Misallocation Cycles*

Cedric Ehouarne† Lars-Alexander Kuehn‡ David Schreindorfer§

November 2022

Abstract

The goal of this paper is to quantify the cyclical variation in firm-specific risk and study its aggregate consequences via the allocative efficiency of capital resources across firms. To this end, we estimate a general equilibrium model with firm heterogeneity and a representative household with Epstein-Zin preferences. Firms face investment frictions and permanent shocks, which feature time-variation in common idiosyncratic skewness. Quantitatively, the model replicates well the cyclical dynamics of the cross-sectional output growth and investment rate distributions. Economically, the model generates business cycles through inefficiencies in the allocation of capital across firms, which amounts to an average output gap of 4.5% relative to a frictionless model. These cycles arise because (i) permanent Gaussian shocks give rise to a power law distribution in firm size and (ii) rare negative Poisson shocks cause time-variation in common idiosyncratic skewness. Despite the absence of firm-level granularity, a power law in the firm size distribution implies that large inefficient firms dominate the economy, which hinders the household's ability to smooth consumption.

*We thank Rüdiger Bachmann, Yasser Boualam, Scott Cederburg, Andrea Eisfeldt, Lukas Schmid, Stephen Terry, seminar participants at Arizona State University (Economics and Finance), Carnegie Mellon University, the University of Southern California, Penn State University, and conference participants at the 6th Advances in Macro-Finance Tepper-LAEF Conference, 1st Arizona Junior Finance Conference, 2016 Duke-UNC Asset Pricing Conference, 2016 Society for Economic Dynamics Meeting, 2016 European Finance Association Meeting, 2016 BYU Red Rock Finance Conference, 2017 American Economic Association Meeting, and 2017 Western Finance Association Meeting for helpful comments and suggestions. The views expressed in this paper are the authors' and do not necessarily reflect the views of Bank of America.

†Bank of America, Model Risk Management, 1133 Av. of the Americas, 39th floor, New York, NY 10036, ehouarne@gmail.com

‡Tepper School of Business, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA15213, kuehn@cmu.edu

§W.P. Carey School of Business, Arizona State University, PO Box 873906, Tempe, AZ 85287, david.schreindorfer@asu.edu

Contents

1 Introduction

A large body of research has shown that the cross-section of firms is characterized by a substantial degree of productivity and capital heterogeneity (e.g., Eisfeldt and Rampini (2006)). While the empirical facts about firm heterogeneity are well known, the aggregate consequences are not well understood. In this paper, we develop and estimate a simple general equilibrium model to illustrate how the dynamics of the cross-section of firms impact aggregate fluctuations and risk premia via the misallocation of capital resources. The key implication of our general equilibrium model is that idiosyncratic shocks do not integrate out at the aggregate level but instead generate cyclical movements in the higher moments of consumption growth and risk premia.

Our model is driven by a cross-section of heterogenous firms, which face irreversible investment decisions, exit, and permanent idiosyncratic and aggregate productivity shocks. The representative household has recursive preferences and consumes aggregate dividends. To generate aggregate consequences from a continuum of idiosyncratic shocks via capital misallocation, our model mechanism requires both a *power law distribution* as well as *common idiosyncratic skewness* in productivity.

While most of the literature assumes a log-normal idiosyncratic productivity distribution arising from mean-reverting Gaussian shocks, idiosyncratic shocks are permanent and follow random walks in our model. With firm exit, distributions of random walks generate power laws as emphasized by Gabaix (1999) and Luttmer (2007). Quantitatively, the endogenous power law for firm size is consistent with the data, as reported in Axtell (2001), such that the largest 5% of firms generate more than 30% of consumption and output in our model.

In addition to the random walk assumption, we model innovations to idiosyncratic productivity not only with Gaussian but also with negative Poisson shocks, which induce common idiosyncratic skewness. These negative Poisson shocks do not capture rare aggregate disaster, as in Gourio (2012), because they wash out at the aggregate level in a frictionless model.[1] Instead, time variation in the size of common idiosyncratic skewness allows us to capture

[1]For disaster risk in consumption see Barro (2006), Gabaix (2012), and Wachter (2013).

the cyclicality in the skewness of cross-sectional sales growth, consistent with the evidence in Salgado et al. (2015).

In the model, these features lead to large occasional inefficiencies in the allocation of capital across firms and it hinders the representative agent's ability to smooth consumption. Intuitively, in recessions aggregate productivity falls and the distribution of output growth becomes negatively skewed. Firms with negative idiosyncratic productivity draws find it difficult to disinvest unproductive capital to raise dividends. At the same time, the representative household would like to reallocate capital to smooth consumption.

Because of the power law distribution in firm size, the share of output coming from large firms contributes disproportionally to aggregate consumption, so that negatively skewed shocks to their productivity are particularly painful. Consequently, the drop in dividends from the mass of constrained firms is large, given that they are large in size. While unconstrained firms increase dividends by reducing investment, they are smaller so that they are not able to offset the impact of large constrained firms on aggregate consumption. This effect implies that in recessions aggregate consumption falls by more than aggregate productivity, causing negative skewness and kurtosis, and it arises purely from the cross-sectional misallocation. In contrast, in models with log-normal productivity distributions, the size difference between constrained and unconstrained firms is small so that the groups offset each other.

While the impact of capital misallocation on output and consumption are short lived under temporary mean-reverting shocks, permanent Poisson shocks render misallocation distortions long lasting. Quantitatively, output and consumption growth become more volatile and persistent, even though the model is only driven by i.i.d. innovations. Importantly, consumption growth is left skewed and leptokurtic, as in the data. Because the household cares about long lasting consumption distortions due to Epstein-Zin preferences, the welfare costs of capital misallocation and aggregate risk premia are large.

Our mechanism to generate aggregate fluctuations from idiosyncratic shocks obeying a power law is distinct from the granular hypothesis of Gabaix (2011). While Gabaix also argues that the dynamics of large firms matters for business cycles, he relies on the fact that the number of firms is finite in an economy so that a few very large firms dominate aggregate output. The impact of these very large firms does not wash at the aggregate level when firm size follows a power law. In contrast, we model a continuum of firms such that each individual firm has zero mass. In our model, the power law in firm size generates aggregate fluctuations based on capital misallocation, arising from the investment friction, and not because the economy is populated by a finite number of firms. In reality, both effects are at work to shape business cycles.[2]

Methodologically, we build on Veracierto (2002) and Khan and Thomas (2008), who find that microeconomic investment frictions are inconsequential for aggregate fluctuations in models with mean-reverting idiosyncratic productivity.[3] We show that a model with permanent shocks and a more realistic firm size distribution not only breaks this irrelevance result, but also produces risk premia that are closer to the data. We are not the first to model permanent idiosyncratic shocks, e.g., Caballero and Engel (1999) and Bloom (2009) do so, but these papers study investment dynamics in partial equilibrium frameworks.

There is substantial empirical evidence that the riskiness of the economy is countercyclical, both at the aggregate and at the firm level. Starting with Bloom (2009), a large literature has used this observation to argue that shocks to uncertainty generate business cycles via wait-and-see effects in firms' investment and hiring policies. Our work differs from this prior literature in two important ways.

First, our theory does not feature wait-and-see effects because we deliberately model shocks to idiosyncratic risk as i.i.d. and unobservable ex ante. While recessions in our model are characterized by higher micro and macro volatility, risk shocks do not *cause* recessions. Rather, they lead to an amplification and propagation of downturns via their persistent effect on capital misallocation. Additionally, while aggregate shocks in our model are symmetrically distributed, aggregate output and consumption growth rates are unconditionally left-skewed because measured aggregate productivity falls more than true productivity during recessions. Our work thus provides an endogenous mechanism for the channel in Berger et al. (2016), who

[2] Related to the granular notion, Kelly et al. (2013) derive firm volatility in sparse networks.

[3] Bachmann and Bayer (2014) show that the same irrelevance result holds with idiosyncratic volatility shocks.

argue that the increased volatility observed during recessions is a consequence of negatively-skewed aggregate productivity technology shocks as opposed to a causal driver of recessions.

Second, while Bloom (2009) models risk shocks as a symmetric increase in the volatility of idiosyncratic risk, we assume that they operate through its left-skewness. In particular, we model idiosyncratic shocks as a combination of homoscedastic Gaussian innovations and negative Poisson jumps with time-varying size. As in Bloom's model, the dispersion of idiosyncratic shocks therefore increases during recessions in our model, but the effect is driven by a widening of the left tail of the shock distribution only. This makes our model consistent with the empirically observed distribution of firms' sales growth, which becomes strongly negatively skewed during recessions.

This empirical fact is also reminiscent of Guvenen et al. (2014), who document that households' income shocks feature procyclical skewness. Constantinides and Ghosh (2015) and Schmidt (2015) show that procyclical skewness is quantitatively important for aggregate asset prices in incomplete market economies. Different from these papers, our paper focuses on heterogeneity on the productive side of the economy and analyzes the effect of skewed shocks on capital misallocation.

The first study to quantify capital misallocation is Olley and Pakes (1996). More recent contributions include Hsieh and Klenow (2009) and Bartelsman et al. (2013). We extend their measure of capital misallocation and derive a frictionless benchmark in a general equilibrium framework. The importance of capital misallocation for business cycles is illustrated by Eisfeldt and Rampini (2006).

Our study also relates to the literature on production-based asset pricing, including Jermann (1998), Boldrin et al. (2001), and Kaltenbrunner and Lochstoer (2010), which aims to make the real business cycle model consistent with properties of aggregate asset prices. While these models feature a representative firm, we incorporate a continuum of firms. This allows us to pay close attention to cross-sectional aspects of the data, thereby providing a more realistic micro foundation for the sources of aggregate risk premia. While Kogan (2001) and Gomes et al. (2003) also model firm heterogeneity, our model provides a tighter link to firm fundamentals such that we estimate model parameters.

