Entry, Exit and the Shape of Aggregate Fluctuations in a General Equilibrium Model with Capital Heterogeneity

Gian Luca Clementi
Stern School of Business, New York University

Aubhik Khan
Ohio State University

Berardino Palazzo
Boston University School of Management

Julia K. Thomas
Ohio State University

March 2015

ABSTRACT

We study the cyclical implications of endogenous firm-level entry and exit decisions in a dynamic, stochastic general equilibrium model wherein firms face persistent shocks to both aggregate and individual productivity. Firms' decisions regarding entry into production and their subsequent continuation are affected not only by their expected productivities, but also by the presence of convex and nonconvex capital adjustment costs, and hence their existing stocks. Our model is unique relative to other DSGE settings in that our incumbent firms face two discrete choices, one involving continuation and one on capital adjustment. As such, we can explore how age, size and selection reshape macroeconomic fluctuations in an equilibrium environment with realistic firm life-cycle dynamics and investment patterns.

Examining standard business cycle moments and impulse responses, we find that changes in entry and exit rates and the age-size composition of firms amplify responses over a typical business cycle driven by a disturbance to aggregate productivity somewhat and, to a greater extent, protract them. Both results stem from an endogenous drag on TFP induced by a *missing generation effect*, whereby an usually small number of entrants fails to replace an increased number of exitors. This effect is most injurious several years out as the reduced cohorts of young firms approach maturity. Declines in the number of firms, and most notably in the numbers of young firms, were dramatic over the U.S. 2007 recession. In an exercise designed to emulate that unusual episode, we consider a second shock that more directly affects entry and the exit decisions of young firms. We find that it sharpens the missing generation effect, delivering a far more anemic recovery, and thus a better match to the U.S. post-2009Q2 experience.

Keywords: entry & exit, selection, (S,s) policies, capital reallocation, propagation, business cycles

1 Introduction

It is well understood that the dynamics of capital investment have enormous implications for an economy's business cycle fluctuations. When endogenous capital accumulation is introduced into a typical equilibrium business cycle model, the consequences of temporary disturbances are amplified and propagated in quantitatively important ways. This observation suggests that the dynamics of other forms of investment might also be important in shaping aggregate fluctuations. When viewed from an aggregate perspective, individual firms' decisions regarding their entry and continuation have the capacity to generate such alternative investment dynamics.

How do endogenous movements in the number of firms and their age, size and productivity composition affect the size and persistence of macroeconomic fluctuations? To explore this question, we develop a dynamic stochastic general equilibrium model with endogenous entry and exit and firm-level capital accumulation. Our firms have persistent differences in idiosyncratic productivity, they face fixed costs to enter production and fixed operating costs to continue, and capital reallocation across them is hindered by microeconomic adjustment frictions. Thus, we can consider how age, size and selection reshape macroeconomic fluctuations in a general equilibrium environment disciplined by realistic firm life-cycle dynamics and investment patterns.

Examining standard business cycle moments and impulse responses, we find that changes in firms' entry and exit decisions amplify ordinary business cycles driven by shocks to aggregate productivity somewhat and subtly protract them when sufficiently long horizons are considered. Both results stem from an endogenous downward pull on TFP induced by a missing generation effect, whereby an usually small number of entrants fails to replace an increased number of exitors. In anticipation of this TFP drag, employment and investment fall more than otherwise. The missing generation effect grows prominent after about 5 years when the reduced cohorts of young firms approach maturity and would ordinarily account for a large share of aggregate production. That episode persists over several years, gradualizing the recovery in GDP.

The effects of an aggregate productivity shock are inherently uniform, in that they directly scale all firms' productivities. We also consider the macroeconomic response to a shock that has an asymmetric impact on the distribution of firms and emulates some aspects of the Great Recession. Declines in the number of firms, the number of young firms, and the overall employment share of small firms were dramatic over the U.S. 2007-9 recession. Our second shock induces such unusual changes through a rise in firms' operating costs. Because the payment of such costs is a discrete decision determined by firm value, this shock most directly affects entry and the exit decisions of younger firms. As such, it sharpens the missing generation effect described above, delivering a far more anemic recovery relative to that following a typical recession.

To be informative about the ways in which firms' entry and exit decisions shape aggregate fluctuations in actual economies, it is essential that our theoretical environment generate firm life-cycle dynamics resembling those in the data. Our model reproduces a key set of stylized facts about the characteristics of new firms, incumbent firms in production, and those exiting the economy. At the core of our setting, we have in essence Hopenhayn's (1992) model of industry dynamics. Potential firms receive informative signals about their future productivities and determine whether to pay fixed costs to become startups. Startups and incumbent firms have productivities affected by a persistent common component and a persistent idiosyncratic component, and they decide whether to pay fixed costs to operate or leave the economy. This set of assumptions immediately implies a selection effect whereby the average productivity, size and value of surviving members within a cohort rise as that cohort ages. Firms that have recently entered production are, on average, smaller, less productive and more likely to exit than are older firms, as consistent with the observations of Dunne, Roberts and Samuelson (1989) and other studies. Moreover, all else equal, large firms are those that have relatively high productivities, so mean-reversion in productivity delivers the unconditional negative relationships between size and growth and between age and growth.

One limitation of the original Hopenhayn framework is its perfect mapping between productivity, size and growth. After controlling for size, this leaves no independent negative relationship between age and growth, in contrast to evidence presented by Evans (1987) and Hall (1987). As in Clementi and Palazzo (2010), we overcome this problem by including capital in the production function and imposing frictions on capital reallocation, so that idiosyncratic productivity and capital become separately evolving state variables for a firm. Because firms cannot immediately adjust their capital stocks following changes in their productivities, those observed to be large in the usual employment-based sense need not be firms with high productivity; some may be large by virtue of their accumulated capital stocks.

Consider a group of firms of common size. Given one-period time-to-build in capital, those among them with the smallest stocks and highest idiosyncratic productivities will exhibit the fastest growth between this period and the next, as they raise their capital toward a level consistent with their high relative productivity. By contrast, those with large stocks and low productivity will shrink as they shed excess capital. To be in the latter position, a firm must have experienced a sufficiently long episode of high productivity to have accumulated a large stock. Such firms are more likely to be old than young, particularly given micro-level investment frictions that gradualize firms' capital adjustments. Thus, conditional on size, employment growth rates are negatively correlated with age.

Given its success in reproducing the essential aspects of firm life-cycle dynamics, the model of Clementi and Palazzo (2010) serves as our starting point.[1] There, changes in entry and exit over the cycle are seen to not only amplify the unconditional variation of aggregate series such as GDP and employment, but also generate greater persistence in the economy's responses to shocks. We extend that environment to general equilibrium by explicit introduction of a representative household supplying labor and savings to firms. One problem we confront in doing so is the fact that aggregate excess demand moves discontinuously in a search for an equilibrium interest rate path if small changes in prices induce sharp changes in the number of operating firms. We overcome this obstacle by introducing randomness in the fixed costs of both entry and operation. We also take a somewhat different approach to the supply of potential entrants, replacing the assumption of a constant per-period supply with one that evolves over time as a function of the number of incumbent firms in operation.

We calibrate the parameters of our model using long-run observations on aggregate and firm-level variables, including a series of moments on age, size and survival rates drawn from the BDS and a separate set of observations from Cooper and Haltiwanger (2006) regarding the average distribution of firm-level investment rates. We then verify that our model is a useful laboratory in which to explore that aggregate implications of selection and reallocation by confirming that its microeconomic predictions are consistent with the above-mentioned regularities. Next, we solve the model using a nonlinear method similar to that in Khan and Thomas (2008).

Nonlinearities are absent in representative agent models, which necessarily abstract from binary decisions. Our setting has three sets of such decisions characterized by (S,s) thresholds.

[1] Lee and Mukoyama (2009) also consider the implications of entry and exit in a model based on the Hopenhayn framework. Aside from the fact that ours is a general equilibrium study, a primary distinction between our work and theirs is our inclusion of capital. Samaniego (2008) studies perfect foresight transitions in a general equilibrium model of entry and exit. He includes capital, but as a single stock managed by the representative household and frictionlessly rented to firms within each period.

When the common exogenous component of TFP is unusually low, a potential firm that might otherwise pay its fixed entry cost sees its expected value reduced. At any given idiosyncratic productivity signal, the set of entry costs a potential firm is willing to accept shrinks. Thus, at the onset of a recession, the number of new startups falls, while their mean expected productivity rises. Next, operating decisions determine which new firms actually enter into production and which incumbent firms remain. Given the drop in all firms' values at the onset of a TFP-led recession, the willingness to pay operating costs to produce and continue in the economy falls at each capital and idiosyncratic productivity pair, implying reduced entry and raised exit. Fewer incumbents remain in production, and they are more selective than usual about continuing from relatively low individual productivity levels. Because similar mechanics deter entry, our model delivers both countercyclical exit and procyclical entry. As noted above, these forces amplify the responses in aggregate production, employment and investment following an aggregate productivity shock. Third, given micro-level capital adjustment frictions, we also have extensive margin decisions on investment. However, in keeping with results in Khan and Thomas (2003, 2008), we find these have negligible impact for macroeconomic fluctuations in our model.

As noted above, changes in firm startup, entry and exit decisions can imply greater persistence in aggregate fluctuations, due to a missing generation effect. Following a negative TFP shock, an unusually small number of young firms are in production. Over subsequent periods, as aggregate productivity begins to revert toward its mean, the typical surviving member of this smaller-than-average group of young firms grows in productivity and size, so the cohort's reduced membership hinders aggregate productivity and production. The extent to which such changes gradualize economic recovery depends on the extent of the disruptions in entry and exit over early dates, which in turn depend on the size of the TFP shock. Examining impulse responses following a conventional aggregate productivity shock, we find that endogenous changes in entry and exit extend the half life of the GDP response by roughly one year.[2] Examining HP-filtered business cycle moments for a 5000 period simulation of our model driven by productivity shocks, our setting

[2] These results support the findings of Clementi and Palazzo (2010) qualitatively, though the magnitudes are less pronounced. Three aspects of our model dampen the aggregate implications of endogenous entry and exit following a productivity shock. The first two, an endogenously evolving household labor supply curve and time-varying equilibrium interest rates, imply more modest changes in entry and exit in response to the shock. The third, an endogenously evolving supply of potential entrants, permits more rapid reversals following damage to the stock of firms. Our findings here are similar to Samaniego (2008), which shows transitional dynamics following productivity shocks are insensitive to changes in entry and exit rates, because such changes are small in equilibrium.

has somewhat higher volatility relative to an otherwise identical model with a fixed measure of firms; however, cross correlations at short horizons reveal little evidence of increased persistence.

There is, by now, a mounting body of firm-level evidence that the most recent U.S. recession had disproportionate negative effects on young firms (Sedlacek (2013), Sedlacek and Sterk (2014)) and on small firms (Khan and Thomas (2013), Siemer (2013)). Indirect evidence suggests this recession originated with a shock in the financial sector (Almeida et al. (2009), Duchin et al. (2010)). Khan and Thomas (2013) examines a shock to the availability of credit in an equilibrium model where a fixed measure of heterogenous firms face real and financial frictions. Predictions there match the 2007 recession well, but the model fails to deliver the subsequent anemic recovery. Several recent equilibrium studies have considered whether changes in the number and composition of firms may have contributed to this. Sedlacek (2013) examines a search and matching model with multi-worker firms and endogenous entry and exit following a TFP shock, while Siemer (2013) considers a credit crunch in a setting where new firms must finance a fraction of their startup costs with debt. Both models predict a missing (or lost) generation effect that propagates the effects of an aggregate shock; however, both abstract from capital and thus its reallocation. Khan, Senga and Thomas (2014) consider a shock to default recovery rates and firms' cash positions in a model with endogenous default, entry and exit and find that endogenous destruction to the stock of firms slows the recovery. However, they abstract from micro-level capital adjustment frictions essential to match key aspects of firm-level investment patterns. Furthermore, because their financial shock simultaneously disrupts investment decisions of continuing incumbent firms, it is difficult to disentangle the damage caused by persistent disruption to the distribution of capital across such firms from that arising as a result of the disruption to firms' entry and exit decisions.

Drawing on evidence from the BDS, three striking observations distinguish the Great Recession relative to a typical recession. First, the total number of firms fell by 5 percent (Siemer (2013)). Second, the number of young (age 5 and below) firms fell by 15 percent (Sedlacek (2013)). Third, total employment among small (fewer than 100 employees) firms fell more than twice as much as it did among large (more than 1000 employees) firms (Khan and Thomas (2013)). When our model is confronted with a shock raising firms' operating costs, we find that its asymmetric impact on the decisions of vulnerable firms generates sharp declines in entry and rises in exit rates concentrated particularly among young and small firms, reducing total employment in such firms disproportionately relative to the aggregate employment decline. The number of firms ultimately falls by roughly 5 percent, with the greatest population losses occurring in young age groups. As noted above, the disparate impact of this shock on young firms sharpens the missing generation effect in our model, and delivers an anemic recovery in GDP.

The remainder of the paper is organized as follows. Section 2 describes our model. Section 3 analyzes the three sets of threshold policy rules arising therein and derives a series of implications useful in developing a numerical algorithm to solve for competitive equilibrium. Section 4 discusses our model's calibration to moments drawn from postwar U.S. aggregate and firm-level data and thereafter describes the solution method we adopt. Section 5 presents results, first exploring aspects of our model's steady state, then considering aggregate fluctuations. Section 6 concludes.

