)} \pi_{ij} \rho(z_j, z_i, \mu) \sum_{n=1}^{N(\varepsilon)} \pi_n^0 v(\varepsilon_n, k', z_j, \mu') . \quad (8)$$

A firm enters if

$$-w(z_i, \mu) \xi_s + \max_{k'} -z_i p(z_i, \mu) k' + v^e(s, k', z_i, \mu') \geq 0 . \quad (9)$$

We will assume that the distribution of entry costs, ξ_s is unbounded above, hence the number of entrants of each type will always be less than potential entrants. Let $\xi_s(s, z_i, \mu)$ represent the threshold entry cost determined by

$$w(z_i, \mu) \xi_s(s, z_i, \mu) = \max_{k'} \left(-z_i p(z_i, \mu) k' + v^e(s, k', z_i, \mu') \right) . \quad (10)$$

The number of entrants of type s will be

$$\vartheta_s(\xi_s(s, z_i, \mu)) \omega(s)(\theta - M(\mu)) \quad (11)$$

2.5 Equilibrium

We begin by explaining the mechanics of the real wage and firms' employment capacity. Choosing a form for the capacity constraint that's proportional to firms' unconstrained labour demand, we set a real wage function that clears the labour market and ensures that firms' labour choices are unconstrained. Given households do not value leisure, this clear the labour market.

Given there are no fixed costs of production, all firms that exist with capital in place at the start of a period will produce. However, they productively use more labour than the

level determined by their employment capacity function. The optimal choice of employment solves

$$\begin{aligned} & \max_n \left((z_i \varepsilon_m)^{1-\alpha} k^\alpha n^\nu - w(z, \mu) n \right) \\ & \text{st.} \\ & n \leq n(\varepsilon_m, k, z). \end{aligned}$$

If the capacity constraint does not bind, then

$$n = \left(\frac{\nu (z_i \varepsilon_m)^{1-\alpha} k^\alpha}{w(z, \mu)} \right)^{\frac{1}{1-\nu}} \quad (12)$$

Denote this as $N(z_i, \varepsilon_m, k, w)$.

Assume the capacity constraint takes the form,

$$n(\varepsilon_m, k, z) = \varphi_0^{\frac{1}{1-\nu}} z_i^{\frac{\gamma_1}{1-\nu}} \left(\nu (z_i \varepsilon_m)^{1-\alpha} k^\alpha \right)^{\frac{1}{1-\nu}}. \quad (13)$$

This form implies that it either binds for all firms, or none. Now, using (12) in (13),

$$n(\varepsilon_m, k, z) = \varphi_0^{\frac{1}{1-\nu}} z_i^{\frac{\gamma_1}{1-\nu}} w(z, \mu)^{-\frac{1}{1-\nu}} N(z_i, \varepsilon_m, k)$$

If $\varphi_0^{\frac{1}{1-\nu}} z_i^{\frac{\gamma_1}{1-\nu}} w(z, \mu)^{-\frac{1}{1-\nu}} = 1$ then firms are unconstrained as $n(\varepsilon_m, k, z) = N(z_i, \varepsilon_m, k)$.

In this case, capacity does not constrain employment at any firm and all firms choose their unconstrained optimal level of employment.

As households do not value leisure, the highest equilibrium wage that will have all firms operating at capacity and maximise employment is

$$w(z, \mu) = \varphi z_i^{\gamma_1}. \quad (14)$$

This wage is procyclical if $\gamma_1 > 0$, and independent of μ . As long as $1 - \alpha - \gamma_1 > 0$, employment will be procyclical at firms.

In *recursive competitive equilibrium*, households' value functions solve (1), existing firms' value functions solve (4) - (7), entry is determined by (11) with $\xi_s(s, z_i, \mu)$ solving (10), firms discount factors are consistent with (3) and the real wage is given by (14).

3 Results

We choose a set of parameter values that illustrate the model's ability to replicate empirical patterns seen in the age and size distribution of firms. The distribution of capital adjustment costs, ϑ_k is assumed to be uniform with an upper bound of 0.005 and a lower bound of 0. The distribution of continuation costs ϑ_c is log-normal with a mean of -1.5 and a variance of 2. Entry costs are also log-normally distributed with a mean of 0.75 and the same variance. These costs are scaled, for each type, s , by as^b where $a = 0.03$ and $b = 0.75$. This ensures more productive firms from ignoring fixed costs in their decisions.

We discretise types into 5 levels using a truncated Pareto distribution with bounds $[1.5, 2.5]$. This Pareto distribution has a shape parameter of 1.6 and a scale parameter equal to 0.375. We set the initial productivity for a type, relative to its long-run level, to be $x_0 = 0.75$ and that, in each period, there is a 10 percent chance of reaching the long run level of $x_1 = 1$. Finally, the persistent idiosyncratic shock is log-normal with a persistence of 0.757 and a standard deviation of innovations of 0.0499. This is discretised into three levels, hence the composite idiosyncratic state has 30 values; 5 long-run types, s , 2 levels of x and 3 levels of ϵ .