Our model mechanism is also related to the works of Gabaix (1999) and Luttmer (2007). Gabaix (1999) explains the power law of city sizes with random walks reflected at a lower bound. Using a similar mechanism, Luttmer (2007) generates a power law in firm size in a steady-state model. We extend this literature by studying the impact of a power law in firm size in a business cycle model with common idiosyncratic skewness shocks.

Starting with the influential paper by Berk et al. (1999), there exists a large literature, which studies the cross-section of returns in the neoclassical investment framework, e.g., Carlson et al. (2004), Zhang (2005), Cooper (2006), and Gomes and Schmid (2010). For tractability, these papers assume an exogenous pricing kernel and link firm cash flows and the pricing kernel directly via aggregate shocks. In contrast, we provide a micro foundation for the link between investment frictions and aggregate consumption.

2 Model

Time is discrete and infinite. The economy is populated by a unit mass of firms. Firms own capital, produce output with a neoclassical technology subject to investment being partially irreversible, and face permanent idiosyncratic and aggregate shocks. The representative household has recursive preferences and consumes aggregate dividends. This section elaborates on these model elements and defines the recursive competitive equilibrium of the economy.

2.1 Production

Firms produce output Y with the neoclassical technology

$$Y = (X\mathcal{E})^{1-\alpha} K^{\alpha}, \tag{1}$$

where X is aggregate productivity, $\mathcal{E}$ is idiosyncratic productivity, K is the firm's capital stock and $\alpha < 1$ is a parameter that reflects diminishing returns to scale. Aggregate productivity X follows a geometric random walk

$$X' = \exp\left\{g_x - \sigma_x^2/2 + \sigma_x \eta_x'\right\} X, \tag{2}$$

where g_x denotes the average growth rate of the economy, σ_x the volatility of log aggregate productivity growth, and η_x an i.i.d. standard normal innovation.

Idiosyncratic productivity growth is a mixture of a normal and a Poisson distribution, allowing for rare but large negative productivity draws. These negative jumps capture, for instance, sudden drops in demand, increases in competition, the exit of key human capital, or changes in regulation. As we will see, they are also essential for allowing the model to replicate the cross-sectional distribution of firms' sales growth. Specifically, idiosyncratic productivity $\mathcal{E}$ follows a geometric random walk modulated with idiosyncratic jumps

$$\mathcal{E}' = \exp\left\{g_\varepsilon - \sigma_\varepsilon^2/2 + \sigma_\varepsilon \eta_\varepsilon' + \chi' J' - \lambda \left(e^{\chi'} - 1\right)\right\} \mathcal{E}, \tag{3}$$

where g_ε denotes the average firm-specific growth rate, σ_ε the volatility of the normal innovations in firm-specific productivity, η an i.i.d. idiosyncratic standard normal shock, and J an i.i.d. idiosyncratic Poisson shock with constant intensity λ. The jump size χ varies with aggregate conditions η_x, which we capture with the exponential function

$$\chi(\eta_x) = -\chi_0 e^{-\chi_1 \eta_x} \tag{4}$$

with strictly positive coefficients χ_0 and χ_1. This specification implies that jumps are negative and larger in worse aggregate times, i.e., for low values of η_x.

Our specification for idiosyncratic productivity warrants a few comments. First, Bloom (2009) structurally estimates the cyclicality in the dispersion of idiosyncratic productivity, which is a symmetric measure of uncertainty. Our specification also leads to time variation in the higher moments of idiosyncratic productivity growth. In particular, equation (4) implies that firm-specific productivity shocks become more left skewed in recessions. Second, different from the uncertainty shocks in Bloom (2009) and Bloom et al. (2014), our assumptions imply that changes in idiosyncratic jump risk are neither known to firms ex ante nor persistent, and therefore do not cause wait-and-see effects. As we will show, however, they induce large changes in measured aggregate productivity via their effect on the efficiency of the cross-sectional capital distribution. Third, in contrast to the consumption-based asset pricing literature with disaster risk in consumption, for instance Barro (2006), Gabaix (2012), and

Wachter (2013), we do not model time variation in the jump probability λ. If the jump probability were increasing in recessions, it would induce rising skewness in productivity and sales growth, while in the data it is falling.[4] Fourth, the idiosyncratic jump risk term χJ is compensated by its mean $\lambda(e^{\chi} - 1)$, so that the cross-sectional mean of idiosyncratic productivity is constant (see equation (5) below). This normalization implies that aggregate productivity is determined solely by η_x-shocks, so that our model does not generate aggregate jumps in productivity as emphasized by, e.g., Gourio (2012). Because the size of the jump risk is common across firms, we refer to it as *common idiosyncratic skewness* in productivity.

Given the geometric growth in idiosyncratic productivity, the cross-sectional mean of idiosyncratic productivity is unbounded unless firms exit. We therefore assume that at the beginning of a period – before production takes place and investment decisions are made – each firm exits the economy with probability $\pi \in (0, 1)$. Exiting firms are replaced by an identical mass of entrants who draws their initial productivity level from a log-normal distribution with location parameter $g_0 - \sigma_0^2/2$ and scale parameter σ_0. Whenever firms exit, their capital stock scrapped and entrants start with zero initial capital.

Since the idiosyncratic productivity distribution is a mixture of Gaussian and Poisson innovations, it cannot be characterized by a known distribution.[5] But two features are noteworthy. First, due to random growth and exit, the idiosyncratic productivity distribution and thus firm size features a power law, as shown by Gabaix (2009). A power law holds when the upper tail of the firm size distribution obeys a Pareto distribution such that the probability of size S greater than x is proportional to $1/x^{\zeta}$ with tail (power law) coefficient ζ.[6]

Second, even though the distribution is unknown, we can compute its higher moments. Let $\mathbb{M}_n$ denote the n-th cross-sectional raw moment of the idiosyncratic productivity distribution

[4]Note that skewness of Poisson jumps J equals $\lambda^{-1/2}$.

[5]Dixit and Pindyck (1994) assume a similar process without Poisson jumps in continuous time and solve for the shape of the cross-sectional density numerically; see their chapter 8.4.

[6]In our model, the tail coefficient solves the nonlinear equation $1 = (1 - \pi)Z(\zeta)$, where $Z(\zeta) = \exp\{\zeta g_{\varepsilon} - \zeta\sigma_{\varepsilon}^2/2 + \zeta^2\sigma_{\varepsilon}^2/2 + \lambda(e^{\zeta\chi} - 1) - \zeta\lambda(e^{\chi} - 1)\}$.

$\mathcal{E}$. It has the following recursive structure

$$\mathbb{M}_n' = (1-\pi)\exp\{ng_\varepsilon - n\sigma_\varepsilon^2/2 + n^2\sigma_\varepsilon^2/2 + \lambda(e^{n\chi'}-1) - n\lambda(e^{\chi'}-1)\}\mathbb{M}_n + \pi\exp\{ng_0 - n\sigma_0^2/2 + n^2\sigma_0^2/2\}. \quad (5)$$

The integral over idiosyncratic productivity and capital determines aggregate output. To ensure that aggregate output is finite, we require that the productivity distribution has a finite mean.[7] Equation (5) states that the mean evolves according to $\mathbb{M}_1' = (1-\pi)e^{g_\varepsilon}\mathbb{M}_1 + \pi e^{g_0}$, which is finite if

$$g_\varepsilon < -\ln(1-\pi) \approx \pi. \quad (6)$$

In words, the firm-specific productivity growth rate has to be smaller than the exit rate. In this case, the first moment is constant and, for convenience, we normalize it to one by setting

$$g_0 = \ln(1 - e^{g_\varepsilon}(1-\pi)) - \ln(\pi). \quad (7)$$

2.2 Firms

To take advantage of higher productivity, firms make optimal investment decisions. Capital evolves according to

$$K' = (1-\delta)K + I, \quad (8)$$

where δ is the depreciation rate and I is investment. As in Khan and Thomas (2013) and Bloom et al. (2014), we assume investment is partially irreversible, which generates spikes and positive autocorrelation in investment rates as observed in firm level data. Quadratic adjustment costs can achieve the latter only at the expense of the former, since they imply an increasing marginal cost of adjustment. Partial irreversibility means that firms recover only a fraction ξ of the book value of capital when they choose to disinvest. These costs arise from resale losses due to transactions costs, asset specificity, and the physical costs of resale.

We show in Section 3 that partial irreversibility yields an (S, s) investment policy such that firms have nonzero investment only when their capital falls outside an (S, s) inactivity

[7] Luttmer (2007) makes a related assumption (Assumption 4), which states that "a firm is not expected to grow faster than the population growth rate" to ensure that the firm size distribution has finite mean.

band.[8] A firm with an unacceptably high capital stock relative to its current productivity will reduce its stock only to the upper bound of its inactivity range. Similarly, a firm with too little capital invests only to the lower bound of its inactivity range to reduce the linear penalty it will incur if it later chooses to shed capital. Thus, partial irreversibility can deliver persistence in firms investment rates by encouraging repeated small investments at the edges of inactivity bands.

We summarize the distribution of firms over the idiosyncratic states $(K, \mathcal{E})$ using the probability measure μ and note that the aggregate state of the economy is given by (X, μ). The distribution of firms evolves according to a mapping Γ, which we derive in Section 3. Intuitively, the dynamics of μ are shaped by the exogenous dynamics of $\mathcal{E}$ and X, the endogenous dynamics of K resulting from firms' investment decisions, and firm entry and exit.