2 Model

Our model economy builds on Clementi and Palazzo (2010), extending their setting to general equilibrium.[3] We have three groups of decision makers: households, firms and potential firms. Households are identical and own all firms. Potential firms face fixed entry costs to access the opportunity to produce in the next period. Firms face fixed operating costs as well as both convex and nonconvex costs of capital adjustment. These costs compound the effects of persistent differences in total factor productivities, yielding substantial heterogeneity in production. We begin this section with a summary of the problems facing firms and potential firms, then follow with a brief discussion of households and a description of equilibrium.

2.1 Firms

Our economy houses a large, time-varying number of firms. Conditional on survival, each firm produces a homogenous output using predetermined capital stock k and labor n, via an increasing and concave production function F. Each such firm's output is $y = z\varepsilon F(k, n)$, where z is exogenous stochastic total factor productivity common across firms, and ε is a persistent firm-specific counterpart. For convenience, we assume that ε is a Markov chain; $\varepsilon \in \{\varepsilon_1, \dots, \varepsilon_{N_\varepsilon}\}$,

[3]Beyond our explicit treatment of households, the main departure in extending that environment to general equilibrium is the introduction of idiosyncratic randomness to fixed costs associated with firm entry and continuation. Given discrete firm-specific productivity shocks, this modification serves to smooth the responses in aggregate excess demand to changes in prices, faciliating the search for equilibrium. Our differing approach in modeling the supply of potential entrants will be explained in section 2.2 below.

where $\Pr(\varepsilon' = \varepsilon_m \mid \varepsilon = \varepsilon_l) \equiv \pi^{\varepsilon}_{lm} \geq 0$, and $\sum_{m=1}^{N_\varepsilon} \pi^{\varepsilon}_{lm} = 1$ for each $l = 1, \ldots, N_\varepsilon$. Similarly, $z \in \{z_1, \ldots, z_{N_z}\}$ with $\Pr(z' = z_j \mid z = z_i) \equiv \pi_{ij} \geq 0$, and $\sum_{j=1}^{N_z} \pi_{ij} = 1$ for each $i = 1, \ldots, N_z$.

At the beginning of any period, each firm is defined by its predetermined stock of capital, $k \in \mathcal{K} \subset \mathbf{R}_+$, and by its current idiosyncratic productivity level, $\varepsilon \in \{\varepsilon_1, \ldots, \varepsilon_{N_\varepsilon}\}$. We summarize the start-of-period distribution of firms over (k, ε) using the probability measure μ defined on the Borel algebra generated by the open subsets of the product space $\mathcal{K} \times \mathcal{E}$; $\mu : \mathcal{B}(\mathcal{K} \times \mathcal{E}) \to [0, 1]$. The aggregate state of the economy will be fully described by (z, μ), with the distribution of firms evolving over time according to an equilibrium mapping, Γ, from the current state; $\mu' = \Gamma(z, \mu)$. The evolution of the firm distribution is determined in part by the actions of continuing firms and in part by the startups of potential firms to be described below.[4]

On entering a period, any given firm (k, ε) observes the economy's aggregate state (hence equilibrium prices) and also observes an output-denominated fixed cost it must pay to remain in operation, φ. This operating cost is individually drawn each period from a time-invariant distribution $H(\varphi)$ with bounded support $[\varphi_L, \varphi_U]$. The firm can either pay its φ to enter current production, or it can immediately and permanently exit the economy. If it chooses to exit, it sells its capital to recover a scrap value $(1 - \lambda)k$, where $\lambda \in [0, 1]$.

If a firm pays its operating cost, it then chooses its current level of employment, n, undertakes production, and pays its wage bill. Next, it observes its realization of a fixed cost associated with capital adjustment, $\xi \in [\xi_L, \xi_U]$, which is denominated in units of labor and individually drawn each period from the time-invariant distribution $G(\xi)$. At that point, the firm chooses its investment in capital for the next period, given the standard accumulation equation,

$$k' = (1 - \delta)\, k + i, \tag{1}$$

where $\delta \in (0, 1)$ is the rate of capital depreciation, and primes indicate one-period-ahead values.

The firm can avoid capital adjustment costs by undertaking zero investment. However, if it chooses $i \neq 0$, it must hire ξ units of labor at equilibrium wage ω to manage the activity, and it must also suffer a convex output-disruption cost $c_q(\frac{i}{k})^2 k$, where $c_q > 0$.

investment	adjustment costs	future capital
$i \neq 0$	$\omega(z, \mu)\xi + c_q \frac{i^2}{k}$	any $k' \in \mathcal{K}$
$i = 0$	0	$k' = (1 - \delta)k$

[4] Our distribution μ includes new business startups (described in the section below). We define entrants in our model as those startups that choose to produce; those that do not are neither entrants nor exitors.

The optimization problem facing each of the economy's firms may be described as follows. Given the current aggregate state, (z_i, μ), let $v^1(k, \varepsilon_l, \varphi; z_i, \mu)$ denote the expected discounted value of a firm that enters the period with capital k and idiosyncratic productivity ε_l just after it observes its current operating cost φ. Let $v^0(k, \varepsilon_l; z_i, \mu)$ be its expected value just beforehand;

$$v^0(k, \varepsilon_l; z_i, \mu) \equiv \int_{\varphi_L}^{\varphi_U} v^1(k, \varepsilon_l, \varphi; z_i, \mu) H(d\varphi). \tag{2}$$

The first decision the firm faces is whether to operate or exit. Defining the flow profit function,

$$\pi(k, \varepsilon; z, \mu) \equiv \max_n [z\varepsilon F(k, n) - \omega(z, \mu)n], \tag{3}$$

the firm solves the following binary maximization problem at the start of the period.

$$v^1(k, \varepsilon_l, \varphi; z_i, \mu) = \max\Big\{(1-\lambda)k,\ \pi(k, \varepsilon_l; z_i, \mu) - \varphi + \int_{\xi_L}^{\xi_U} v^2(k, \varepsilon_l, \xi; z_i, \mu) G(d\xi)\Big\} \tag{4}$$

Since the firm cannot observe its fixed capital adjustment cost until it produces, the ex-production continuation value in (4) computed at the start of the period involves an expectation over the possible realizations of ξ. In some places below, we find it convenient to represent the continuation decision of an incumbent firm using an indicator function χ.

$$\chi(k, \varepsilon, \varphi; z, \mu) = \begin{cases} 1 \text{ if } \pi(k, \varepsilon; z, \mu) - \varphi + \int_{\xi_L}^{\xi_U} v^2(k, \varepsilon, \xi; z, \mu) G(d\xi) \geq (1-\lambda)k \\ 0 \text{ otherwise} \end{cases}$$

The value function v^2 represents an operating firm's discounted continuation value net of investment and capital adjustment costs. The firm faces a second binary decision at the end of the current period as it chooses its investment. Let $d_j(z_i, \mu)$ represent the discount factor each firm applies to its next-period value conditional on $z' = z_j$ and the current aggregate state (z_i, μ). Taking as given the evolution of ε and z according to the transition probabilities defined above, and taking as given the evolution of the firm distribution, $\mu' = \Gamma(z, \mu)$, the firm solves the optimization problem in (5) - (6) to determine its future capital.

$$\begin{aligned} v^2(k, \varepsilon_l, \xi; z_i, \mu) &= \max\Big\{\sum_{j=1}^{N_z}\sum_{m=1}^{N_\varepsilon} \pi_{ij}\pi_{lm}^{\varepsilon} d_j(z_i, \mu) v^0((1-\delta)k, \varepsilon_m; z_j, \mu'), \\ &\qquad -\omega(z_l, \mu)\xi + e(k, \varepsilon_l; z_i, \mu)\Big\}, \text{ where} \end{aligned} \tag{5}$$

$$\begin{aligned} e(k, \varepsilon_l; z_i, \mu) &= \max_{k' \in \mathcal{K}} \Big[-[k' - (1-\delta)k] - \frac{c_q}{k}[k' - (1-\delta)k]^2 \\ &\quad + \sum_{j=1}^{N_z}\sum_{m=1}^{N_\varepsilon} \pi_{ij}\pi_{lm}^{\varepsilon} d_j(z_i, \mu) v^0(k', \varepsilon_m; z_j, \mu')\Big] \end{aligned} \tag{6}$$

The firm can select line 1 of (5), avoiding all capital adjustment costs, and continue to the next period with the remains of its current capital after depreciation. Alternatively, by selecting line 2, it can pay its random fixed cost ξ (converted to output units by the wage) and select a k' that maximizes its continuation value net of investment and convex adjustment costs.

In section 3, we will revisit the incumbent firm problem from (2) - (6) and characterize the resulting decision rules. For now, note that there is no friction associated with a firm's employment choice, since the firm pays its current wage bill after production takes place, and its capital choice for next period also has no implications for current production. Thus, conditional on paying the fixed costs to operate, firms sharing in common the same (k, ε) combination select a common employment and output, which we denote by $n(k, \varepsilon; z, \mu)$ and $y(k, \varepsilon; z, \mu)$, respectively. By contrast, they make differing investment decisions, given differences in their fixed capital adjustment costs. We denote their choices of next-period capital by $g(k, \varepsilon, \xi; z, \mu)$.[5]

2.2 Potential firms

There is a fixed stock of production locations in the economy, Q. Any production location not in use by operating firms (one location per firm) may be used to create a potential firm. Thus, in any date t, there are M_t potential firms, where:

$$M_t \equiv M(z, \mu) = Q - \int_{\mathcal{K}\times\mathcal{E}} \int_{\varphi_L}^{\varphi_U} \chi(k, \varepsilon, \varphi; z, \mu) H(d\varphi) \mu(d[k \times \varepsilon]). \tag{7}$$

Each potential firm draws a productivity signal and chooses whether to pay a fixed entry cost to become a startup firm. Any such startup chooses a capital stock with which it will appear in the firm distribution at the start of next period.

A potential firm observes the current aggregate state, its output-denominated fixed entry cost, γ, and its productivity signal, s_l. Entry costs are individually drawn from the time-invariant distribution $H_e(\gamma)$ with bounded support $[\gamma_L, \gamma_U]$. Signals are individually drawn from a distribution with the same support as incumbent firm productivities, $\{s_1, \ldots, s_{N_\varepsilon}\} = \{\varepsilon_1, \ldots, \varepsilon_{N_\varepsilon}\}$, and with probability weights $\pi^e(s_l) \equiv \Pr(s = s_l)$. The transition probabilities from signals to future productivities match those for incumbent firms: $\Pr(\varepsilon' = \varepsilon_m \mid s = s_l) = \pi^{\varepsilon}_{lm}$, and startups choose their capital stocks accordingly.

[5] Absent the convex cost of capital adjustment, the same k' would solve (6) for all firms sharing the same current productivity, ε. In that case, an operating firm of type (k, ε) would adopt either $k'(\varepsilon; z, \mu)$ or $(1 - \delta)k$.

Equations 8 - 9 describe the optimization problem for a potential firm identified by $(s_l, \gamma; z_i, \mu)$. The first line reflects a binary choice of whether to become a startup. In the second line, a startup firm selects capital for the next period, when it will have its first opportunity to produce.

$$
\begin{aligned}
v^p(s_l, \gamma; z_i, \mu) &= \max\Big\{0, -\gamma + v^e(s_l; z_i, \mu)\Big\} && (8)\\
v^e(s_l; z_i, \mu) &= \max_{k' \in \mathcal{K}} \Big[-k' + \sum_{j=1}^{N_z}\sum_{m=1}^{N_\varepsilon} \pi_{ij}\pi^{\varepsilon}_{lm} d_j\left(z_i, \mu\right) v^0(k', \varepsilon_m; z_j, \mu')\Big] && (9)
\end{aligned}
$$

We let $g^e(s_l; z_i, \mu)$ denote the capital solving (9).[6] At points below, we reflect the entry decision of a potential firm using the indicator function χ^e.

$$
\chi^e(s_l, \gamma; z_i, \mu) = \begin{cases} 1 \text{ if } -\gamma + v^e(s_l; z_i, \mu) \geq 0 \\ 0 \text{ otherwise} \end{cases}
$$

2.3 Households

The economy is populated by a unit measure of infinitely-lived, identical households. Household wealth is held as one-period shares in firms, which we denote using the measure λ.[7] Given the prices they receive for their current shares, $\rho_0\left(k, \varepsilon; z_i, \mu\right)$, and the real wage they receive for their labor, $\omega\left(z_i, \mu\right)$, households determine their current consumption, c, hours worked, n^h, and the numbers of new shares, $\lambda'\left(k', \varepsilon'\right)$, to purchase at prices $\rho_1\left(k', \varepsilon'; z_i, \mu\right)$. The lifetime expected utility maximization problem of the representative household is listed below.

$$
W\left(\lambda; z, \mu\right) = \max_{c, n^h, \lambda'} \Big[U\left(c, 1 - n^h\right) + \beta \sum_{j=1}^{N_z} \pi_{ij} W\left(\lambda'; z_j, \mu'\right)\Big] \tag{10}
$$

subject to

$$
c + \int_{\mathcal{K}\times\mathcal{E}} \rho_1\left(k', \varepsilon'; z, \mu\right) \lambda'\left(d\left[k' \times \varepsilon'\right]\right) \leq \omega\left(z, \mu\right) n^h + \int_{\mathcal{K}\times\mathcal{E}} \rho_0\left(k, \varepsilon; z, \mu\right) \lambda\left(d\left[k \times \varepsilon\right]\right).
$$

Let $C\left(\lambda; z, \mu\right)$ describe the household consumption choice, and let $N\left(\lambda; z, \mu\right)$ be its choice of hours worked. Finally, let $\Lambda\left(k', \varepsilon', \lambda; z, \mu\right)$ be the quantity of shares purchased in firms that will begin the next period with k' units of capital and idiosyncratic productivity ε'.