Firm level capital adjustment, in the absence of fixed costs, involves the cost function

$$\phi(z_k k', k) = a_2 \left(\frac{z_k k'}{k} - 1 \right)^{a_3} k.$$

We set $a_3 = 3$ and $a_2 = 0.00002$. These convex costs apply when the firm does not pay $\xi_k^{s(\epsilon)}$. The investment goods producing firm faces diminishing marginal productivity of producing capital goods. It's production function is

$$\Phi \left(\frac{z_k k' - (1 - \delta) k}{k} \right) k = a_0 + \frac{a_1}{\omega} \left(\frac{z_k k' - (1 - \delta) k}{k} \right)^\omega k.$$

We set $\omega = 0.5$ and then a_0 and a_1 are chosen to imply that the relative price of capital goods, $p(z, \mu)$ is 1 in the steady state and no adjustment costs.

Lastly we set $\sigma = 2/3$ and $\gamma = 5$, which implies a Sharpe ratio, for a representative agent version of our model, that is similar to the empirical value. Given these values for the inverse of the elasticity of intertemporal substitution and the coefficient of relative risk

aversion, β is chosen to imply a 4 percent annual risk free real interest rate in the steady state where TFP grows at 1.6 percent per year.

The stationary distribution has a capital-to-output ratio of 2.1 and investment to GDP ratio of 0.18 and aggregate hours worked that are 1/3 of the households' time endowment. We adjust the number of potential entrants to imply a unit measure of firms in production. The entry and exit rate is 8 percent. This is close to the average, in the BDS, over the period 1984 - 2015, where the exit rate of all firms was 8.5 percent.

The model captures the smaller relative size of entrants, though not to the extent in the data. In the BDS, over the period 1982 - 2015, the relative size of entrants compared to all firms, measured in terms of the number of employees, was 27.68 percent. In the model the corresponding value is 38.3 percent. Examining a cohort of firms aged 0 to 5, the model predicts a relative size of 0.56 while the corresponding value in the BDS is 37.5. Thus young firms are small, in the model, and grow gradually over time. This is the result of assuming all firms, independent of their long run mean productivity, s , begin with xs where $x = 0.75$. Our assumption that each firm with $x < 1$ has a 10 percent change of maturing in any period is too high. As a result, cohorts grow too fast.

Turning to the size distribution, over the years 1982 - 2014, the BDS shows the following distribution.

	small	medium	large
number of employees	1 – 19	20 – 499	< 500
share of firms	88.30	11.36	0.37
share of employment	19.67	31.94	48.46

A Pareto distribution of permanent productivity types helps the model match this skewness in the size distribution of firms. However the largest 0.4 percent of firms has only 8.4 percent of employment while the next 11.4 percent of firms have 44 percent of employment. As seen above, the corresponding numbers for the data are 48.5 and 32 percent, respectively. Thus the distribution of long-term productivity, s , needs to be more skewed than is assumed in our present example. Nonetheless, in this example, the largest

12 percent of firms employ more than half of the labour force. This is dramatically more skewed than would be implied by the conventional log-normally distributed productivity shock process chosen to reproduce the volatility of firm level output or investment.

The recursive equilibrium of the model is characterised by an aggregate state with a high-dimensional object, the distribution of firms over productivity and capital. As is well known, such models must be numerically solved using some type of aggregate state space approximation. Here, we proxy for the distribution with a one-dimensional object as is common in the quantitative analysis of dynamic stochastic equilibrium in heterogeneous agent economies. However, instead of following the common approach of Krusell and Smith (1999) applied to complete markets models with a distribution of firms by Khan and Thomas (2008), we solve our model using the Backwards Induction method of Reiter (2010). This eliminates the need for a simple forecasting rule. As a result, we are able to avoid long simulations of the model which are costly to solve.

Entry and exit, implying a time-varying number of firms, presents a challenge over and beyond incomplete markets models with a constant number of households. Our approach is to replace μ , the distribution of firms over (ϵ, x, k) with the average capital per firm, $m = K/M(\mu)$. This aggregate state space approximation is used with a non-parametric law of motion agents use to forecast the future proxy aggregate state. In Backwards induction, in contrast to the Krusell and Smith method, the aggregate law of motion and individual decision rules are solved simultaneously. Over a grid on the approximate aggregate state (z, m) we forecast a future state pointwise, $m' = g(z, m)$. Given this forecast, we solve for household and firm value functions. This implies an actual end of period distribution and an actual future approximate aggregate state, m'_1 . We iterate on the forecast until $m'_0 = m'_1$. This requires a proxy distribution $\mu(\epsilon, x, k; z, m)$ at each point (z, m) . Given this proxy distribution we can aggregate individual decision rules into m' .