Firms maximize the present value of their dividend payments to shareholders by solving

$$V(K, \mathcal{E}, X, \mu) = \max_{I} \left\{ D + (1-\pi)\mathbb{E}\big[M' V(K', \mathcal{E}', X', \mu')\big] \right\}, \tag{9}$$

where

$$D = Y - I \; 1_{\{I \geq 0\}} - \xi I \; 1_{\{I<0\}} \tag{10}$$

denotes the firm's dividends and M is the equilibrium pricing kernel based on aggregate consumption and the household's preferences, which we derive in Section 3.1.

2.3 Household

The representative household of the economy maximizes recursive utility U over consumption C as in Epstein and Zin (1989):

$$U(X, \mu) = \max_{C} \left\{ (1-\beta) C^{1-\frac{1}{\psi}} + \beta \left(\mathbb{E}\left[U(X', \mu')^{1-\gamma} \right] \right)^{(1-\frac{1}{\psi})/(1-\gamma)} \right\}^{1/(1-\frac{1}{\psi})} \tag{11}$$

where $\psi > 0$ denotes the elasticity of intertemporal substitution (EIS), $\beta \in (0, 1)$ the subjective discount factor, and $\gamma > 0$ the coefficient of relative risk aversion. In the special case when risk aversion equals the inverse of EIS, the preferences reduce to the common power

[8]See Andrew B. Abel (1996) for a continuous time model of partial irreversibility.

utility specification. The household's resource constraint is

$$C = \int Y \, \mathrm{d}\mu - \int I \times 1_{\{I>0\}} \, \mathrm{d}\mu - \xi \int I \times 1_{\{I<0\}} \, \mathrm{d}\mu + \frac{\pi\xi}{1-\pi} \int K \, \mathrm{d}\mu, \tag{12}$$

where the last term captures the liquidating dividends of exiting firms.[9]

2.4 Equilibrium

A recursive competitive equilibrium for this economy is a set of functions (C, U, V, K, Γ) such that:

(i) Firm optimality: Taking M and Γ as given, firms maximize firm value (9) with policy function K subject to (8) and (10).

(ii) Household optimality: Taking V as given, household maximize utility (11) subject to (12) with policy function C.

(iii) The good market clears according to (12).

(iv) Model consistency: The transition function Γ is induced by K, aggregate productivity X, equation (2), idiosyncratic productivity $\mathcal{E}$, equation (3), and entry and exit.

3 Analysis

In this section, we characterize firms' optimal investment policy and the transition dynamics of the cross-sectional distribution of firms. We also derive closed-form solutions for a frictionless version of the model, which serves as a benchmark for quantifying the degree of capital misallocation and the wedge between actual and measured aggregate productivity. Because aggregate productivity contains a unit root, we solve the model in detrended units, such that detrended consumption c and wealth w are given by

$$c = C/X \qquad w = W/X.$$

[9] To understand this term, note that exiting firms are not contained in the current μ. Since entrants do not own capital, the aggregate capital stock at the end of the previous period (before exit shocks materialized) was $\frac{1}{1-\pi}\int K \, \mathrm{d}\mu$. Because exit shocks are equally likely to hit any firm, the capital of exiting firms equals $\frac{\pi}{1-\pi}\int K \, \mathrm{d}\mu$, so that the resale value of this capital equals $\xi \frac{\pi}{1-\pi}\int K \, \mathrm{d}\mu$.

3.1 Household Optimization

The household's first order condition with respect to the optimal asset allocation implies the usual Euler equation

$$\mathbb{E}\big[M'R'\big] = 1 \tag{13}$$

where M' is the pricing kernel and R' is the return on equity, defined by $V'/(V-D)$. The pricing kernel is given by

$$M' = \beta^{\theta}(x')^{-\gamma}\left(\frac{c'}{c}\right)^{-\theta/\psi}\left(\frac{w'}{w-c}\right)^{\theta-1}, \tag{14}$$

where $\theta = \frac{1-\gamma}{1-1/\psi}$ is a preference parameter and $x' = X'/X$ is i.i.d. log-normal distributed. In the case of power utility, θ equals one and wealth drops out of the pricing kernel. With Epstein-Zin preferences, the dynamics of both consumption and wealth evolve endogenously and are part of the equilibrium solution.

Consistent with the Euler equation (13), wealth is defined recursively as the present value of future aggregate consumption:

$$w = c + \beta\mathbb{E}\left[(x')^{1-\gamma}(w')^{\theta}\left(\frac{c'}{c}\right)^{-\theta/\psi}\right]^{1/\theta}. \tag{15}$$

Firm exit introduces a wedge between wealth and the aggregate market value of firms. This stems from the fact that wealth captures the present value of both incumbents and entrants, whereas aggregate firm value relates to the present value of dividends of incumbent firms only.

3.2 Firm Optimization

Having solved for the functional form of the pricing kernel, we can characterize firms' optimal investment policy. The homogeneity of the value function and the linearity of the constraints imply that we can detrend the firm problem by the product of both permanent shocks $X\mathcal{E}$, as for instance in Bloom (2009). We define the firm-specific capital to productivity ratio $\kappa = K/(X\mathcal{E})$, the capital target to productivity ratio $\tau = K'/(X\mathcal{E})$, and the firm value to productivity ratio $v = V/(X\mathcal{E})$.

Given the linear cost structure, one can divide the value function into three regions. In the investing region $((1-\delta)\kappa \leq \tau)$, firms increase their capital to productivity ratio and the optimal firm value solves v_u; in the disinvesting region ($\tau \leq (1-\delta)\kappa$), firms decrease their capital to productivity ratio and the optimal firm value solves v_d; otherwise, firms are inactive. Firm value v is thus the maximum of the value of investing v_u, disinvesting v_d, or inactivity:

$$v_u(\kappa,\mu) = \max_{(1-\delta)\kappa\leq\tau} \left\{\kappa^\alpha - (\tau-(1-\delta)\kappa) + (1-\pi)\mathbb{E}\Big[M'x'\varepsilon' v\big(\kappa',\mu'\big)\Big]\right\}, \quad (16)$$

$$v_d(\kappa,\mu) = \max_{\tau\leq(1-\delta)\kappa} \left\{\kappa^\alpha - \xi(\tau-(1-\delta)\kappa) + (1-\pi)\mathbb{E}\Big[M'x'\varepsilon' v\big(\kappa',\mu'\big)\Big]\right\}, \quad (17)$$

$$v(\kappa,\mu) = \max\left\{v_u(\kappa,\mu),\ v_d(\kappa,\mu),\ \kappa^\alpha + (1-\pi)\mathbb{E}\Big[M'x'\varepsilon' v\big((1-\delta)\kappa/(x'\varepsilon'),\mu'\big)\Big]\right\}, \quad (18)$$

where $\varepsilon' = \mathcal{E}'/\mathcal{E}$. Because both growth rates ε' and x' are i.i.d., the state space of the detrended firm problem reduces to (κ,μ). Importantly, for adjusting firms next period's capital to productivity ratio $\kappa' = \tau/(x'\varepsilon')$ is independent of the current capital to productivity ratio. This fact implies that firms share a common time-varying capital target τ, which is independent of their own characteristic κ. The optimal capital targets for the investing and disinvesting regions is given by $\mathcal{T}_u(\mu)$ and $\mathcal{T}_d(\mu)$, respectively, and solves

$$\mathcal{T}_u(\mu) = \arg\max_\tau \left\{-\tau + (1-\pi)\mathbb{E}\Big[M'x'\varepsilon' v\big(\tau/(x'\varepsilon'),\mu'\big)\Big]\right\},$$

$$\mathcal{T}_d(\mu) = \arg\max_\tau \left\{-\xi\tau + (1-\pi)\mathbb{E}\Big[M'x'\varepsilon' v\big(\tau/(x'\varepsilon'),\mu'\big)\Big]\right\}.$$

Given these capital targets, the optimal policy of the firm-specific capital to productivity ratio can be characterized by an (S,s) policy and is given by

$$\kappa' = \max\big\{\mathcal{T}_u(\mu), \min\{\mathcal{T}_d(\mu),(1-\delta)\kappa\}\big\}/(x'\varepsilon') \quad (19)$$

where the max operator characterizes the investing region and the min operator the disinvesting one. Conditional on adjusting, the capital to productivity ratio of every firm is either $\mathcal{T}_u$ or $\mathcal{T}_d$, independent of their own characteristic κ but dependent on the aggregate firm distribution μ.

The optimal investment rate policy, implied by (19), can be summarized by the same three regions of investment, inactivity, and disinvestment:

$$\frac{I}{K} = \begin{cases} \frac{\mathcal{T}_u(\mu)-\kappa}{\kappa} + \delta & (1-\delta)\kappa < \mathcal{T}_u \quad \text{investing,} \\ 0 & \mathcal{T}_u \leq (1-\delta)\kappa \leq \mathcal{T}_d \quad \text{inactive,} \\ \frac{\mathcal{T}_d(\mu)-\kappa}{\kappa} + \delta & \mathcal{T}_d < (1-\delta)\kappa \quad \text{disinvesting.} \end{cases}$$

In Figure 1, we plot both the optimal capital to productivity and investment rate policies for two arbitrary capital targets. Intuitively, when a firm receives a positive idiosyncratic productivity draw, its capital to productivity ratio κ falls. If the shock is large enough and depreciated κ is less than $\mathcal{T}_u$, it will choose a positive investment rate, which reflects the relative difference between target and current capital to productivity ratio as well as the depreciation rate. As a result, next period's capital to productivity ratio will reach $\mathcal{T}_u$ in the investment region.