[6] If incumbent firms faced no convex costs of capital adjustment ($c_q = 0$), any entrant with signal s_l would select the same k' as every incumbent firm with productivity ε_l currently undertaking nonzero investment. That convenient result does not hold for the current model, since $c_q > 0$ implies incumbents' intensive margin investment decisions are affected by their current capital levels.

[7] Households also have access to a complete set of state-contingent claims. As there is no heterogeneity across households, these assets are in zero net supply in equilibrium, so we do not explicitly model them here.

2.4 Recursive equilibrium

A *recursive competitive equilibrium* is a set of functions,

$$\Big(\omega, (d_j)_{j=1}^{N_z}, \rho_0, \rho_1, v^1, n, g, \chi, v^p, g^e, \chi^e, W, C, N, \Lambda\Big),$$

that solve firm and household problems and clear the markets for assets, labor and output, as described by the following conditions.

(i) v^1 solves (4) - (6), and (χ, n, g) are the associated policy functions for firms

(ii) v^p solves (8) - (9), and χ^e and g^e are the resulting policy functions for potential firms

(iii) W solves (10), and (C, N, Λ) are the associated policy functions for households

(iv) $\Lambda(k', \varepsilon', \mu; z, \mu) = \mu'(k', \varepsilon'; z, \mu)$, for each $(k', \varepsilon') \in \mathcal{K} \times \mathcal{E}$

(v) $$N(\mu; z, \mu) = \int_{\mathcal{K}\times\mathcal{E}} \int_{\varphi_L}^{\varphi_U} \chi(k, \varepsilon, \varphi; z, \mu) \Bigg[n(k, \varepsilon; z, \mu) + \textstyle\int_{\xi_L}^{\xi_U} \xi \mathcal{J}\Big(g(k, \varepsilon, \xi; z, \mu) - (1-\delta) k \Big) G(d\xi) \Bigg] H(d\varphi)\mu(d[k \times \varepsilon]),$$

where $\mathcal{J}(x) = 0$ if $x = 0$; $\mathcal{J}(x) = 1$ otherwise.

(vi) $$C(\mu; z, \mu) = \int_{\mathcal{K}\times\mathcal{E}} \int_{\varphi_L}^{\varphi_U} \chi(k, \varepsilon, \varphi; z, \mu) \Bigg[z\varepsilon F(k, n(k, \varepsilon; z, \mu)) - \varphi - \textstyle\int_{\xi_L}^{\xi_U} \Big[g(k, \varepsilon, \xi; z, \mu) - (1-\delta) k + \frac{c_q}{k}\Big(g(k, \varepsilon, \xi; z, \mu) - (1-\delta)k\Big)^2 \Big] \mathcal{J}\Big(g(k, \varepsilon, \xi; z, \mu) - (1-\delta) k \Big) G(d\xi) \Bigg] H(d\varphi)\mu(d[k \times \varepsilon])$$
$$+ \int_{\mathcal{K}\times\mathcal{E}} \int_{\varphi_L}^{\varphi_U} \Big[1 - \chi(k, \varepsilon, \varphi; z, \mu) \Big] \Big[(1-\lambda)k \Big] H(d\varphi)\mu(d[k \times \varepsilon])$$
$$- M(z, \mu) \sum_{l=1}^{N_\varepsilon} \pi^e(s_l) \int_{\gamma_L}^{\gamma_U} \chi^e(s_l, \gamma; z, \mu) \Big[\gamma + g^e(s_l; z, \mu) \Big] H_e(d\gamma),$$

(vii) $$\mu'(D, \varepsilon_m) = \int_{\{(k, \varepsilon_l, \xi) \,|\, g(k, \varepsilon_l, \xi; z, \mu) \in D\}} \chi(k, \varepsilon_l, \varphi; z, \mu) \pi_{lm}^{\varepsilon} G(d\xi)\, H(d\varphi)\mu(d[\varepsilon_l \times k])$$
$$+ M(z, \mu) \sum_{\{s_l \,|\, g^e(s_l; z, \mu) \in D\}} \pi^e(s_l) \pi_{lm}^{\varepsilon} \int_{\gamma_L}^{\gamma_U} \chi^e(s_l, \gamma; z, \mu) H_e(d\gamma),$$

for all (D, ε_m) measurable, defines Γ

3 Analysis

Let C and N represent the market-clearing values of household consumption and hours worked satisfying conditions (v) and (vi) above. It is straightforward to show that market-clearing requires that (a) the real wage equal the household marginal rate of substitution between leisure and consumption, $\omega(z,\mu) = D_2U(C, 1-N)/D_1U(C, 1-N)$, and that (b) firms' (and potential firms') state-contingent discount factors agree with the household marginal rate of substitution between consumption across states. Letting C'_{ij} denote household consumption next period given current state (z_i, μ) and future state $(z_j, \mu'(z_i, \mu))$ and with N'_{ij} as the corresponding labor input, the resulting discount factors are: $d_j(z_i, \mu) = \beta D_1U\left(C'_{ij}, 1-N'_{ij}\right)/D_1U(C, 1-N)$.

We may compute equilibrium by solving a single Bellman equation that combines the firm profit maximization problem with the equilibrium implications of household utility maximization from above. Here, we effectively subsume households' decisions into the problems faced by firms. Without loss of generality, we assign $p(z, \mu)$ as an output price at which firms and potential firms value current profits and payments, and we correspondingly impose that their future values are discounted by the household subjective discount factor. Given this alternative means of expressing equilibrium discount factors, the following two conditions ensure all markets clear in our economy.

$$p(z,\mu) = D_1U(C, 1-N) \tag{11}$$

$$\omega(z,\mu) = D_2U(C, 1-N)/p(z,\mu) \tag{12}$$

To develop a tractable numerical algorithm with which to solve our economy, it is useful to characterize the optimizing decisions of incumbent and potential firms in ways convenient for aggregation. As we consider firms' and potential firms' binary choice problems, we find it convenient to start with the continuous decision problems contingent on each action, then work backward to the binary choice. Throughout this section, we suppress aggregate state arguments in the p and ω functions to shorten the equations, and we continue abbreviating $\mu'(z,\mu)$ by μ'.

We begin by reformulating (2) - (6) to describe each firm's value in units of marginal utility, with no change in the resulting decision rules. Exploiting the fact that the choice of n is independent of the k' choice, suppressing the indices for current aggregate and idiosyncratic productivity, and defining $V^0(k,\varepsilon;z,\mu) \equiv \int_{\varphi_L}^{\varphi_U} V^1(k,\varepsilon,\varphi;z,\mu)\,H(d\varphi)$, we have the following recursive representation for the start-of-period value of a type (k,ε) firm drawing operating cost φ.

$$V^1(k,\varepsilon,\varphi;z,\mu) = \max\Big\{p(1-\lambda)k,\ p[\pi(k,\varepsilon;z,\mu)-\varphi] + \int_{\xi_L}^{\xi_U} V^2(k,\varepsilon,\xi;z,\mu)G(d\xi)\Big\} \tag{13}$$

$$V^2(k,\varepsilon,\xi;z,\mu)=\max\Big\{\beta\sum\sum\pi_{ij}\pi^{\varepsilon}_{lm}V^0((1-\delta)k,\varepsilon_m;z_j,\mu'),-p\omega\xi+E(k,\varepsilon;z,\mu)\Big\} \tag{14}$$

$$\begin{aligned}E(k,\varepsilon;z,\mu) &= \max_{k'\in\mathcal{K}}\Big[-p[k'-(1-\delta)k]-\frac{pc_q}{k}[k'-(1-\delta)k]^2 \\ &\qquad +\beta\sum\sum\pi_{ij}\pi^{\varepsilon}_{lm}V^0(k',\varepsilon_m;z_j,\mu')\Big]\end{aligned} \tag{15}$$

The problem of a potential firm from (8) - (9) is analogously reformulated.

$$V^p(s_l,\gamma;z_i,\mu) = \max\Big\{0,-p\gamma+V^e(s_l;z_i,\mu)\Big\} \tag{16}$$

$$V^e(s_l;z_i,\mu) = \max_{k'\in\mathcal{K}}\Big[-pk'+\beta\sum_{j=1}^{N_z}\sum_{m=1}^{N_\varepsilon}\pi_{ij}\pi^{\varepsilon}_{lm}V^0(k',\varepsilon_m;z_j,\mu')\Big] \tag{17}$$

3.1 Continuing firms' investment decisions

Consider first the end-of-period decision made by a continuing firm that has chosen to pay its adjustment cost and undertake a nonzero investment. Any such firm will adopt a *target capital* consistent with its current productivity and the aggregate state, which we denote by $k^*(k,\varepsilon;z,\mu)$.

$$k^*(k,\varepsilon;z,\mu)\equiv\arg\max_{k'\in\mathcal{K}}\Big[-pk'-\frac{pc_q}{k}[k'-(1-\delta)k]^2+\beta\sum_{j=1}^{N_z}\sum_{m=1}^{N_\varepsilon}\pi_{ij}\pi^{\varepsilon}_{lm}V^0(k',\varepsilon_m;z_j,\mu')\Big] \tag{18}$$

The gross adjustment value associated with this action is $E(k,\varepsilon;z,\mu)$ from equation 15.

If there were no convex adjustment costs, notice that the target capital choice would be independent of a firm's current capital, since the price of investment goods (p) is unaffected by its level of investment and the current capital adjustment cost draw ξ carries no information about future draws (and thus does not enter V^0). In that case, all firms with the same current productivity level undertaking nonzero investment would move to the next period with a common capital stock, and their gross adjustment values would be linear in k; both observations could be used to expedite model solution. However, given $c_q>0$, the scale of adjustment affects the level of adjustment costs; hence, target capitals depend on not only ε but also k.

Next, we turn to the binary capital adjustment decision. For a continuing firm of type (k,ε), the ex-production value of undertaking no adjustment is $\beta\sum\sum\pi_{ij}\pi^{\varepsilon}_{lm}V^0((1-\delta)k,\varepsilon_m;z_j,\mu')$, while the value of adjustment is $-p\omega\xi+E(k,\varepsilon;z,\mu)$. The firm pays its capital adjustment cost only if the net benefit of doing so is non-negative, i.e., if:

$$[-p\omega\xi+E(k,\varepsilon;z,\mu)]-\beta\sum\sum\pi_{ij}\pi^{\varepsilon}_{lm}V^0((1-\delta)k,\varepsilon_m;z_j,\mu')\geq 0.$$

The firm's capital decision rule can be described as a threshold policy. Define $\widetilde{\xi}(k,\varepsilon;z,\mu)$ as the fixed cost that leaves the firm indifferent to adjustment, and define $\xi^T(k,\varepsilon;z,\mu)$ as the resulting threshold cost confined to the support of the cost distribution.

$$\begin{aligned}
\widetilde{\xi}(k,\varepsilon;z,\mu) &= \frac{E(k,\varepsilon;z,\mu) - \beta\sum\sum \pi_{ij}\pi^{\varepsilon}_{lm}V^0((1-\delta)k,\varepsilon_m;z_j,\mu')}{p\omega} \\
\xi^T(k,\varepsilon;z,\mu) &= \max\{\xi_L, \min\{\xi_U, \widetilde{\xi}(k,\varepsilon;z,\mu)\}\} \qquad (19)
\end{aligned}$$

If the firm draws a fixed cost at or below its threshold, ξ^T, it pays that cost and adopts the target $k^*(k,\varepsilon;z,\mu)$. Otherwise, it undertakes zero investment. The resulting capital decision rule is listed below.

$$g\left(k,\varepsilon,\xi;z,\mu\right) = \begin{cases} k^*\left(k,\varepsilon,z,\mu\right) & \text{if} \quad \xi \leq \xi^T(k,\varepsilon;z,\mu) \\ (1-\delta)k & \text{otherwise} \end{cases}$$

All else equal, a firm tends to be more willing to pay adjustment costs when its existing stock is farther away from its target. When this is so, the threshold cost is higher, which in turn implies a greater likelihood that the firm will adopt its k^*. Thus, our model implies (S,s) capital decisions and rising adjustment hazards as in Caballero and Engel (1999), Khan and Thomas (2003, 2008) and other studies involving nonconvex microeconomic investment decisions.