Proxy distributions are derived from simulations of the model, and must be moment consistent. Thus the distribution of firms associated with (z, m) must have $m = \frac{\int k \mu(d[\epsilon \times k]; z, m)}{M(\mu(\epsilon, k; z, m))}$. We present preliminary results below involving a stochastic process for z with $N(z) = 3$. This is a discretisation of a log-normal continuous process with a persistence of 0.5 and a

standard deviation of innovations of 0.0012.

In the table below we present HP-filtered results, using a weight of 400, of a 500 period simulation of the model. All series are in levels, and the reported moments are percent standard deviations and contemporaneous correlations with output. Aggregate fluctuations are driven by growth shocks to aggregate total factor productivity, and propagated by responses in the aggregate investment, in the distribution of this investment across firms, and in entry and exit over the distribution of firms itself. This leads to nonlinear dynamics in the model following a positive shock to tfp.

In a representative firm model, or a heterogeneous firm model without entry and exit, there would be a monotone response in output and investment, and a hump-shaped response in consumption. In the present model, the familiar forces that propagate a shock are present. Persistence in TFP leads to a substitution from current to future consumption and there is a rise in investment. Consumption smoothing implies that this increase in capital is gradual over time. As a result, the variability of the shock is less than the variability of GDP.

In our model with entry and exit, a persistent shock to TFP growth leads to a rise in entry and a decline in exit. As seen below, entry is weakly procyclical, as in the data, and exit is more pronouncedly countercyclical. The number of firms in production is then procyclical, driven by a rise in entry and a fall in exit. However, capital adjustment costs and the gradual rise in individual firms' productivity, over time, imply that a cohort of entrants grows in importance over time. As the tfp growth shock decays, the effect of its initial rise is propagated by a persistent increase in several generations of firms whose economic importance rises gradually. In our model with decreasing returns to scale at the level of an individual firm, this increase in the scale of production leads to a prolonged increase in measured tfp over and beyond that implied by the shock. This lengthens the expansion.

	GDP	Investment	Employment	Wage	Entry	Exit	Firms
std. dev.	1.96	10.32	2.11	1.17	1.37	2.05	0.57
correlation	1.0	0.47	0.84	0.08	0.38	-0.78	0.51

It is worth noting that the persistent rise in the entry rate is very persistent in the model, with an autocorrelation of 0.54. This gradualism in entry is important in introducing persistence through dynamics in the distribution of firms. Importantly, it is the result of persistent equilibrium changes in real interest rates.

References

- [1] Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., & Kerr, W. R. (2017). Innovation, Reallocation, and Growth. Stanford mimeo.
- [2] Aghion, P., Akcigit, U., & Howitt, P. (2014). What do we learn from Schumpeterian growth theory?. In Handbook of economic growth (Vol. 2, pp. 515-563). Elsevier.
- [3] Barrero, J. M., Bloom, N., & Wright, I. (2017). Short and long run uncertainty (No. w23676). National Bureau of Economic Research.
- [4] Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3), 623-685.
- [5] Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., & Terry, S. J. (forthcoming). Really uncertain business cycles. *Econometrica*
- [6] Favilukis, J., & Lin, X. (2013). Long run productivity risk and aggregate investment. *Journal of Monetary Economics*, 60(6), 737-751.
- [7] Garcia-Macia, D. (2017). The financing of ideas and the great deviation. International Monetary Fund.
- [8] Gourio, F., Messer, T., & Siemer, M. (2016). Firm entry and macroeconomic dynamics: A state-level analysis. *American Economic Review*, 106(5), 214-18.
- [9] Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics*, 95(2), 347-361.
- [10] Khan, A. and J. K. Thomas (2008). Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, 76(2), pages 395–436.
- [11] Kung, H., & Schmid, L. (2015). Innovation, growth, and asset prices. *The Journal of Finance*, 70(3), 1001-1037.
- [12] Krusell, Per and Anthony A. Smith (1998) ‘Income and Wealth Heterogeneity in the Macroeconomy’ *Journal of Political Economy*, 106(5), pages 867 - 896.

- [13] Moreira, S. (2017). Firm Dynamics, Persistent Effects of Entry Conditions, and Business Cycles. Working Papers 17-29, Center for Economic Studies, U.S. Census Bureau.
- [14] Nakamura, E., Sergeyev, D., & Steinsson, J. (2017). Growth-rate and uncertainty shocks in consumption: Cross-country evidence. *American Economic Journal: Macroeconomics*, 9(1), 1-39.
- [15] Reiter, Michael (2010) ‘Solving the Incomplete Markets Model with Aggregate Uncertainty by Backwards Induction’ *Journal of Economic Dynamics & Control* 34, pages 28–35.
- [16] Schmitz, T. (2014). Fluctuations in R&D investment and long-run growth: The role of the size distribution of innovating firms. manuscript, Bocconi University.
- [17] Sedláček, P., & Sterk, V. (2017). The growth potential of startups over the business cycle. *American Economic Review*, 107(10), 3182-3210.
- [18] Siemer, M. (2016). Firm entry and employment dynamics in the great recession. Working Paper.