When a firm experiences an adverse idiosyncratic productivity draw, its capital to productivity ratio κ increases and it owns excess capital. If the shock is severe enough and depreciated κ is greater than $\mathcal{T}_d$, it will choose a negative investment rate, which reflects the relative difference between target and current capital to productivity ratio as well as the depreciation rate. As a result, next period's capital to productivity ratio will fall to $\mathcal{T}_d$ in the disinvestment region. For small enough innovations, the depreciated capital to productivity ratio remains within $\mathcal{T}_u$ and $\mathcal{T}_d$. In this region, firms are inactive and have a zero investment rate.

An important features of our model is that there is heterogeneity in the duration of disinvestment constraintness. This feature arises because adverse idiosyncratic productivity shocks can arise either from a normal distribution or from a Poisson distribution. While adverse normal distributed shocks are short lasting, Poisson shocks are rare and large and therefore long lasting. As a result of Poisson shocks, the capital to productivity ratio rises dramatically, indicating a long duration of disinvestment constraintness.

3.3 Aggregation

In the previous section, we showed that the firm-specific state space of the firm's problem reduces to the capital-to-productivity ratio κ. In contrast, the univariate distribution of firms over κ is not sufficient to determine aggregate quantities in the model because output in equation (1) cannot be expressed in terms of κ as the single idiosyncratic state. To derive aggregates, we thus define idiosyncratic variables that are detrended by aggregate productivity only, which we denote by the corresponding lower case letters:

$$k \equiv \frac{K}{X}, \tag{20}$$

and similarly for Y, I, and D. The transition dynamics for detrended capital follow by multiplying the transition of the capital-to-productivity ratio in (19) by $\mathcal{E}'$,

$$k' = \max\big\{\mathcal{E}\mathcal{T}_u(\mu), \min\{\mathcal{E}\mathcal{T}_d(\mu), (1-\delta)k\big\}\big\}/x', \tag{21}$$

Note that, due to detrending, k' is not contained in the current period's information set. Detrended investment follows by dividing the capital accumulation equation (8) by X and substituting $\frac{K'}{X} = k'x'$:

$$i = \max\{\mathcal{E}\mathcal{T}_u(c) - (1-\delta)k, 0\} + \min\{\mathcal{E}\mathcal{T}_d(c) - (1-\delta)k, 0\}. \tag{22}$$

We summarize the distribution of firms over the detrended idiosyncratic states $(k, \mathcal{E})$ using the probability measure μ, which is defined on the Borel algebra $\mathcal{S}$ for the product space $\mathbf{S} = \mathbb{R}_0^+ \times \mathbb{R}^+$.[10] Using this measure, detrended aggregate quantities can be obtained by integrating over the respective firm-level variables,

$$\bar{k} = \int k \, \mathrm{d}\mu, \tag{23}$$

and similarly for Y, I, and D, so that the detrended aggregate resource constraint reads

$$c = \bar{y} - \int i \mathbf{1}_{\{i>0\}} \, \mathrm{d}\mu - \xi \int i \mathbf{1}_{\{i<0\}} \, \mathrm{d}\mu + \frac{\pi\xi}{1-\pi}\bar{k}. \tag{24}$$

The measure μ evolves over time according to the mapping $\Gamma : (\mu, \eta_x') \mapsto \mu'$, which results from the dynamics of idiosyncratic productivity $\mathcal{E}$ in equations (3) and (4), and the transition law for firms' detrended capital in equation (21). To derive this mapping, note that k' is predetermined with respect to the firm-level productivity shocks (η_ε', J'). This implies that, conditional on current information and next period's aggregate shock η_x', next period's characteristics $(k', \mathcal{E}')$ are cross-sectionally independent of one another. Therefore, for any $(\mathbf{K}, \mathbf{E}) \in \mathbf{S}$,

$$\mu'(\mathbf{K}, \mathbf{E}|\eta_x') = \mu_k'(\mathbf{K}|\eta_x') \times \mu_{\mathcal{E}}'(\mathbf{E}|\eta_x'), \tag{25}$$

[10] Note that, with a slight abuse of notation but for better readability, we continue to use the symbols μ and Γ to denote the distribution of firms and its transition in the detrended economy.

where μ_k and $\mu_{\mathcal{E}}$ are the marginal distributions of capital and productivity, respectively. The measure of firms with a capital stock of $k' \in \mathbf{K}$ next period is simply the integral over the measure of firms who choose k' as their optimal policy this period and survive, plus the mass of entrants in the case $0 \in \mathbf{K}$:

$$\mu_k'(\mathbf{K}|\eta_x') = (1-\pi) \int \mathbf{1}_{\{k' \in \mathbf{K}\}} \, \mathrm{d}\mu + \pi \mathbf{1}_{\{0 \in \mathbf{K}\}} \tag{26}$$

The measure of firms with an idiosyncratic productivity of $\mathcal{E}' \in \mathbf{E}$ next period follows from the fact that, conditional on $(\mathcal{E}, J', \eta_x')$, $\mathcal{E}'$ is log-normally distributed for both continuing firms and new entrants. The distribution of $\mathcal{E}'$ conditional on η_x' can therefore be computed as

$$\begin{aligned} \mu_{\mathcal{E}}'(\mathbf{E}|\eta_x') \quad = \quad & \int_{\mathcal{E}' \in \mathbf{E}} \Bigg\{ (1-\pi) \int \sum_{j=0}^{\infty} p_j \phi \left(\frac{\ln(\mathcal{E}') - \left(\ln(\mathcal{E}) + g_\varepsilon - \frac{\sigma_\varepsilon^2}{2} + \chi' j - \lambda \left(e^{\chi'} - 1 \right) \right)}{\sigma_\varepsilon} \right) \mathrm{d}\mu_{\mathcal{E}} \\ & + \pi \phi \left(\frac{\ln(\mathcal{E}') - \left(g_0 - \sigma_\varepsilon^2/2 \right)}{\sigma_\varepsilon} \right) \Bigg\} d\mathcal{E}' \end{aligned} \tag{27}$$

where $p_j = \lambda^j e^{-\lambda}/j!$ is the Poisson probability of receiving j jumps and ϕ the standard normal density. Equations (25)–(27) define the transition function Γ.

3.4 Efficiency of the Cross-Sectional Allocation

We are interested in the extend to which aggregate output and consumption dynamics are determined by time-variation in the efficiency of the capital allocation across firms. A natural benchmark for quantifying this efficiency is an allocation that maximizes aggregate output by equating the marginal products of capital across firms, as suggested by Hsieh and Klenow (2009). We will refer to this allocation as the frictionless (FL) benchmark. The marginal product in our model equals $\alpha\kappa^{\alpha-1}$, and it is equalized across firms when firms' capital stocks are proportional to their idiosyncratic productivities, $k_{FL} = \bar{k}\mathcal{E}$. This results in an aggregate output of

$$\bar{y}_{FL} = \int \mathcal{E}^{1-\alpha} (\bar{k}\mathcal{E})^\alpha \, \mathrm{d}\mu = \bar{k}^\alpha. \tag{28}$$

Following Hsieh and Klenow, we quantify the efficiency of the cross-sectional allocation with the output gap, defined as

$$\mathcal{G}_Y(\mu, \eta_x) = \frac{\bar{y}}{\bar{y}_{FL}} = \frac{\int \mathcal{E}^{\alpha-1} k^\alpha \, \mathrm{d}\mu}{\bar{k}^\alpha}. \tag{29}$$

Hsieh and Klenow interpret the output gap as a measure of capital misallocation, but in our model it arises as a dynamically optimal outcome from the interplay of idiosyncratic risk and investment frictions.[11] Three channels play a role.

First, the common assumption that capital stocks are predetermined implies that there is always a contemporaneous mismatch between firms' productivity levels and the capital stocks that would equate their marginal products. The severity of this mismatch increases in the dispersion and asymmetry of idiosyncratic shocks. Similarly, time-variation in these higher moments results in time-variation in the allocative efficiency.

Second, existing mismatches between capital and productivity carry over to future periods for firms that do not adjust their capital stocks. Inactivity therefore reduces the average allocative efficiency. It also creates persistence when aggregate shocks affect the higher moments of idiosyncratic risk. The reason is that firms that experience more extreme idiosyncratic shocks move further away from their adjustment triggers, therefore becoming inactive for longer periods. Episodes of volatile or asymmetric idiosyncratic shocks therefore lower the extensive margin for multiple periods into the future. As a consequence, the allocative efficiency becomes a function of the recent history of aggregate shocks.

Third, the assumption that new firms enter the economy with positive productivity but with zero capital stocks reduces the average allocative efficiency because these firms have infinite marginal products.

To separate the effect of these channels, we decompose the output gap into a part due to adjustment costs and a part due to predetermined capital stocks and exit as

$$\mathcal{G}_Y(\mu, \eta_x) = \frac{\bar{y}}{\bar{y}_{NC}} \times \frac{\bar{y}_{NC}}{\bar{y}_{FL}}. \tag{30}$$

Here, $\bar{y}_{NC}$ denotes the level of output that would be feasible to produce based on the current aggregate capital stock in an economy without adjustment costs but with predetermined capital stocks and exit. We will refer to this allocation as the no cost (NC) benchmark. The term $\frac{\bar{y}}{\bar{y}_{NC}}$ in equation (30) isolates the part of the output gap due to adjustment costs. In the

[11] The fact that static measures of misallocation do not necessarily reflect true misallocation in a dynamic framework was first pointed out by **?**

appendix, we characterize the solution to the NC benchmark and show that

$$\bar{y}_{NC} = A(\eta_x) \times (1-\alpha)^{1-\alpha}\bar{k}^{\alpha}, \tag{31}$$

where $A(\eta_x) = \mathbb{E}[\varepsilon^{1-\alpha}|\eta_x]$. Combing this with (28) shows that the second factor in the output gap equals $\frac{\bar{y}_{NC}}{\bar{y}_{FL}} = A(\eta_x) \times (1-\pi)^{1-\alpha}$, where A captures the part of the output gap due to predetermined capital stocks and $(1-\pi)^{1-\alpha}$ the part due to exit.