Observe from (19) that all firms of type (k,ε) share in common the same threshold cost ξ^T. Thus, each of them has the same probability of capital adjustment and hence the same expected ex-production continuation value before the individual ξ draws have been realized. Let $\alpha^k(k,\varepsilon;z,\mu)$ denote any such firm's probability of capital adjustment, which is simply the probability of drawing $\xi \leq \xi^T$, and let $\Phi^k(k,\varepsilon;z,\mu)$ denote the conditional expectation of the fixed cost to be paid.

$$\begin{aligned}
\alpha^k(k,\varepsilon;z,\mu) &\equiv G\Big(\xi^T(k,\varepsilon;z,\mu)\Big) \qquad (20) \\
\Phi^k(k,\varepsilon;z,\mu) &\equiv \int_{\xi_L}^{\xi^T(k,\varepsilon;z,\mu)} \xi G(d\xi) \qquad (21)
\end{aligned}$$

3.2 Operating decisions

As firms make their operating decisions at the start of a period, recall that they do not yet know their current fixed capital adjustment costs. As such, they use (18) - (21) from above to compute their expected ex-production continuation values.

$$\begin{aligned}
\int_{\xi_L}^{\xi_U} V^2(k,\varepsilon,\xi;z,\mu)G(d\xi) &= [1-\alpha^k(k,\varepsilon;z,\mu)]\beta\sum_{j=1}^{N_z}\sum_{m=1}^{N_\varepsilon}\pi_{ij}\pi^{\varepsilon}_{lm}V^0((1-\delta)k,\varepsilon_m;z_j,\mu') \qquad (22) \\
&\quad + \alpha^k(k,\varepsilon;z,\mu)E(k,\varepsilon;z,\mu) - p\omega\Phi^k(k,\varepsilon;z,\mu)
\end{aligned}$$

Given the expected continuation value from equation 22, we can solve any firm's start-of-period operating decision. If the firm exits the economy, it achieves a scrap value $p(1-\lambda)k$. If it operates, it achieves the flow profits $\pi(k,\varepsilon;z,\mu)$ from (3) and the expected continuation value from (22). The firm continues into production only if the value of its current operating cost does not exceed the net benefit of doing so;

$$\Big[p\pi(k,\varepsilon;z,\mu)+\int_{\xi_L}^{\xi_U}V^2(k,\varepsilon,\xi;z,\mu)G(d\xi)\Big]-p(1-\lambda)k\geq p\varphi.$$

The firm's binary operating decision can be described as a threshold policy. Define $\widetilde{\varphi}(k,\varepsilon;z,\mu)$ as the cost that leaves the firm indifferent to continuing, and define $\varphi^T(k,\varepsilon;z,\mu)$ as the resulting threshold cost confined to the support of H.

$$\begin{aligned}\widetilde{\varphi}(k,\varepsilon;z,\mu) &= \pi(k,\varepsilon;z,\mu)-(1-\lambda)k+\frac{1}{p}\int_{\xi_L}^{\xi_U}V^2(k,\varepsilon,\xi;z,\mu)G(d\xi)\\ \varphi^T(k,\varepsilon;z,\mu) &= \max\{\varphi_L,\min\{\widetilde{\varphi}(k,\varepsilon;z,\mu),\varphi_U\}\}\end{aligned}\tag{23}$$

If the firm realizes a φ above the threshold, φ^T, it exits the economy. Otherwise, it hires and produces according to the decision rules $n(k,\varepsilon;z,\mu)$ and $y(k,\varepsilon;z,\mu)$ that maximize its current flow profits (see equation 3).

Before leaving this subsection, note that (23) implies that all firms entering the period with the same (k,ε) pair have the same threshold operating cost. This means that, as they are entering the period, each of them has equal probability of survival, α^c, and equal conditional expectation of the operating costs they will pay, Φ^c;

$$\begin{aligned}\alpha^c(k,\varepsilon;z,\mu) &= H\Big(\varphi^T(k,\varepsilon;z,\mu)\Big)\\ \Phi^c(k,\varepsilon;z,\mu) &= \int_{\varphi_L}^{\varphi^T(k,\varepsilon;z,\mu)}\varphi H(d\varphi).\end{aligned}$$

Combining the results above (and recalling that $V^0(k,\varepsilon;z,\mu)\equiv\int_{\varphi_L}^{\varphi_U}V^1(k,\varepsilon,\varphi;z,\mu)H(d\varphi)$), we can compute the start-of-period expected value of any firm as it enters a period:

$$\begin{aligned}V^0(k,\varepsilon;z,\mu) &= [1-\alpha^c(k,\varepsilon;z,\mu)]p(1-\lambda)k-p\Phi^c(k,\varepsilon;z,\mu)\\ &\quad+\alpha^c(k,\varepsilon;z,\mu)[p(z,\mu)\pi(k,\varepsilon;z,\mu)-p\omega\Phi^k(k,\varepsilon;z,\mu)]\\ &\quad+\alpha^c(k,\varepsilon;z,\mu)\alpha^k(k,\varepsilon;z,\mu)E(k,\varepsilon;z,\mu)\\ &\quad+\alpha^c(k,\varepsilon;z,\mu)[1-\alpha^k(k,\varepsilon;z,\mu)]\beta\sum_{j,m}\pi_{ij}\pi^{\varepsilon}_{lm}V^0((1-\delta)k,\varepsilon_m;z_j,\mu'),\end{aligned}$$

where $E(k,\varepsilon;z,\mu)$ is defined in (15).

3.3 Entry decisions

Conditional on paying its entry cost to become a startup, a potential firm with productivity signal s_l adopts the capital stock solving (17) above. We denote that choice by $k_e^*(s_l; z, \mu)$ here forward. The potential firm pays its entry cost, γ, if:

$$\beta \sum_{j=1}^{N_z} \sum_{m=1}^{N_\varepsilon} \pi_{ij} \pi_{lm}^{\varepsilon} V^0(k_e^*(s_l; z, \mu), \varepsilon_m; z_j, \mu') - pk_e^*(s_l; z, \mu) \geq p\gamma.$$

Define $\widetilde{\gamma}(\varepsilon_l; z, \mu)$ as the entry cost implying indifference, and define $\gamma^T(\varepsilon_l; z, \mu)$ as the associated threshold entry cost confined to the support of H_e.

$$\begin{aligned}
\widetilde{\gamma}(\varepsilon_l; z, \mu) &= \frac{\beta}{p} \sum \sum \pi_{ij} \pi_{lm}^{\varepsilon} V^0(k_e^*(s_l; z, \mu), \varepsilon_m; z_j, \mu') - k_e^*(s_l; z, \mu) \\
\gamma^T(\varepsilon_l; z, \mu) &= \max\{\gamma_L, \min\{\widetilde{\gamma}(\varepsilon_l; z, \mu), \gamma_U\}\}
\end{aligned}$$

Only if the potential firm draws an entry cost at or below γ^T will it become a startup. Thus, we have the fraction of potential firms with signal s_l that will choose to invest toward next period entry, as well as the expected cost paid by each.

$$\begin{aligned}
\alpha^e(s_l; z, \mu) &= H\Big(\gamma^T(\varepsilon_l; z, \mu)\Big) \\
\Phi^e(s_l; z, \mu) &= \int_{\gamma_L}^{\gamma^T(\varepsilon_l; z, \mu)} \gamma H_e(d\gamma)
\end{aligned}$$

3.4 Aggregation

Given the probabilities of entry, continuation, and capital adjustment from above, alongside the conditional fixed cost expectations, and the accompanying labor, output and capital decision rules, aggregation is straightforward. Aggregate production and employment are

$$\begin{aligned}
Y(z, \mu) &= \int\limits_{\mathcal{K} \times \mathcal{E}} \Big[\alpha^c(k, \varepsilon; z, \mu) y(k, \varepsilon; z, \mu)\Big] \mu(d[k \times \varepsilon]) \\
N(z, \mu) &= \int\limits_{\mathcal{K} \times \mathcal{E}} \Big[\alpha^c(k, \varepsilon; z, \mu) n(k, \varepsilon; z, \mu)\Big] \mu(d[k \times \varepsilon]) + \Psi_n^k(z, \mu),
\end{aligned}$$

where Ψ_n^k is total labor-denominated fixed costs associated with capital adjustment;

$$\Psi_n^k(z, \mu) = \int_{\mathcal{K} \times \mathcal{E}} \Big[\alpha^c(k, \varepsilon; z, \mu) \Phi^k(k, \varepsilon; z, \mu)\Big] \mu(d[k \times \varepsilon]).$$

Aggregate investments across incumbent firms (I^c) and startup firms (I^e) are

$$\begin{aligned}
I^c(z,\mu) &= \int_{\mathcal{K}\times\mathcal{E}} \alpha^c(k,\varepsilon;z,\mu)\alpha^k(k,\varepsilon;z,\mu)\Big[k^*\left(\varepsilon;z,\mu\right)-(1-\delta)k\Big]\mu(d[k\times\varepsilon]) \\
&\quad -\int_{\mathcal{K}\times\mathcal{E}} \Big[\Big(1-\alpha^c(k,\varepsilon;z,\mu)\Big)(1-\lambda)k\Big]\mu(d[k\times\varepsilon]) \\
I^e(z,\mu) &= M(z,\mu)\sum_{l=1}^{N_\varepsilon}\pi^e(s_l)\alpha^e\left(s_l;z,\mu\right)k^*\left(s_l;z,\mu\right),
\end{aligned}$$

with the measure of potential firms given by $M(z,\mu)=Q-\int\limits_{\mathcal{K}\times\mathcal{E}}\alpha^c(k,\varepsilon;z,\mu)\mu(d\left[k\times\varepsilon\right])$. Household consumption is

$$C(z,\mu)=Y(z,\mu)-[I^c(z,\mu)+I^e(z,\mu)]-[\Psi^e(z,\mu)+\Psi^c(z,\mu)+\Psi^k_y(z,\mu)],$$

where Ψ^e, Ψ^c, and Ψ^k_y are the total output-denominated costs associated with startup entry, firm operations, and capital adjustment (Ψ^k_y), respectively.

$$\begin{aligned}
\Psi^e(z,\mu) &= M(z,\mu)\sum_{l=1}^{N_\varepsilon}\pi^e(\varepsilon_l)\Phi^e\left(s_l;z,\mu\right) \\
\Psi^c(z,\mu) &= \int\limits_{\mathcal{K}\times\mathcal{E}}\Phi^c(k,\varepsilon;z,\mu)\mu(d[k\times\varepsilon]) \\
\Psi^k_y(z,\mu) &= \int_{\mathcal{K}\times\mathcal{E}}\Big[\alpha^c(k,\varepsilon;z,\mu)\alpha^k(k,\varepsilon;z,\mu)\Big]\cdot \\
&\qquad \left[\frac{c_q}{k}\Big(k^*(k,\varepsilon;z,\mu)-(1-\delta)k\Big)^2\right]\mu(d[k\times\varepsilon]).
\end{aligned}$$

Finally, before turning to the calibration, we identify *incumbents*, *entrants*, and *exitors* in our model for comparison with firm-level data. Here forward, an incumbent is a firm that produced in the previous period, an entrant is a firm that has not produced before and does so in the current period, and an exitor is an incumbent that does produce in the current period. Given current aggregate state (z,μ) and next period state (z',μ'), the number of producers next period will be $\int\limits_{\mathcal{K}\times\mathcal{E}}\alpha^c(k,\varepsilon;z,\mu)\mu(d[k\times\varepsilon])$, the number of entrants will be:

$$M(z,\mu)\sum_{l=1}^{N_\varepsilon}\pi^e(s_l)\alpha^e\left(s_l;z,\mu\right)\sum_{m=1}^{N_\varepsilon}\pi^\varepsilon_{lm}\alpha^c\Big(k^*\left(s_l;z,\mu\right),\varepsilon_m;z',\mu'\Big),$$

and incumbent producers will be the difference between these two. To compute total exit next period, we count all potential producers leaving the economy, then discard those that have never produced (startups that do not enter):

$$\int\limits_{\mathcal{K}\times\mathcal{E}}[1-\alpha^c(k,\varepsilon;z,\mu)]\mu(d[k\times\varepsilon])-M(z,\mu)\sum_{l=1}^{N_\varepsilon}\pi^e(s_l)\alpha^e(s_l;z,\mu)\sum_{m=1}^{N_\varepsilon}\pi^\varepsilon_{lm}\Big[1-\alpha^c\Big(k^*(s_l;z,\mu),\varepsilon_m;z',\mu'\Big)\Big].$$

We measure entry and exit rates at each date in our model as they are measured in the data. The entry rate is the current number of entrants divided by the average number of firms across the current and previous date; the exit rate is current exitors over the same denominator.

4 Calibration and solution

In the sections to follow, we will at points consider how the mechanics of our model compare to those in a reference model with an exogenously fixed measure of firms. Aside from the changes noted here for that reference, we will select a common parameter set by targeting our full model economy at a series of moments drawn from postwar U.S. aggregate and firm-level data discussed below. To construct our no-entry/exit reference, we then reset the upper bound on continuation costs to 0 and adjust the fixed stock of production locations Q to imply a constant number of producers matching that obtained in the steady state of our full model.

4.1 Functional forms and aggregate targets

We assume that the representative household's period utility is the result of indivisible labor (Rogerson (1988)): $u(c,L) = \log c + \theta L$. Firm-level production is Cobb-Douglas: $z\varepsilon F(k,n) = z\varepsilon k^\alpha n^\nu$. In specifying our exogenous stochastic process for aggregate productivity, we begin by assuming a continuous shock following a mean zero AR(1) process in logs: $\log z' = \rho_z \log z + \eta'_z$ with $\eta'_z \sim N\left(0,\sigma^2_{\eta_z}\right)$. Next, we estimate the values of ρ_z and σ_{η_z} from Solow residuals measured using NIPA data on US real GDP and private capital, together with the total employment hours series constructed by Cociuba, Prescott and Ueberfeldt (2012) from CPS household survey data over 1959-2002. Next, we discretize the productivity process using a grid with 3 shock realizations to obtain (z_i) and (π_{ij}). We determine the firm-specific productivity shocks (ε_l) and the Markov Chain governing their evolution (π^ε_{lm}) similarly by discretizing a log-normal process, $\log\varepsilon' = \rho_\varepsilon \log\varepsilon + \eta'_\varepsilon$ using 15 values, and we assign the initial distribution of productivity signals, $T(s)$, as a discretized Pareto distribution with curvature parameter $\mathfrak{p}$.

We set the length of a period to correspond to one year, and we determine the values of β, ν, δ, α, and θ using moments from the aggregate data as follows. First, we set the household discount factor, β, to imply an average real interest rate of 4 percent, consistent with recent findings by Gomme, Ravikumar and Rupert (2008). Next, we set the production parameter ν to imply an average labor share of income at 0.60 (Cooley and Prescott (1995)). The depreciation rate, δ, is taken to imply an average investment-to-capital ratio at 0.069, corresponding to the average value for the private capital stock between 1954 and 2002 in the U.S. Fixed Asset Tables, controlling for growth. Given that value, we determine capital's share, α, so that our model matches the average private capital-to-output ratio over the same period, at 2.3, and we set the parameter governing the preference for leisure, θ, to imply an average of one-third of available time is spent in market work. The parameter set obtained from this part of our calibration exercise is summarized below.

β	ν	δ	α	θ	ρ_z	σ_{η_z}
0.962	0.60	0.069	0.26	2.58	0.909	0.015

4.2 Firm-level targets

The remaining parameters are jointly determined using moments from U.S. firm- and establishment-level data. Most of our target moments are drawn from the Business Employment Dynamics (BDS) database constructed from the Quarterly Census of Employment and Wages and maintained by the Bureau of Labor Statistics for the period 1977-2011. Beyond its public availability, an advantage of this annual data set relative to the establishment data in the Longitudinal Research Database (LRD) is that it includes all firms covered by state unemployment insurance programs, which accounts for roughly 98 percent of all nonfarm payrolls. Over 1979-2007, the average exit rate among firms in the BDS is 8.7 percent. Over the same period, the employment sizes of new and one-year old firms relative to employment in a typical firm are 28.46 and 37.07 percent, respectively, while the population shares of firms aged 1 and 2 years are 8.7 and 7.4 percent, respectively.

To discipline the extent of idiosyncratic volatility in our model, and to select the capital adjustment parameters, we target some establishment-level investment moments reported by Cooper and Haltiwanger (2006) from the LRD.[8] These include the standard deviation of investment rates (0.337) and the fraction of establishments with investment rates exceeding 20 percent (0.186). While our model has life-cycle aspects affecting firms' investments, the Cooper and Haltiwanger

[8] The distinction between firm and establishment may be relatively unimportant here; over 95 percent of firms in the BDS have fewer than 50 employees. A more problematic distinction is the fact that the LRD includes only manufacturing establishments.

(2006) dataset includes only large manufacturing establishments that remain in operation throughout their sample period. Thus, in undertaking this part of our calibration, we must select an appropriate model sample for comparability. This we do by simulating a large number of firms for 37 years, retaining only those firms that survive throughout, and then restricting the dates over which investment rates are measured to eliminate life-cycle effects.

We use Foster, Haltiwanger and Syverson's (2008) unweighted estimate of the annual persistence of firm-level total factor productivity to set $\rho_\varepsilon = 0.757$. Next, we assume that the random costs of entry, continuing in production and capital adjustment are drawn from uniform distributions and set the lower bounds of each to 0; $\gamma_L = \varphi_L = \xi_L = 0$. The remaining parameters are listed in the table below. First, there are the stock of production locations available for use by potential firms and those currently producing, Q, and the curvature of the Pareto distribution of potential firms' productivity signals, $\mathfrak{p}$. Next, there are the upper bounds of the costs of entry, operation and capital adjustment, (γ_U, φ_U, ξ_U). Finally, there are the quadratic capital adjustment cost term, c_q, the fraction of capital lost when a firm exits, λ, and the standard deviation of innovations to firm-specific productivity, $\sigma_{\eta_\varepsilon}$. We determine these parameter values by targeting the 7 empirical moments discussed above and jointly selecting Q to normalize the steady state number of firms in production at 1.0.

Q	$\mathfrak{p}$	γ_U	φ_U	ξ_U	c_q	λ	$\sigma_{\eta_\varepsilon}$
1.84	5.2	0.072	0.13	0.014	0.064	0.03	0.182

Here are the results for several additional firm-level moments not targeted in our calibration. The mean establishment-level investment rate is 12.2 percent in Cooper and Haltiwanger's (2006) dataset, and its autocorrelation is 0.058; the corresponding moments in our model are 11.0 percent and 0.096. The survival rate to age 5 is 50.73 percent among firms in the BDS over 1981-2007, whereas our model predicts 62.06 percent. Continuing to focus on young firms, the average relative employment size among firms aged 0 to 5 is 38.78 percent in the data, and their overall employment share is 17.05 percent; the corresponding moments in our model are 54.09 and 20.83 percent, respectively. In the data, the average exit rate among firms of age 1 is 21.33 percent, and that among age-2 firms is 15.85 percent. Exit rates at ages 1 and 2 in our model are 17.18 percent and 10.93 percent.

4.3 Numerical method

The distribution μ in the aggregate state vector of our model economy is a large object. In general, discrete choices imply that this distribution is highly non-parametric. For each level of productivity, we store the conditional distribution using a fine grid defined over capital. However, firms' choices of investment are not restricted to conform to this grid. To allow the possibility that nonconvex capital adjustment may interact with endogenous entry and exit over the business cycle in a way that delivers aggregate nonlinearities, we adopt a nonlinear solution method. Given μ, an exact solution is obviously numerically intractable; thus, we use selected moments of μ as a proxy for the distribution in the aggregate state vector when computing expectations.

Our solution method is an adaptation of that in Khan and Thomas (2008). Following the approach developed by Krusell and Smith (1997, 1998), we assume that firms approximate the distribution in the aggregate state vector with a vector of moments, $m = (m_1, ..., m_I)$, drawn from the true distribution. Because our model implies a discrete distribution over k and over ε, conditional means from I equal-sized partitions of the capital distribution work well, implying small forecasting errors.

As in Krusell and Smith (1997), we solve our model by iterating between an inner loop step and an outer loop step until we isolate forecasting rules satisfyingly consistent with equilibrium outcomes. In the inner loop, we take as given a current set of forecasting rules for p and m' and use them to solve incumbent firms' expected value functions V^0 (from section 3). This we do by combining value function iteration with multivariate piecewise polynomial cubic spline interpolation allowing firms to evaluate and select off-grid options. We next move to the outer loop to simulate the economy for 5000 periods. The current set of m' forecast rules are used in the outer loop, while p is endogenously determined in each date. Each period in the simulation begins with the actual distribution of firms over capital and productivity implied by the decisions of the previous date. Given incumbent firms' value functions from the most recent inner loop, and the aggregation of section 3.4, we determine equilibrium prices and quantities, and thus the subsequent period's distribution. Once the simulation has finished, we use the resulting data to update the forecasting rules, with which we return to the inner loop.[9]

[9] Despite the rich distribution of firms, agents closely predict prices and the proxy state with the distribution summarized by only mean capital. Recall the forecast rules for p, conditional on current z, are used only in solving for firm value functions in the inner loop. The R^2 for these forecast rules are all above 0.9999, and standard errors are below 0.0002. The conditional forecast rules for m' have R^2 above 0.9987 and standard errors below 0.0013.

5 Results

5.1 Steady state

In this section, we explore aspects of our model in its steady state and briefly compare it to an otherwise identical reference model without entry and exit. In the reference model, an exogenous stock of firms produces each period exempt from fixed costs of operation. As noted above, the stock of firms there is fixed at the steady state number of firms in our full model.

On average, our model economy forfeits roughly 9 percent of its GDP to operating costs. However, the average level of consumption is 1.03 times that in the reference model with no such costs. This is achieved in part by the fact that households work roughly 13.1 percent more in our economy. However, the more direct explanation lies in the distribution of firms over productivity levels, which encourages this higher work effort and supports 16.5 percent more investment.

FIGURE 1. How selection shapes the productivity distribution

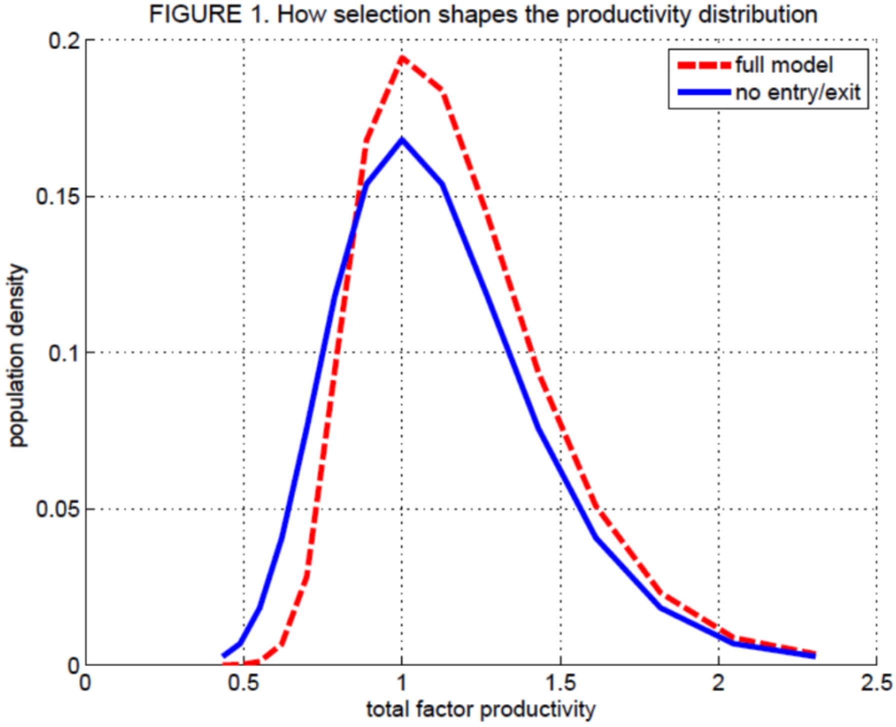


Figure 1 compares our model's stationary distribution of firms over TFP to that in the model without entry and exit. All else equal, firms with relatively low productivities are induced to exit our model economy by the costs they must pay to remain. Furthermore, fixed entry costs induce those potential firms with relatively low productivity signals to stay out. As such, the typical exiting firm is replaced by an entrant with higher productivity. Given these aspects of selection, the stationary distribution of firms in our model economy has less mass over lower productivity levels and more mass in higher regions of productivity than does the reference model. This raises average productivity by 7.2 percent, encouraging households to work and save more.

Figures 2 and 4 (below) display the stationary distributions of firms in our economy at the start of a period and at the time of production, respectively. In each of these figures, population density increases as one looks toward the back right corner representing the highest levels of capital and productivity. Comparing the start-of-period distribution to that remaining at production time, we see how selection generates these shapes.

FIGURE 2. Start of period potential producer distribution

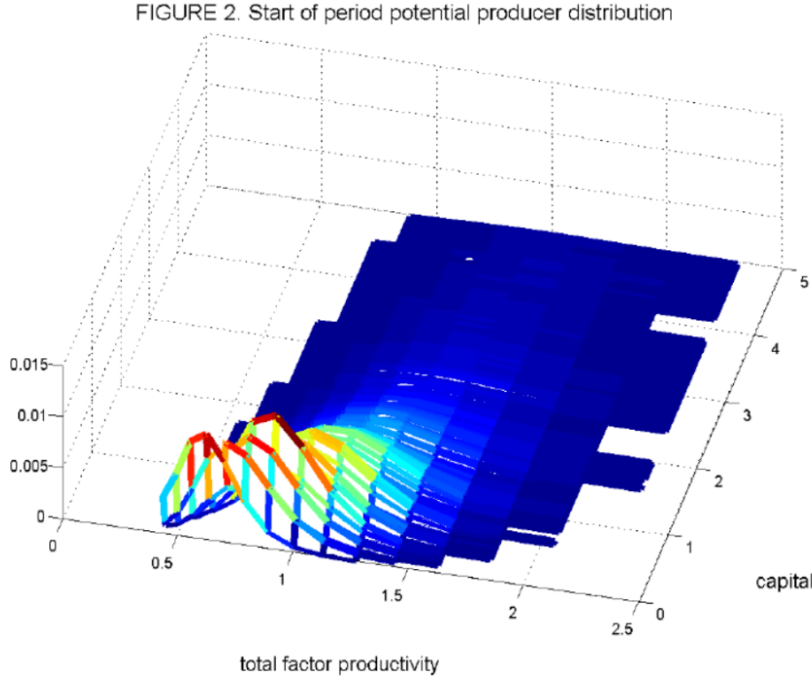


Figure 3 shows the steady state distribution of entrants in their first year of production, the startups from Figure 2 that choose to produce. Given the distribution of initial productivity signals, entrants are mostly concentrated in the lower ranges of productivity and capital. Looking to the distribution of all producers in Figure 4, we see the mass of firms expanding into higher productivities and capital levels. This shows that our model is consistent with the empirical evidence that young firms are smaller and less productive than the typical firm. Conditional on survival, young firms become more productive and larger over time as they approach maturity.