Because we measure the efficiency of the cross-sectional allocation based on a fully specified general equilibrium model, we are also able to quantify the associated welfare costs. We quantify the welfare losses due to adjustment costs with the welfare gap

$$\mathcal{G}_U(\mu, \eta_x) = \frac{U^*}{U^*_{NC}}, \tag{32}$$

where U^* denotes social welfare in the full model and U^*_{NC} denotes social welfare in the NC benchmark. Note that our definition of the output gap includes the parts due to predetermined capital stocks and exit, whereas the welfare gap measures the effect due to adjustment costs only.

3.5 Numerical Method

As in Krusell and Smith (1998), we approximate the firm-level distribution μ with a finite-dimensional aggregate state variable to make the model solution computable. However, instead of relying of cross-sectional moments of capital as most of the previous literature, we use detrended aggregate consumption c. For two reasons, this approach is better suited for models with significant time-variation in efficiency of the cross-sectional allocation. First, consumption captures the joint distribution of capital and productivity, whereas aggregate capital (and higher moments of capital) only capture the marginal distribution of capital. Second, using consumption as a state variable eliminates the need for a second approximation rule that maps capital into marginal utilities. Below, we discuss each of these points in more detail and summarize our numerical approach.

To illustrate the importance of capturing both dimensions of μ, consider the stylized example in Table 1, where both idiosyncratic productivity and capital can only take on two

Table 1: Two stylized firm-level distributions

Case I

k_{high}	0	0.5
k_{low}	0.5	0
	ε_{low}	ε_{high}

Case II

k_{high}	0.5	0
k_{low}	0	0.5
	ε_{low}	ε_{high}

values. The table entries are the probability mass for each point in the support of μ. We assume that the aggregate shock is identical in both scenarios. Case I shows an efficient allocation, where productive firms hold a high capital stock, unproductive firms hold a low capital stock, and aggregate output is high. Case II shows an inefficient allocation that results in low aggregate output. Importantly, the marginal distribution of capital is identical in both cases and aggregate capital stock equals $(k_{low} + k_{high})/2$. Because the Krussel and Smith algorithm predicts next period's capital stock solely based on today's marginal capital distribution (and the current aggregate shock), it incorrectly predicts the same value in both cases. In contrast, being a policy, consumption reflects both dimensions of the idiocyncratic state space and can therefore distinguish between the two cases.

The second issue related to the Krussel and Smith algorithm arises only when it is applied to models with firm heterogeneity, as in Khan and Thomas (2008, 2013) and Bloom et al. (2014). Because the decentralized firm problem involves the pricing kernel, it is necessary to compute the representative agent's marginal utility in order to solve the decentralized firm problem. When μ is approximated with $\bar{k}$, one has to introduce a second, contemporaneous approximation that maps $\bar{k}$ into marginal utility. For example, Khan and Thomas (2008) specify $u'(c)$ as a log-linear function of $\bar{k}$. When misallocation becomes quantitatively important, this approximation is poor because c is in general a function of both dimensions of μ, whereas $\bar{k}$ only reflects one marginal distribution.[12] In contrast, specifying c as an aggregate state variable implies that no additional approximation is required.

Methodologically, the main difference between aggregate capital compared to consumption as state variable arises when specifying their law of motions. Tomorrow's aggregate capital

[12]For example, in Bloom et al. (2014), the mapping to marginal utility results in R^2's as low as 88% for some states – see their Table B1.

stock is contained in the current information, so that it is possible to approximate the law of motion for aggregate capital with a deterministic function. In contrast, tomorrow's consumption depends on tomorrow's realization of the aggregate shock η'_x, and we with the stochastic log-linear rule

$$\ln(c') = \varphi_0(\eta'_x) + \varphi_1(\eta'_x)\ln(c). \tag{33}$$

These forecasting functions imply intercepts and slope coefficients that depend on the future shock to aggregate productivity, i.e., they yield forecasts conditional on η'_x.

In a model based on a representative household with time-separable utility, the consumption rule (33) is sufficient to close the model. Because we model a representative household with recursive utility, we also have to solve for the wealth dynamics to be able to compute the pricing kernel (14). In the absence of arbitrage opportunities, the Euler equation for the return on wealth (13) implies a strict consistency requirement between the dynamics of consumption and wealth. In particular, wealth has to equal the present value of future consumption. To impose this requirement, we define wealth as a nonparametric function of current consumption, $w(c)$, that we determine by iterating on the Euler equation (15). To do so, we specify a fine grid for current consumption, impose the dynamics specified in (33), and use cubic splines to evaluate $w(c)$ on off grid values. In contrast to the algorithm used by Khan and Thomas (2008) and many subsequent papers, our model solution therefore does not allow for dynamic inconsistencies in the form of arbitrage opportunities.

To summarize, our algorithm works as follows. Starting with a guess for the coefficients of the equilibrium consumption rule (33), we first solve for the wealth rule and then the firm's problem (16)–(18) by value function iteration. To update the coefficients in the equilibrium rule (33), we simulate a continuum of firms. Following Khan and Thomas (2008), we impose market clearing in the simulation, meaning that firm policies have to satisfy the aggregate resource constraint (24). The simulation allows us to update the consumption dynamics and we iterate on the procedure until the consumption dynamics have converged.

4 Estimation

The main goal of our paper is to relate aggregate fluctuations and risk premia to time variation in the efficiency of factor allocations at the firm level. Because such variation results from the interplay of idiosyncratic risk and frictions, it is crucial for our model to capture the cyclicality in the shocks that individual firms face. We therefore estimate productivity parameters based on a set of moments that reflects both the shape and cyclicality of the cross-sectional distribution. In particular, our simulated method of moments (SMM) estimation targets the cross-sectional distribution of firms' sales growth and investment rates, along with a set of aggregate quantity moments. Our paper is the first to estimate a general equilibrium model with substantial heterogeneity based on such a set of endogenous moments. This estimation is made feasible largely due to modeling shocks as permanent, which allows us to reduce the dimensionality of the state space relative to earlier studies such as Khan and Thomas (2008), Bachmann and Bayer (2014), or Bloom et al. (2014).

4.1 Data

Our estimation relies on both aggregate and firm-level data over the period from 1976 to 2015. We use quarterly time series but rely on overlapping 4-quarter moments. This allows us to make use of the higher information content of quarterly relative to annual data, while avoiding the seasonal variation of quarterly moments.

We define aggregate output as gross value added of the non-financial corporate sector, aggregate investment as private nonresidential fixed investment, and aggregate consumption as the difference between the two. All series are per capita, deflated with their respective price indices, and taken from NIPA. Moments are based on 4-quarter log growth rates.

In addition to aggregate moments, our estimation uses cross-sectional moments of firms' sales growth and investment rates to identify the parameters associated with idiosyncratic productivity and adjustment costs. Firm-level data is taken from the merged CRSP-Compustat database. All firm-level variables are converted to per capita units to make them consistent with the aggregate series. We eliminate firms in the finance, insurance, and real estate sectors

(NAICS sectors 52 and 53) because their balance sheets differ substantially from those of other firms. We also eliminate utilities (NAICS sector 22) because government regulation of this industry implies that the profit maximization assumption is likely not to hold. Additionally, we only consider firms with at least 10 years of data to ensure that time-variation in cross-sectional statistics is mostly driven by shocks to existing firms as opposed to changes in the composition of firms. In reality, such changes are driven by firms' endogenous entry and exit decisions, a channel that is outside of our model.

Sales growth is defined as the four quarter change in log SALEQ, deflated by the implicit price deflator for GDP. The investment rate is defined as the sum of four quarterly investment observations divided by the beginning capital stock. We compute quarterly investment as the change in gross property, plant and equipment (PPEGTQ), deflated by the implicit price deflator for private fixed investment in the corresponding NAICS sector, subsector, or industry group (henceforth "NAICS group").[13,14] Firms' capital stocks are computed via a perpetual inventory method: $K_{i,t} = (1 - \delta_{i,t})K_{i,t-1} + I_{i,t}$. To account for investment good- and period-specific differences in depreciation rates, we rely on the BEA's estimates for each year and each NAICS group.[15] The recursion is initialized at net property, plant and equipment (PPENTQ), deflated by the price index for the corresponding NAICS group, and multiplied by a subsector (3 digit NAICS) specific constant ϕ. As in Bachmann and Bayer (2014), the constant corrects deflated PPENTQ for the fact that it tends to underestimate economic capital because (i)

[13]Investment good deflators for individual NAICS codes are inferred from the BEA's fixed asset tables. In particular, the difference between the log growth rate in investment (from Table 3.7ESI) and the log growth rate in the chained quantity index for investment (from Table 3.8ESI) equals an estimate of the inflation rate. To check the accuracy of this calculation, we compute a capital-weighted average between the sector-specific inflation rates (excluding sectors 22, 52, and 53), and determine its correlation with the aggregate inflation time series for nonresidential private fixed investment that is reported by the BEA. This correlation equals 92.8%, confirming the validity of our approach. Lastly, to map the annual BEA data to the quarterly frequency, we assume the inflation rate is constant throughout each calendar year.