FIGURE 3. Production-time distribution: New firms

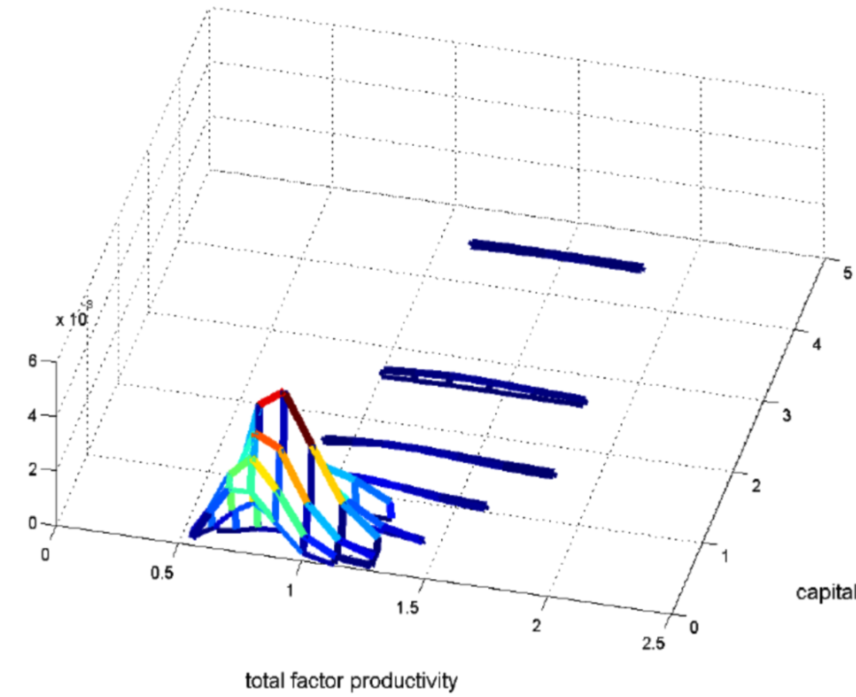


FIGURE 4. Production-time distribution: All Firms

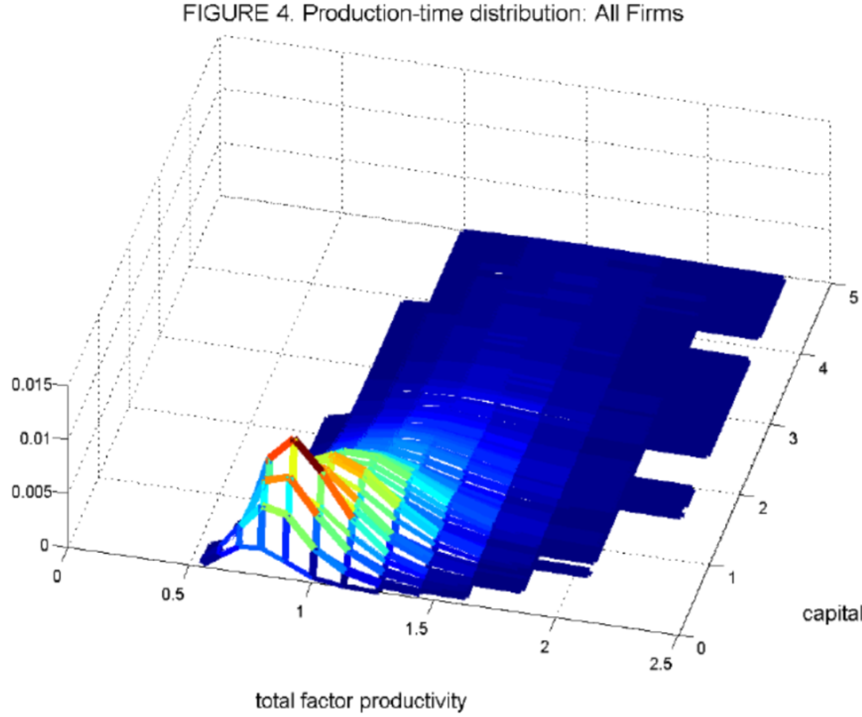
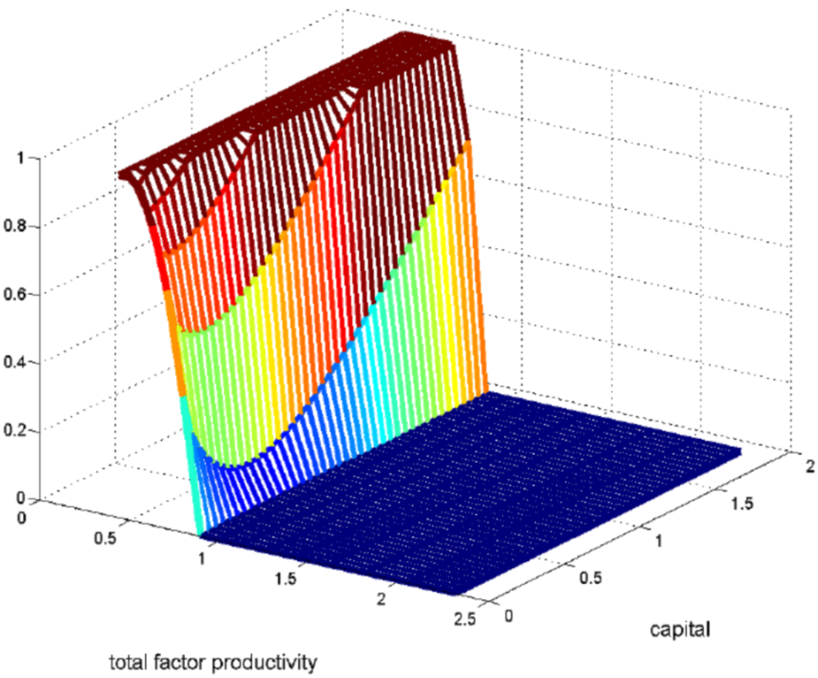


Figure 5 displays our steady state exit hazard for startups and incumbent firms. The patterns here arise naturally from two facts: (i) firm values are increasing in both capital and productivity, while (ii) convex and nonconvex adjustment costs distort optimal capital reallocation.

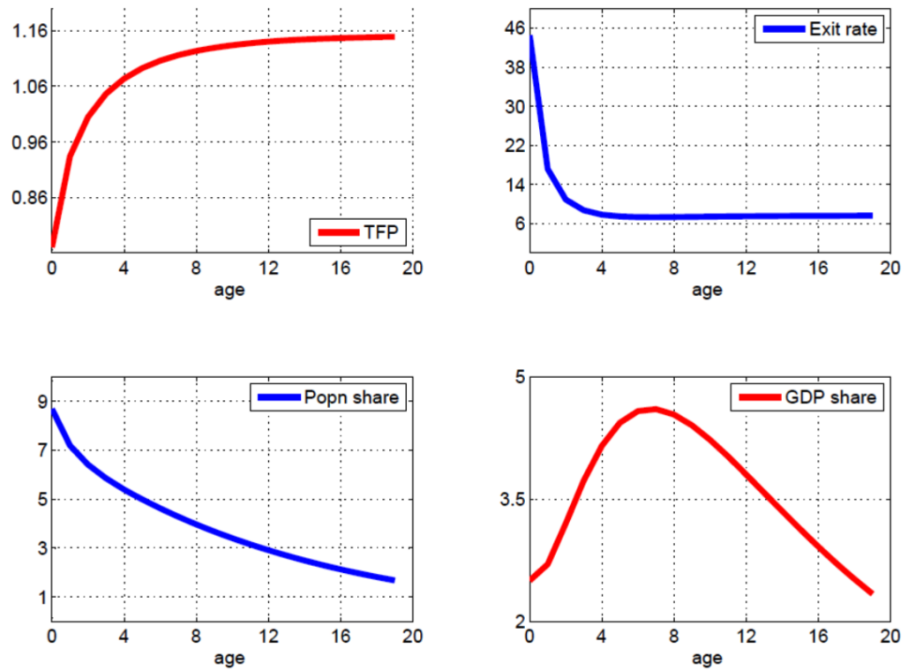
FIGURE 5. Exit hazard



At productivity ranges above 0.89, irrespective of capital, all firms are willing to pay the highest operating costs; so no firm exits. Elsewhere, for any given capital stock, selection implies that exit probabilities rise as TFP falls. On the other hand, if we condition on productivity, the probability of exit is non-monotone in firm size. Absent costs of capital adjustment, the hazard would always fall in capital (given higher flow profits and the fact that a fraction of the firm's capital is lost when it exits). Here, however, some firms with large capital stocks and low productivity prefer to exit rather than pay operating costs and also suffer adjustment costs to shed their excess capital. As a result, at firm-TFP levels below around 0.8, the exit hazard is u-shaped.

We conclude this section with a more direct look at firm life-cycle dynamics. Figure 6 tracks an initially large cohort of startup firms as it ages across 20 years in our model's steady state. By allowing mean-reverting idiosyncratic productivities alongside fixed entry and operating costs, our model obtains the selection-based successes of Hopenhayn's (1992) original model of industry dynamics. Here, as there, the average productivity and value of surviving members within a cohort rise as the cohort ages, so exit rates fall with age.

FIGURE 6. Productivity, Exit, Population and Output By Age



From the top right panel of Figure 6, we see exit rates fall off sharply from the 44 percent failure rate of startups, to roughly 17 percent exit among one-year old firms, about 11 percent among firms at age 2, and less than 9 percent by age 3. It takes the typical firm roughly 14 years to reach its ultimate productivity in the top left panel of the figure, although the half-life from its first date of production is only about 3 years. Older firms tend to have higher productivity than young firms, and they experience mean-reversion in their productivities. Thus, we easily obtain an unconditional negative relationship between firm size and growth, and between firm age and growth, as found in the data by Dunne, Roberts and Samuelson (1989), Haltiwanger, Jarmin and Miranda (2013) and other studies.

The inclusion of one-period time-to-build capital stocks breaks the perfect mapping between firm productivity and size inherent in the Hopenhayn model. Thus, our model is capable of a negative correlation between age and growth conditional on size consistent with the empirical findings of Evans (1987), Hall (1987) and Haltiwanger, Jarmin and Miranda (2013). Firms cannot immediately adjust their capital inputs in response to changes in their productivities, so those with large employment levels need not necessarily have high productivity. This muddying effect is compounded by the presence of capital adjustment frictions. Consider two firms that are small - one young, one old. Given the productivity profile for the typical firm in a cohort as it ages, the young firm is likely to be small simply because it has not yet reached maturity. For any such firm, the joint rise in TFP and capital will imply sharp employment growth. The average growth rate for firms between age 1 and 2 is about 30 percent; whereas the average growth rate between age 8 and 9 is about 6 percent. On the other hand, an old firm can be small only due to a series of poor productivity draws. So long as this firm's productivity does not rise, it will either reduce its capital stock or simply maintain it, so its employment growth is negative or zero.

Finally, consider the implications of Figure 6 for the timing of when the greatest damage from a 'lost generation' of entrants might be felt in our economy. In ordinary times, a young cohort closes roughly three-quarters of the gap to its ultimate productivity by age 4, when its population share is still relatively high. Taking into account the gradualism implied by firm-level capital accumulation, the cohort has its greatest contribution to aggregate production (cohort output/GDP) at ages 6 and 7 (lower right panel). Thus, to the extent that an aggregate shock to our economy causes a large reduction in firm entry, we may expect to see the largest effects of those losses roughly six years on.

5.2 Aggregate fluctuations

We begin this section by considering how endogenous firm life-cycle dynamics alter the cyclical movements in GDP, employment and other series when fluctuations are driven solely by aggregate productivity shocks. In response to a fall in the exogenous component of aggregate TFP, potential firms realizing any given firm-level productivity signal anticipate lower value relative to an ordinary date. Thus, fewer among them choose to invest toward becoming startups in the next period. For the same reasons, the numbers of startups choosing to pay their operating costs to enter production in the current period also fall, while the numbers of incumbent firms exiting rise. We will see below that these choices drive procyclical entry and countercyclical exit in our model, as in the data. Such changes have the potential to exacerbate the movements in employment and GDP; however, note that the most affected firms are those with low relative productivities, so selection should have some stabilizing influence. Among firms with the same productivity, one might expect that larger firms would be more likely to survive. Recall from Figure 5 that this need not be the case, however, given the implications of micro-level capital adjustment frictions.

To consider how entry and exit decisions reshape the typical business cycle, we first compare HP-filtered moments from our model to those from the reference model described above wherein a fixed set of firms lives forever. Table 1 examines volatility and contemporaneous comovement in the two settings. Despite hindrances to capital reallocation, both economies have the usual traits of an equilibrium business cycle model in terms of their relative volatilities and contemporaneous correlations with GDP. Our economy has a bit more volatility in overall GDP and consumption in comparison with the reference model. The differences in employment and capital investment are more pronounced. Changes in the number and age composition of firms drive higher volatility in both series and weaken their correlations with GDP.