[14]The information in the BEA's fixed asset tables is provided in heterogeneous levels of granularity. Most categories are reported at the subsector level (3 digits). Others (23 Construction, 42 wholesale trade, 44-45 retail trade, 55 management of companies and enterprises, 61 educational services, and 81 other services except government) are reported at the sector level (2 digits) only. Lastly, a few categories (5411 legal services, 5415 computer systems design and related services) are reported at the industry group level (4 digits) or as sets of several industry groups. Overall, we make use of estimates for 55 distinct groups of firms.

[15]The BEA's estimation procedure is described in detail in the note "BEA Depreciation estimates", available at www.bea.gov/national/FA2004/Tablecandtext.pdf. While the BEA's depreciation rate estimates are not directly available as a dataset, they can be inferred by dividing current cost depreciation (Table 3.4ESI) by the current cost net stock of private fixed assets (Table 3.1ESI).

accounting depreciation is motivated by tax incentives and typically overestimates economic depreciation and (ii) accounting capital stocks are reported at historical prices. We determine the constant such that the ratio $\frac{K}{\phi \times PPENTQ^*}$ equals one on average across the firm/quarter observations within each NAICS subsector, where $PPENTQ^*$ denotes the deflated value.[16] The initial value used in the perpetual inventory method, $\phi \times PPENTQ^*$, therefore equals an unbiased estimate of firm's economic capital stock.

Lastly, we delete all firm/quarter observations for which either the four quarter sales growth rate or the four quarter investment rate is missing. The cross-sectional dimension of our final sample includes 658 firms in Q1-1977, peaks at 1,586 firms in Q2-1997, and then drops back to 660 firms by Q4-2015.

4.2 Cyclical Properties of the Cross-Section of Firms

In this section, we document how the cross-section of firms moves over the business cycle, and we discuss implication of the associated empirical facts. Figure 3 shows the evolution of the cross-sectional distributions of firms' sales growth (left column) and investment rates (right column) over time. We summarize both distributions with robust versions of their first three moments, i.e., we measure centrality with the median, dispersion with the inter quartile range (IQR), and asymmetry with Kelly skewness.[17] The two top panels of the figure show that recessions are characterized by sizable declines in sales growth and investment rates for the median firm. This observation is unsurprising. However, recessions are further characterized by pronounced changes in the *shape* of the cross-sectional distributions.

Sales growth becomes more disperse during recessions and its skewness switches sign from positive to negative. This evidence suggests that recessions coincide with an increase in idiosyncratic risk. Bloom (2009) and Bloom et al. (2014) provide ample additional evidence for the increase in dispersion and model it as an increase in the volatility of firms' Gaussian productivity shocks. However, the pronounced change in the skewness of sales growth shows

[16] The average adjustment constant across subsectors equals 1.0355, which implies that PPENTQ underestimates the true capital stock by about 3.55% on average.

[17] Kelly skewness is defined as KSK= $\frac{(p_{90}-p_{50})-(p_{50}-p_{10})}{p_{90}-p_{10}}$, where p_x denotes the x-th percentile of the distribution. It measures asymmetry in the center of the distribution as opposed to skewness that can result from tail observations. Similar to the median and IQR, Kelly skewness is thus robust to outliers.

that the countercyclicality of idiosyncratic risk is better described as resulting from an expansion of the left tail of the shock distribution as opposed to a symmetric widening of the whole distribution. Intuitively, recessions are times where a subset of firms receives very negative shocks, but it is not the case that an equal proportion of firms receives very positive shocks.

Another characteristic of recessions is the fact – first documented by Bachmann and Bayer (2014) – that the dispersion in firms' investment rates *declines*. This procyclicality is suggestive of nonconvexities in firms' capital adjustment cost because in the absence of such frictions, the increase in the dispersion of firms' productivity would lead to a *larger* dispersion in investment rates.

Bachmann and Bayer (2014) argue that this fact is also informative about the cyclicality of idiosyncratic risk. Intuitively, when Gaussian volatility increases during recessions, the subset of firms receiving large positive shocks will undertake large positive investments, which leads to an increase in the cross-sectional dispersion of investment rates. When changes in uncertainty are large enough, this effect dominates the real options effect that causes firms to delay their investments in the face of increased uncertainty, which all else equal reduces the dispersion in investment rates. On the other hand, when uncertainty shocks are more moderately sized, the model becomes consistent with the procyclical investment rate dispersion, but uncertainty shocks no longer induce business cycles. Note, however, that this conclusion relies on a model with (a) real option effects and (2) time-variation in idiosyncratic risk that results from a *symmetric* change in the dispersion of shocks. Neither of these features are present in our model.

4.3 Simulated Method of Moments

This section explains how we estimate the model parameters. The full set of model parameters includes preference (β, γ, ψ), technology (δ, α), entry and exit (π, σ_0), and productivity $\theta = (\chi_0, \chi_1, \lambda, g_\varepsilon, \sigma_\varepsilon, g_x, \sigma_x)$ parameters. Since it is not feasible computationally to estimate the full set of parameters, we focus on estimating the vector of productivity parameters θ.

Values for the remaining parameters are taken from previous literature and are summarized in Panel A of Table 3. Following Bansal and Yaron (2004), we assume that the representative agent is fairly risk averse, $\gamma = 10$, and has a large EIS, $\psi = 2$. The time discount rate of $\beta = 0.995$ is chosen to achieve a low average risk-free rate. Capital depreciates at a rate of 2.5% and the curvature of the production function equals 0.65, similar to Cooper and Haltiwanger (2006). Firms exit the economy with a rate of 2%, similar to the value reported in Dunne et al. (1988). The productivity draws of entrants has a mean pinned down by condition (7) and a volatility of 10%. As estimated by Bloom (2009), we assume partial irreversibility costs of $\xi = 0.7$.

Productivity parameters are estimated with SMM, which minimizes a distance metric between key moments from actual data, Ψ^D, and moments from simulated model data, $\Psi^M(\theta)$. Given an arbitrary parameter vector θ, the model is solved numerically as described in Section 3.5. In solving the model, we use an equilibrium simulation of length $3040 + 600$ quarters, which equals twenty times the time dimension of the actual data plus an initial 600 periods that we discard to eliminate dependence on initial conditions. We then fix the equilibrium path of consumption that results from the equilibrium simulation and simulate a finite panel of firms for the same path of the economy.[18] Based on the simulated data panel, we calculate the model moments $\Psi^M(\theta)$ as well as the objective function $[\Psi^D - \Psi^M(\theta)]'W[\Psi^D - \Psi^M(\theta)]$. The parameter estimate $\hat{\theta}$ is found by searching globally over the parameter space. We use an identity weighting matrix and implement the global minimization via a genetic algorithm. Computing standard errors for the parameter estimate requires the Jacobian of the moment vector, which we find numerically via a finite difference method.

4.4 Moment Selection and Parameter Identification

To identify the parameter vector θ, we rely on a combination of aggregate and cross-sectional moments. First, we include the time series means of the six cross-sectional moments depicted in Figure 3: median, IQR, and Kelly skewness of sales growth and investment rates. Doing

[18]While the simulation step of the solution algorithm is based on a continuum of firms that are tracked on a bivariate histogram, this approach is not feasible for determining cross-sectional moments that span multiple quarters. The reason is that multi-period transition functions become too high dimensional to be computationally manageable. However, the Monte Carlo sample can be interpreted as a subsample of the continuum of firms because it is based on the same path for aggregate shocks and aggregate consumption as the model solution. We choose the number of simulated firms high enough to ensure that the simulated cross-sectional moments are not significantly affected by Monte Carlo noise.

so ensures that we capture the average shape of the conditional distributions of sales growth and investment rates, and therefore also the shape of their long run distributions. Second, we include the time series standard deviations of the same six cross-sectional moments to capture the amount of time-variation in the conditional cross-sectional distributions. Third, we rely on time series correlations between three cross-sectional moments and aggregate output growth to capture the cyclicality of the cross-section. In particular, we include the cyclicality of the dispersion in sales growth documented by Bloom (2009), the cyclicality of the skewness in sales growth documented by Salgado et al. (2015), and the cyclicality in the dispersion of investment rates documented by Bachmann and Bayer (2014). Relying on three measures of cyclicality jointly ensures that we capture various aspects of how the cross section co-moves with the cycle.

At the aggregate level, we include the mean growth rate of aggregate output, and the standard deviations of aggregate output, consumption, and investment to ensure that productivity parameters reflect not only the cross-section but also remain consistent with macro aggregates. In total, we estimate 7 productivity parameters based on the 19 moments shown in the data column of Table 4. In what follows, we discuss the main sources of identification for each estimated parameter.

The drift and volatility parameters of aggregate productivity, g_x and σ_x, are pinned down by the mean and volatility of aggregate output growth. While the drift and volatility parameters of idiosyncratic productivity, g_ε and σ_ε, are identified by the average median, IQR, and Kelly skewness of sales growth and investment rates, the jump process parameters, χ_0, χ_1 and λ, affect the standard deviation of IQR and Kelly skewness of sales growth and investment rates as well as their cyclicality.

4.5 Baseline Estimates

SMM parameter estimates are shown in Table 3, whereas data and model moments are shown in Table 4. Our benchmark specification is shown in the columns labeled Spec-1 (we will return to the alternative specifications below). As Table 3 shows, all estimated parameters are well-identified as indicated by very small standard errors. The estimated jump intensity of $\lambda = 0.0941$ implies that firms receive negative jumps in productivity about once every 11 quarters, whereas the estimated parameters of the jump size function ($\chi_0 = 0.2384$ and $\chi_1 = 0.7027$) imply that the average jump size is about -31%. The log growth rate of idiosyncratic productivity has a Gaussian volatility of $\sigma_\varepsilon = 5.39\%$ and drift parameter of $g_\varepsilon = 1.46\%$ per quarter, well below the threshold of $\pi = 2\%$ that is required to ensure finiteness of the cross-sectional mean of productivity – see Equation 6. Lastly, the log growth rate of aggregate productivity has a drift parameter of $g_x = 0.30\%$ and a volatility of $\sigma_x = 3.56\%$ per quarter.