TABLE 1. Volatilities and contemporaneous correlations with GDP

	GDP	Consumption	Investment	Employment
A: RELATIVE STD. DEV.				
No Entry/Exit	(1.994)	0.501	3.797	0.559
Full Model	(2.001)	0.505	3.931	0.568
B: CORRELATION				
No Entry/Exit	1.000	0.933	0.966	0.947
Full Model	1.000	0.924	0.956	0.941

Table 2 presents the cross-date correlations of entry and exit rates and the measure of firms with GDP in our model. The first column confirms that our model delivers procyclical entry and countercyclical exit. Given the one-period time-to-build nature in the creation of entrants, the strongest relationship between GDP and the number of entrants comes with a one-year lag. This explains the significantly positive correlation between GDP and the date $t+1$ entry rate, despite procyclical movements in the number of firms at date t. Exit is less persistent. Nonetheless, on the whole, movements in the number of producers are protracted; there, the contemporaneous correlation with GDP is lower than the correlations at both date t+1 and date t+2.

TABLE 2. Cross-date correlations with current GDP

	$t+0$	$t+1$	$t+2$	$t+3$	$t+4$
Entry rate	0.774	0.474	-0.238	-0.504	-0.526
Exit rate	-0.919	-0.126	0.236	0.303	0.270
Firms	0.636	0.791	0.645	0.412	0.188

Across Tables 1 and 2, we have seen that cyclical changes in firms' entry and exit decisions amplify TFP-driven business cycles. Given the usual firm life-cycle patterns presented in Figure

6, such changes also have the potential to add persistence to the movements in GDP. Consider the fact that an unusually small number of new firms enters into production following a negative TFP shock. Over subsequent periods, the typical surviving member of this smaller-than-average cohort of young firms grows in productivity and size. We noted in closing section 5.1 that, in ordinary times, a young cohort contributes increasingly to GDP as it nears age 6. Thus, early reductions in the numbers of entering firms in response to a negative aggregate TFP shock can hold aggregate production down at later dates, even as the exogenous component of aggregate productivity reverts toward its mean, thereby protracting a TFP-driven recession. We will see evidence of this phenomenon below in Figure 8. However, it is hard to detect in conventional second moments. Table 3 shows no evidence in our model's GDP correlations over short time horizons; indeed, the autocorrelations of GDP over leads 1 - 3 are weaker than in the reference model without entry and exit. The only suggestion of increased propagation comes in the final column, where our model has weaker negative correlations of GDP with itself and exogenous TFP at date t+5 than the reference model. This is broadly consistent with our reasoning above suggesting that changes in entry rates following a shock in date t should become important in gradualizing mean reversion of aggregate production after 4 to 6 periods.[10]

TABLE 3. Persistence and the propagation of shocks

	GDP(t+1)	GDP(t+2)	GDP(t+3)	GDP(t+4)	GDP(t+5)
A: CORRNS WITH GDP(t)					
No Entry/Exit	0.505	0.138	−0.084	−0.200	−0.254
Full Model	0.457	0.087	−0.112	−0.199	−0.231
B: CORRNS WITH Z(t)					
No Entry/Exit	0.510	0.145	−0.076	−0.192	−0.247
Full Model	0.464	0.097	−0.103	−0.192	−0.226

For some further insight into the moments reported in Tables 1 - 3, we examine impulse responses following a persistent negative shock to the exogenous component of aggregate productivity. We begin with the reference model in Figure 7. There, GDP falls roughly 3.2 percent at the shock's impact, while employment and capital investment fall roughly 1.8 percent and 11.6 percent, respectively. Thereafter, these series display the usual mean-reversion seen in business

[10] The correlation of GDP with itself at date $t+6$ is -0.264 in the reference model and -0.227 in our model.

cycle models driven by AR(1) shocks. Consumption and the real wage also exhibit the usual u-shape of a business cycle model with indivisible labor preferences.

Examining the lower right panel of Figure 7, notice that there is virtually no difference in the response of measured versus exogenous productivity in the reference model despite the presence of capital reallocation frictions. This is consistent with results in Khan and Thomas (2003, 2008), although those studies did not include convex adjustment costs. The payment of such costs is included in broad investment (as will be entry and operation costs in our full model below); however these are sufficiently minor and unaffected by the shock as to imply little difference between the narrow and broad investment responses.[11]

FIGURE 7. Productivity shock: No Entry & Exit

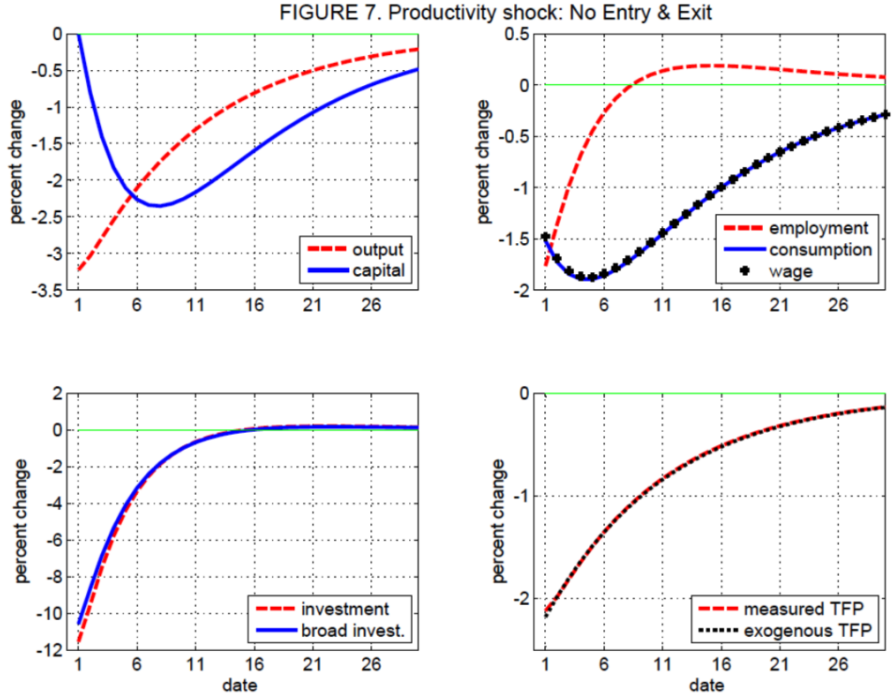


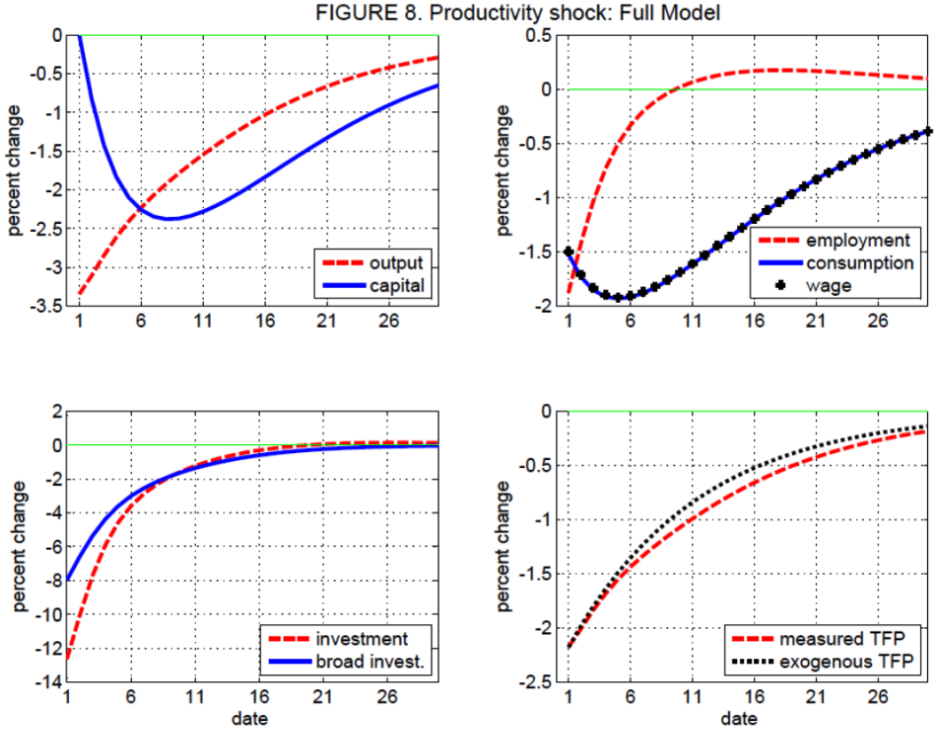
Figure 8 shows responses to the same productivity shock in our full model economy. GDP falls about 0.13 percent more in Figure 8, and there is a similar difference in the employment response. Capital investment falls about 1.1 percent further relative to the reference model, whereas broad investment falls 2.6 percent less.[12] The latter reflects the fact that firms' investment in the form

[11] The same is true of the labor-denominated nonconvex adjustment costs in both models. Thus, we do not report responses for the narrow employment series.

[12] The size of the TFP shock here matches the detrended measured Solow residual at the trough of the 2007 recession. Differences grow with larger shocks; under a one standard deviation negative shock, GDP and hours fall about 0.2 percent more in our model than in the reference model, while capital investment falls 1.7 percent more.

of fixed entry and operating costs is far less responsive to the shock than is their investment in capital. The former arises from an anticipated endogenous drag on aggregate productivity evident in the lower right panel. Its absence from Figure 7 implies that it is driven entirely by the disruption in firm-level entry and exit decisions.

FIGURE 8. Productivity shock: Full Model

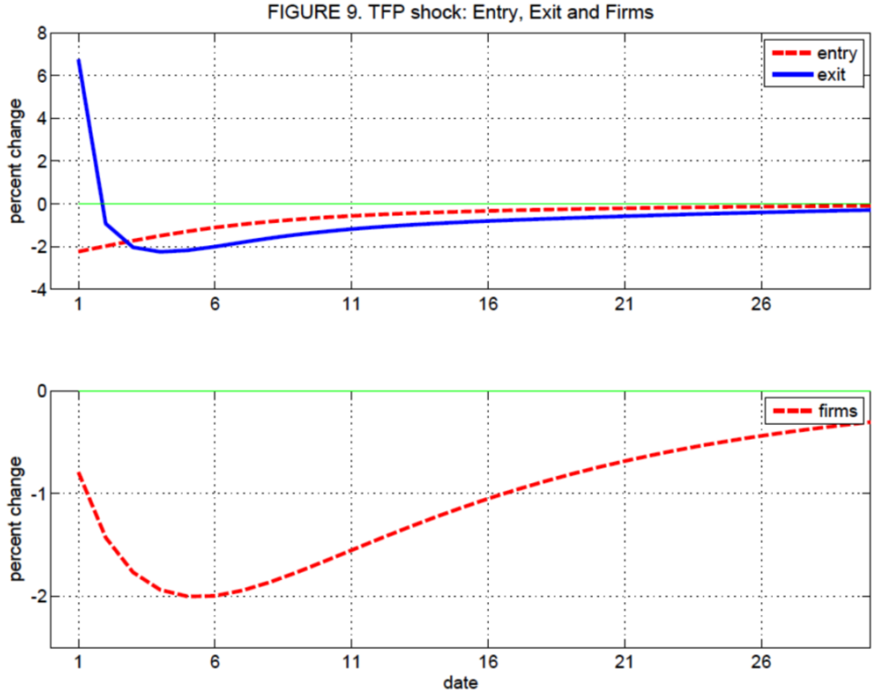


Beyond the modest amplification, our model also implies increased persistence. The half-life of GDP's response is 8.8 years in the reference model, whereas it is 9.9 years in our model. Thereafter, our recovery is more appreciably gradualized; for example, GDP reaches 1 percent below normal two years later in our model. This is a direct result of the persistent TFP wedge in Figure 8, and the fact that it steadily widens up until date 13.

To understand why the overall propagation of a TFP shock is altered in our model, we turn to the responses in market participation that distinguish it in Figure 9. Recall from Figure 6 that the essential mechanism we anticipated would offset mean-reversion in exogenous TFP and hold aggregate production down longer was a missing generation effect, the growth phase of a smaller-than-usual cohort of young firms following the shock. Figure 9 shows that the number of entrants falls roughly 2.2 percent below normal at the date of the shock and recovers about half way over the next five years. Thus, several cohorts of young firms are appreciably reduced.

These new cohorts fail to replace an initially large number of exiting firms. Exit rises about 6.8 percent at date 1, and the rate of failure among firms remains unusually high for several periods; hence, we see a u-shaped response in the measure of operating firms.[13]

FIGURE 9. TFP shock: Entry, Exit and Firms



So far, we have studied how entry, exit and selection contribute to the mechanics of a typical recession. We next consider their role in a Great Recession such as the U.S. 2007-9 experience. A growing body of evidence suggests that this recession was unusual not simply in its very large and persistent declines in GDP, employment and investment, but also in the disparate effects on employment in firms of different age and size. Examining BDS data, Sedlacek (2013) and Sedlacek and Sterk (2014) document disproportionate negative effects on young firms, while Khan and Thomas (2013) and Siemer (2013) show small firms were far more affected than large firms. Elsewhere, indirect evidence suggests that the recession originated in a shock in the financial sector (Almeida et al. (2009), Duchin et al. (2010)). To the extent that young, small firms are more reliant on external finance or have more difficulty accessing it, the financial shock story fits well with the findings from the BDS noted here.

[13] A startup in place at the date of the shock faces a 45.8 percent chance of immediate failure (versus 44.5 percent chance in ordinary times); conditional on entry it faces a 17.7 percent chance of exiting at age 1 (versus 17.2 percent) and, conditional on survival at age 1, an 11.2 percent chance of exit at age 2 (versus 10.9 percent).