To gain a better understanding of how the model matches the sales growth and investment rate distributions, we plot the idiosyncratic productivity, sales growth, and investment rate distributions averaged across expansions (black line) and recessions (red line) in Figure 4.[19] Intuitively, sales growth is a weighted average of idiosyncratic productivity shocks and changes in capital, i.e., investment rates. Due to costly reversibility, a large fraction of firms is inactive with zero investment rates. Firms with large positive idiosyncratic productivity draws have positive investment rates, causing a positively skewed investment rate distribution. At the same time, investment rates are lower on average in recessions but more dispersed in expansions.

The positive skewness in investment rates also makes sales growth positively skewed in expansions. When the economy switches into a recession due to a sequence of adverse aggregate productivity draw, the jump size χ increases and the productivity distribution becomes strongly left skewed. Importantly, the mass of firms receiving a negative Poisson draw does not increase. This negative skewness in productivity dominates the positive skewness in investment rates and thus sales growth becomes left skewed as in the data. The negative skewness in productivity also raises the dispersion in productivity and sales growth.

This mechanism explains why the model cannot fit all target moments in Table 4 perfectly. While the average median sales growth and investment rate are too small relative to the data, increasing the drift of idiosyncratic productivity would yield a better fit. But the drift also

[19] We define an expansion as a period where four quarter aggregate output growth falls below its' unconditional 25-th percentile. Correspondingly, expansions are defined as output being above its' 75-th percentile.

has a positive impact on the average IQR of sales growth and investment rates and especially the latter is already too large. Increasing idiosyncratic volatility would yield a better fit for the average IQR of sales growth and investment rates but, as explained above, it also raises the average skewness of sales growth and investment rates, which are closely matched.

The parameters that govern Poisson jumps in idiosyncratic productivity have a strong effect on the volatility of cross-sectional moments and their cyclicality. While making Poisson jumps more likely increase the volatility of IQR and Kelly skewness of sales growth and investment rates, this effect also leads to counterfactual large aggregate consumption volatility because the capital misallocation hinders consumption smoothing – see our discusssion in Section 5.2. The cyclicality and size of Poisson jumps also have a strong positive impact on the volatility of IQR and Kelly skewness of sales growth and investment rates, these parameters also affect the cyclicality of IQR and Kelly skewness of sales growth. Increasing the jump size parameters would add more time variation in terms of IQR of sales growth and investment rates at the cost of too much cyclicality.

4.6 Alternative Specifications

In this section, we illustrate estimation results for a number of alternative model specifications in order to highlight the role of pre-specified parameters. The results of these experiments are contained in the additional columns of Tables 3 and 5.

Columns labeled Spec-2 show results for a preference parameter calibration that implies time-separable utility. In particular, we calibrate the EIS to a low value of $\psi = 0.5$ and the relative risk aversion coefficient to 2, similar to the values typically assumed in the macroeconomics literature. Table 3 shows that the estimated productivity parameters are very similar to those in our benchmark specification, with the exception of χ_1 and g_x. The estimated parameter values imply that the size of productivity jumps is less cyclical than in the benchmark, whereas aggregate quantities grow at a lower rate. Table 4 shows that, while cross-sectional moments are matched similarly well as in the benchmark specification, the volatilities of aggregate output, consumption, and investment are very different from their data counterparts. Therefore, the low EIS leads to a tension between matching cross-sectional and aggregate facts. While one could certainly match quantity volatilities by assigning a larger weight to these moments in the criterion function, this would come at the expense of no longer matching cross-sectional facts.

The next alternative specification labeled Spec-3 changes the benchmark specification by assuming a higher exit rate of $\pi = 3\%$ as opposed to 2% in the benchmark. While this only leads to small differences in estimated productivity parameters and fit of the targeted moments, the higher exit rate implies a lower power law coefficient, i.e., a lower concentration of large firms. As shown in Table 5 and Figure 8 (both of which we discuss below), this results in an improved ability to smooth consumption, lower risk premia, and less misallocation relative to the benchmark.

5 Model Implications

In this section, we study the consumption dynamics and asset pricing implications, arising from capital misallocation. Quantitatively, the amount of capital misallocation in the model depends on the firms' life cycle effect combined with power law in firms size. Intuitively, the aggregate consequences of capital misallocation are worse when large old firms dominate the economy because older firms tend to be less efficient than young firms.

Finally, we can also study the empirical misallocation through the lens of over model. To this end, we fit the output gap to cross-sectional moments of the sales growth and investment rate densities in the model. Given this fit, we use the empirical counterparts to construct an empirical output gap measure, as in Eisfeldt and Muir (2016).

5.1 Firms' Life Cycle

To understand the nature of aggregate fluctuations in our model, it is useful to first characterize the behavior of a typical firm over its life cycle. Figure 6 shows how various firm characteristics change as firms age. Firms enter the economy with zero capital and positive productivity, so that young firms tend to have high investment (top-left) and sales growth rates (top-right), and low payout (bottom-left) and capital-to-productivity ratios (bottom-right). Due to the geometric growth in idiosyncratic productivity, firms' investment rates continue to exceed their depreciation rates as they age, and older firms are larger on average. At the same time, older firms are more likely to have experienced a negative Poisson shock to productivity during their lifetime. After experiencing such a shock, firms' capital-to-productivity ratio increases and, because disinvestment is costly, firms tend to become inactive and let their capital stock depreciate over the coming periods. A large reason why older firms have lower investment rates is therefore that they are more likely to be constrained, i.e. the extensive margin is lower among older firms. To summarize, the typical old firm is larger and more constrained relative to the typical young firm.

5.2 Power Law and Consumption Dynamics

The combination of a unit root in idiosyncratic productivity and random exit in our model results in a firm size distribution, whose right tail exhibits a power law. In particular, the log right tail probabilities (above the 50est percentile), with firm size measured by log capital, lie on a straight line with a slope of -1.24 for our benchmark estimation results. This means that the firm size distribution is well approximated in its right tail by a Pareto distribution with right tail probabilities of the form $1/x^{\zeta}$, with a tail index ζ of 1.24. To illustrate the economic effect of the power law, Figure 7 shows the degree of output concentration implied by our benchmark parameter estimates. Specifically, it shows the fraction of aggregate output produced by firms in various percentiles of the capital distribution.[20] On average, the largest 5% of firms (in terms of capital) produce almost 50% of aggregate output.

Computing the data analogue of this statistic requires a count of the number of firms in the U.S. economy. The 2013 U.S. census reported a total of 5.8 million firms. The majority of firms has between 0 and 4 employees. Given a minimum of 100 employees, the US economy consisted of 103,900 firms; for a minimum of 500 employees, it had 18,636 firms. In the same year, CRSP-Compustat universe included about 2,500 firms (excluding financial firms and utilities), whose sales accounted for 55% of U.S. GDP. Depending on the definition of a firm in the data, the output concentration in the data is even larger than in our model.

[20] To produce the figure, we simulate a continuum of firms and record the fraction of output produced by the 5% smallest firms (in terms of capital), firms between the 5th and 10th size percentiles, etc. in each period. We then average these fractions across all periods in the simulation.

Due to the importance of large firms for output and consumption in the model, permanent negative shocks to their productivity are particularly painful. The discussion in Section 5.1 highlighted that large firms tend to be more constrained on average. Thus, constrained firms are responsible for a large fraction of the drop in consumption during recessions. While unconstrained firms increase dividends by reducing investment, they are smaller on average so that they are not able to offset the impact of large constrained firms on aggregate consumption. In other words, the firm size distribution in combination with negative jumps in productivity implies that it is difficult for the representative household to smooth consumption during recessions. In contrast, in models with log-normal productivity distributions the size difference between constrained and unconstrained firms is small so that the groups offset each other.

Figure 8 illustrates this channel quantitatively via a comparative statics exercise, that varies the exit probability π. A larger exit probability implies that firms survive shorter on average, which reduces the mass of large firms. The top-left panel shows that the power law coefficient $(-\zeta)$ decreases as π increases, meaning that the right tail of the firm size distribution becomes thinner. This implies that it becomes easier to smooth consumption by offsetting the losses in consumption from large firms during recessions with the dividend payments of unconstrained (and relatively smaller) firms. As a consequence, the unconditional left skewness of consumption growth is reduced (top-right panel), the loss in output relative to the frictionless benchmark decreases (bottom-right panel), and risk premia decline (bottom-left panel).

In Table 5, we summarize consumption growth moments, risk premia, and misallocation measures across the three specifications of Tables 2 and 3. In the benchmark specification 3, consumption growth is positively autocorrelated, left skewed, and exhibits excess kurtosis. In the frictionless economy, consumption growth is normal distributed. Relative to the power utility case (specification 2), this increased tail risk has a significant effect on the expected return on wealth, which increases to 1.8%, on the risk-free rate, which decreases to 1.5%, and on the Sharpe ratio, which increases to 0.57.

We next explore the effect of misallocation in the model. As discussed in Section 3.4, the frictionless model does not feature any misallocation, so that $\mathcal{M}$ and $\mathcal{D}$ are both constant and equal to zero. Given our parameter estimates for productivity, we find that capital misallocation is 4.6% on average. It is also very volatile, persistent, and cyclical. Capital misallocation also leads to a significant output gap of the full model relative to the frictionless economy of 4.5%. By eliminating capital frictions, the economy would exhibit a large one time rise in wealth.