Three striking observations regarding the firm distribution distinguish the Great Recession relative to a typical recession. First, the total number of firms fell 5 percent (Siemer (2013)). Second, the number of young (age 5 and below) firms fell 15 percent (Sedlacek (2013)). Third, the relative employment decline among small firms (those with less than 100 employees) was disproportionate in comparison to such firms' usual employment share and roughly double the decline among large firms (Khan and Thomas (2013)). Irrespective of whether the impetus was a financial shock as conventionally perceived, it caused a large disruption in the distribution of firms and distorted firm life-cycle dynamics.

In our final set of impulse responses, we proxy for the implications of financial disruption by adding an aggregate shock with uneven effects on firms' decisions. Specifically, alongside the TFP shock above, we consider a 5 percent rise in the upper support of the operating cost distribution; φ_U rises to .273 and thereafter follows an AR(1) return with persistence $\rho = \rho_z$. Unlike an aggregate productivity shock, this disturbance has minimal direct consequence for large, old firms in the economy; while increased operating costs reduce their cash flows, such costs are typically small relative to the value they place on continuing in operation. On the other hand, given the ordinary life-cycle productivity and exit patterns shown in Figure 6, raised operating costs are quite important for the decisions of young, small firms and startups; on average, these firms have lower productivity and are far more vulnerable to failure than others.

FIGURE 10. TFP & Operating Cost shock: Entry, Exit and Firms

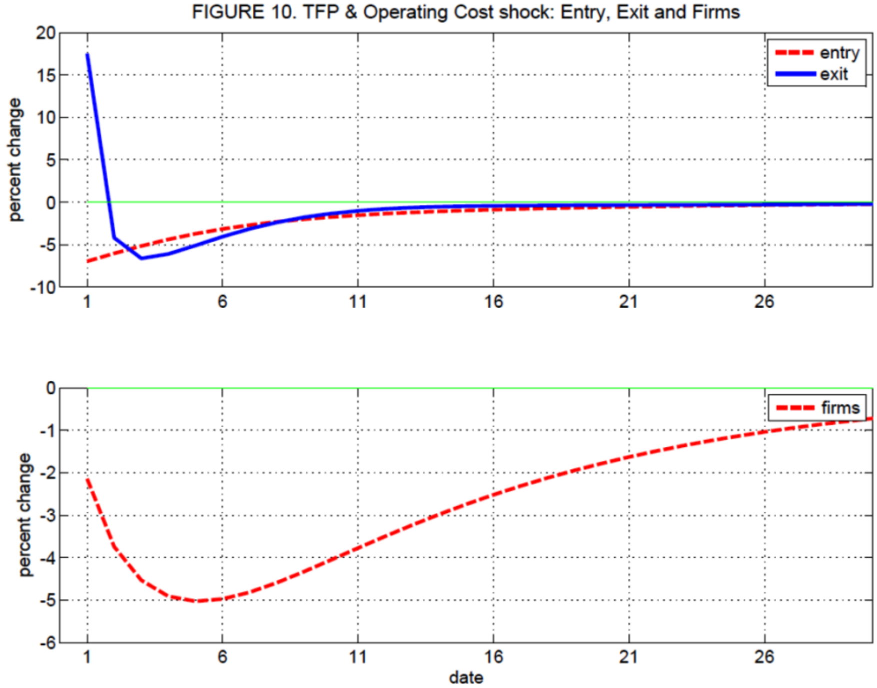


Figure 10 shows the overall effect our second shock has on entry, exit and the number of firms. With an increase to the costs of operating, the rise in exit and the fall in entry more than double in comparison to Figure 9. This carries over into the number of firms in the bottom panel, and generates an ultimate drop matching the 5 percent fall over the Great Recession. The overall number of operating firms is an important input into aggregate production, given decreasing returns at the firm. However, recall from Figure 6 that firms of different ages are not equally valuable. The greatest contributions to GDP come from firms aged 6-8 as they move toward maturity with growing productivity and size, while still relatively large in their numbers. Thus, an important aspect of our proxy financial shock is the fact that it disproportionately eliminates young firms, exacerbating the missing generation effect we saw above.[14]

FIGURE 11. TFP & Operating Cost shock: Full Model

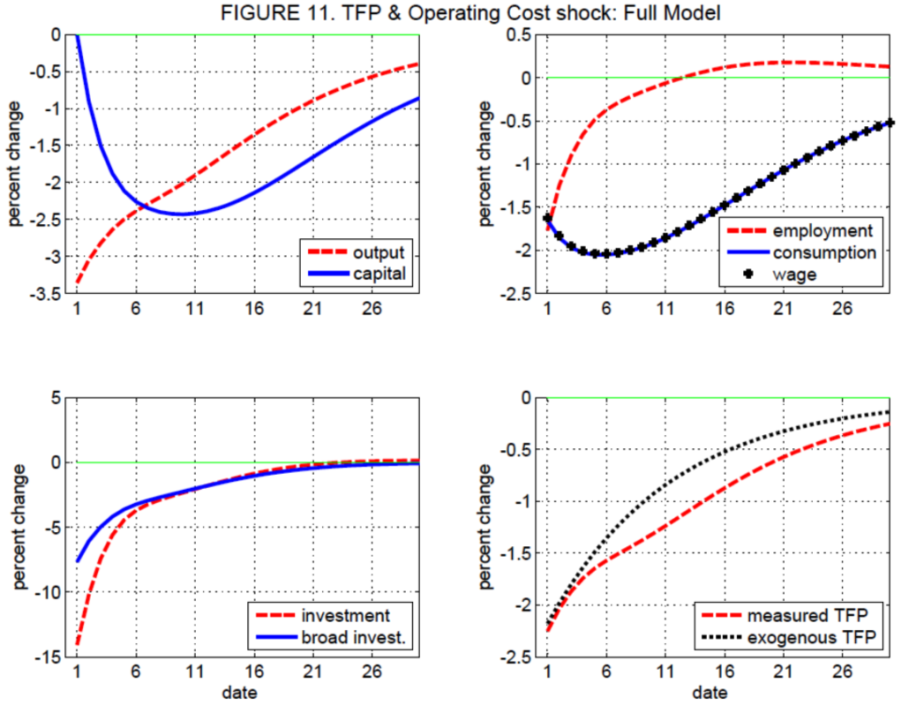


Figure 11 is our model counterpart to the Great Recession. In comparison to Figure 9, the more direct destruction of young firms here amplifies the downturn, as it yields an endogenous drop in aggregate TFP at its impact. Larger differences are seen in later dates as the aggregate shocks mean revert. The consequences of the missing generation grow prominent in measured TFP after

[14]A startup deciding whether to produce at date 1 now has a 48.4 percent chance of failure (versus 44.5 percent in steady state); conditional on entry it has an 18.5 percent chance of exiting at date 2 when it is aged 1 (versus 17.2 percent) and, conditional on survival through age 1, an 11.6 percent chance of exit at age 2 (versus 10.9 percent).

about 5 periods. The effects steadily grow over the next 6 years, increasing the downward pull on productivity to a roughly 0.4 percent TFP wedge that lasts over many subsequent dates. This slows GDP's return markedly, extending its half-life by more than 3 years.

6 Concluding remarks

In the sections above, we have developed a dynamic, stochastic general equilibrium model allowing for time-varying entry and exit in a setting where firms face persistent shocks to aggregate and individual productivity, and they must pay fixed costs to enter and to continue in production. Our firms' decisions regarding entry and their subsequent continuation are affected not only by their expected productivities, but also by capital reallocation frictions, and thus by their existing stocks. We have explored this model toward arriving at a better understanding of whether and how changes in firm entry and exit rates and the composition of firms affects aggregate fluctuations in an environment with realistic firm-level investment patterns and life-cycle dynamics.

Based on an examination of second moments and impulse responses, we have seen that changes in entry and exit amplify responses over a typical business cycle driven by a disturbance to aggregate productivity, and they protract them. Our model amplifies a standard recession because it delivers an endogenous TFP wedge through procyclical movements in entry and countercyclical exit. Recovery is gradualized because the endogenous productivity effects grow over time. That, in turn, happens because young firms fail to replace a raised number of exiting firms in early dates following a shock, and the overall measure of producers falls over time. This missing generation effect is most prominent in GDP at the time when the reduced young cohorts are nearing maturity, and it lingers many years thereafter.

Changes in the number of firms, and particularly in the numbers of young firms, were dramatic over the U.S. 2007 recession. In an exercise designed to emulate this unusual episode, we have also considered a shock to firms' fixed operating costs. This might be interpreted as a loose proxy for a disruption to external finance in that it most directly affects entry decisions and the exit decisions of younger firms, given selection and their relatively low average productivity levels. We have seen that such a shock sharpens the missing generation effect, delivering a more pronounced cumulating drag on aggregate productivity and thus a far more anemic economic recovery.

On balance, the abstraction from endogenous firm entry and exit decisions in existing business cycle studies may be fairly innocuous where aggregate fluctuations driven by shocks of a first-moment, level nature are concerned. However, from the simple alternative example we have considered here, such abstraction is more costly when we try to reconcile our models' predictions with macroeconomic responses following shocks of a second-moment nature that disproportionately influence the decisions of vulnerable firms, as for example a large financial shock.

References

[1] Almeida, H., Campello, M. , Laranjeira, B. and S. Weisbenner. 2009. "Corporate Debt Maturity and the Real Effects of the 2007 Credit Crisis." NBER working paper No. 14990.

[2] Caballero, R. J and E. M. R. A. Engel (1999) "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S, s) Approach," *Econometrica* 67, 783-826.

[3] Clementi, G. and D. Palazzo (2010) "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations," NYU working paper.

[4] Cociuba, S. E., Prescott, E. C. and A. Ueberfeldt (2012) "U.S. Hours and Productivity Behavior Using CPS Hours Worked Data: 1947-III to 2011" University of Western Ontario working paper.

[5] Comin, D. and T. Philippon (2005) "The Rise in Firm-Level Volatility: Causes and Consequences," NBER Working Paper No. 11388.

[6] Cooley, T. F., and E. C. Prescott (1995) "Economic Growth and Business Cycles." in *Frontiers of Business Cycle Research*, Cooley, T. F. Editor, Princeton University Press.

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

[8] Cooper, R. W. and J. C. Haltiwanger (2006) "On the Nature of Capital Adjustment Costs," *Review of Economic Studies* 73, 611-633.

[9] Steven J. Davis, S.J. and J.C. Haltiwanger (1992) "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics* 107(3), 19-63.

[10] Duchin, R., Ozbas, O. and B. A. Sensoy. 2010. "Costly external finance, corporate investment, and the subprime mortgage credit crisis," *Journal of Financial Economics* 97, 418-35.

[11] Dunne, T., M. Roberts, and L. Samuelson (1989) "The Growth and Failure of US Manufacturing Plants," *Quarterly Journal of Economics* 94, 671–98.

[12] Evans, D. (1987) "Tests of Alternative Theories of Firm Growth," *Journal of Political Economy* 95(4), 657–74.

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

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

[15] Gomme, P., Ravikumar, B. and P. Rupert (2008) "The Return to Capital and the Business Cycle," Concordia University Working Paper 08-002.

[16] Haltiwanger, J., Jarmin, R. and J. Miranda (2013) "Who Creates Jobs? Small vs. Large vs. Young," *Review of Economics and Statistics* 95(2), 347-61.

[17] Hopenhayn, H. (1992) "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica* 60(5), 1127–50.

[18] Johnson, S. A. (1989) *Spline approximation in discrete dynamic programming with application to stochastic multi-reservoir systems* Unpublished dissertation (Cornell, Ithaca, NY).

[19] Khan, A., Senga, T. and J. K. Thomas (2014) "Default Risk and Aggregate Fluctuations in an Economy with Production Heterogeneity," Ohio State University working paper.

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

[21] Khan, A. and J. K. Thomas (2008) "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics" *Econometrica* 76, 395-436.

[22] Krusell, P. and A.A. Smith Jr. (1997) "Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns," *Macroeconomic Dynamics* 1, 387-422.

[23] Krusell, P., and A. A. Smith Jr. (1998) "Income and wealth heterogeneity in the macroeconomy," *Journal of Political Economy* 106, 867–96.

[24] Lee, Y., and T. Mukoyama (2009) "Entry, Exit, and Plant–level Dynamics over the Business Cycle," University of Virginia working paper.

[25] Prescott, E. C. (1986) "Theory Ahead of Business Cycle Measurement," *Federal Reserve Bank of Minneapolis Quarterly Review* 10, 9-22.

[26] Rogerson, R. (1988) "Indivisible Labor, Lotteries and Equilibrium," *Journal of Monetary Economics* 21, 3-16.

[27] Samaniego, R. M. (2008) "Entry, exit and business cycles in a general equilibrium model," *Review of Economic Dynamics* 11, 529–41.

[28] Sedlacek, P. (2013) "Firm Age, Business Cycles and the Aggregate Labor Market," University of Bonn working paper.

[29] Sedlacek, P. and V. Sterk (2014) "The Growth Potential of Startups over the Business Cycle," University of Bonn working paper.

[30] Siemer, M. (2013) "Firm Entry and Employment Dynamics in the Great Recession" Boston University working paper.