5.3 Measuring Misallocation Empirically

Our model economy allows us to quantify misallocation in the data by providing a link between the output gap and cross-sectional densities. In particular, the output gap in the model is well-described by a function of cross-sectional moments. We can therefore estimate this relationship in the model and then apply it to the empirically observed moments to back out the output gap in the data. Based on model-generated data, we estimate a linear projection of the output gap on its first lag and the six moments that summarize the cross-sectional sales growth and investment rate densities.[21] We then compute a fitted value based on the estimated coefficients and the cross-sectional moments in the data. The initial value of the lagged output gap is set to the model-implied unconditional average.

The resulting time series for the empirical output gap in Figure 5 shows that misallocation spiked up during all recessions in our sample period. The recessions of 2001 and 2008 were associated with particularly high levels of misallocation, suggesting that cross-sectional inefficiencies were an important driver of these downturns. This view is further supported by the fact that misallocation remained elevated for an extended period in both episodes, thereby contributing to the observed slow recoveries.

5.4 Shock persistence in the data

In our model, firms that are hit by large negative jump shocks either sell capital or become inactive for an extended period, thereby contributing to an increase in misallocation cross-sectionally. Because this effect results from the assumption that shocks are permanent, a

[21] The population R^2 in the model is approximately 82%.

potential concern is that the model overstates the true effect that these shocks have on firms' investment. If extreme shocks to sales were completely transitory for example, one would expect to see little to no effect in investment rates.

To investigate this issue in the data, we analyze the investment behavior of firms around severe shocks to their sales. Both in the model and the data, the most extreme such shocks occur during periods in which the cross-sectional skewness of sales growth becomes very negative. We thus create a pooled sample of all firm-quarter observations for which (1) the cross-sectional skewness falls below -0.1 and (2) the sales growth rate of the individual firm is at the 10-th percentile of the cross-sectional distribution – the cutoff that defines the left tail in Kelly's skewness.[22] To isolate the effect of firm-specific shocks from aggregate conditions, we further demean investment rates with the period-specific cross-sectional average.

The top panel of Figure 9 shows the average demeaned investment rate of our pooled sample, along with a 95% confidence region around the cross sectional average (which is normalized to zero). The gray-shaded region defines the event window and extends from -3 to +3 because investment rates are based on four quarters of data. The figure shows that the average investment rate of affected firms differs insignificantly from the cross-sectional average pre event, drops to about 10% below average on impact of the sales shock, and then remains significantly below average for the following 20 quarters. This evidence supports the assumption that extreme shocks to sales are well-described as being very persistent.

To examine whether the effect is quantitatively consistent with the model, we repeat the same experiment based on simulated model data, using both a large cross-section of firms and a long time series. The bottom panel of Figure 9 shows that the dynamic response of investment rates in the model, while slightly more pronounced than in the data, is very comparable to that in the data. We interpret this evidence as very supportive of the main mechanism for generating misallocation in our model.

[22]There were four episodes during which Kelly's skewness fell below -0.1 in our sample (see figure 3), during some of which it remained below the threshold for a few quarters. Because we are interested in the response to the initial shock, we only include the first quarter during each such sequence. Additionally, we include all firms between the 5-th and 15-th percentile to capture the behavior of firms around the 10-th percentile in order to reduce noise.

6 Conclusion

[to be added]

7 Appendix A: Model without adjustment costs

In this appendix, we derive firms' optimal investment policy as well as aggregate quantities in the absence of adjustment costs, i.e. for the nested case $\xi = 1$. This analysis shows that the aggregate state in the detrended economy collapses to the detrended aggregate capital stock $\bar{k}$ and the current aggregate shock η_x, so that the model can be solved via the equivalent social planner problem.

7.1 Firm optimization

The firm solves

$$v(\kappa, \mu) = \max_{\tau} \left\{ \kappa^{\alpha} - \tau + (1-\delta)\kappa + \mathbb{E}\left[M' \left((1-\pi) x'\varepsilon' v\left(\frac{\tau}{x'\varepsilon'}, \mu' \right) + \pi\tau \right) \right] \right\},$$

subject to $\kappa' = \frac{\tau}{x'\varepsilon'}$ and the transition of μ. The FOC equals

$$1 = \mathbb{E}\left[M' \left((1-\pi) \frac{\partial v'}{\partial \kappa'} + \pi \right) \right].$$

Combining this with the envelope condition $\frac{\partial v}{\partial \kappa} = \alpha\kappa^{\alpha-1} + 1 - \delta$ gives the Euler equation

$$1 = (1-\pi)\mathbb{E}\left[M' \left(\alpha \left(\frac{\mathcal{T}(\mu)}{x'\varepsilon'} \right)^{\alpha-1} + 1 - \delta + \frac{\pi}{1-\pi} \right) \right],$$

which in turn can be solved for the optimal target

$$\mathcal{T}(\mu) = \left(\frac{1 - (1-\pi)(1-\delta+\frac{\pi}{1-\pi})\mathbb{E}\left[M'\right]}{(1-\pi)\alpha\mathbb{E}\left[M'(x'\varepsilon')^{1-\alpha}\right]} \right)^{\frac{1}{\alpha-1}}. \tag{34}$$

While the problem is written in terms of μ as the aggregate state, we show below that $(\bar{k}, \eta_x)$ is a sufficient state. Based on the optimal capital target, the law of motion for firms' capital can be expressed in terms of the capital-to-productivity ratio as $\kappa' = \frac{\mathcal{T}}{x'\varepsilon'}$, or in terms of detrended capital as

$$k' = \frac{\mathcal{E}\mathcal{T}}{x'}. \tag{35}$$

7.2 Aggregation

Using the capital policy of an individual firm allows us to express detrended aggregate capital as

$$
\begin{aligned}
\bar{k} &= \int k \,\mathrm{d}\mu \\
&= (1-\pi)\int \frac{\mathcal{E}_{-1}\mathcal{T}_{-1}}{x}\,\mathrm{d}\mu_{-1} \\
&= (1-\pi)\frac{\mathcal{T}_{-1}}{x}. \qquad (36)
\end{aligned}
$$

The second equality follows from splitting today's μ into the mass $(1-\pi)$ of incumbents and the mass π of entrants, substituting (35) for the capital of incumbents, and noting that entrants have a capital stock of zero. The third equality follows from $\mathbb{E}[\mathcal{E}] = 1$. Detrended aggregate output is given by

$$
\begin{aligned}
\bar{y} &= \int \mathcal{E}^{1-\alpha}k^{\alpha}\,\mathrm{d}\mu \\
&= (1-\pi)\int \mathcal{E}_{-1}^{1-\alpha}\left(\frac{\mathcal{E}_{-1}\mathcal{T}_{-1}}{x}\right)^{\alpha}\mathrm{d}\mu_{-1} \times \mathbb{E}[\varepsilon^{1-\alpha}|\eta_x] \\
&= (1-\pi)\left(\frac{\mathcal{T}_{-1}}{x}\right)^{\alpha} \times \mathbb{E}[\varepsilon^{1-\alpha}|\eta_x] \\
&= (1-\pi)^{1-\alpha}\bar{k}^{\alpha} \times \mathbb{E}[\varepsilon^{1-\alpha}|\eta_x],
\end{aligned}
$$

where the second equality follows from once again splitting μ into incumbents and entrants, and then factoring $\mathcal{E}$ as $\mathcal{E}_{-1}\varepsilon$. The third equality follows from $\mathbb{E}[\mathcal{E}] = 1$, and the last equality follows from (36). The term $\mathbb{E}[\varepsilon^{1-\alpha}|\eta_x]$ captures the effect of predetermined capital stocks on aggregate output and is given by

$$
\begin{aligned}
\mathbb{E}[\varepsilon^{1-\alpha}|\eta_x] &= \int \exp\{g_\varepsilon - \sigma_\varepsilon^2/2 + \sigma_\varepsilon\eta_\varepsilon + \chi J - \lambda(e^{\chi}-1)\}^{1-\alpha}\,\mathrm{d}\eta_\varepsilon\,\mathrm{d}J \\
&= \exp\{(1-\alpha)g_\varepsilon - (1-\alpha)\sigma_\varepsilon^2/2 + (1-\alpha)^2\sigma_\varepsilon^2/2 + \lambda(e^{(1-\alpha)\chi}-1) - (1-\alpha)\lambda(e^{\chi}-1)\}.
\end{aligned}
$$

Note that this term depends on the shock to aggregate productivity η_x via its effect on the jump size χ. Detrended aggregate investment equals

$$
\begin{aligned}
\bar{i} &= \int \mathcal{E}\mathcal{T} - (1-\delta)k\,\mathrm{d}\mu \\
&= \mathcal{T} - (1-\delta)\bar{k},
\end{aligned}
$$

so that detrended consumption is given by

$$c = \bar{y} - \mathcal{T} + (1-\delta)\bar{k} + \frac{\pi}{1-\pi}\bar{k}. \tag{37}$$

The above analysis shows that both consumption and the law of motion for aggergate capital depends on μ only via $\bar{k}$ and η_x. As we show next, this implies that the model can be solved via an equivalent social planner problem.

7.3 Social planner

The planner maximizes social welfare by solving

$$U(\bar{k}, \eta_x) = \max_c \left\{ (1-\beta)c^{1-\frac{1}{\psi}} + \beta \mathcal{R}(\bar{k}, \eta_x)^{1-\frac{1}{\psi}} \right\}^{\frac{1}{1-\frac{1}{\psi}}},$$

subject to the resource constraint (37) and the law of motion (35), where $\mathcal{R}(\bar{k}, \eta_x) \equiv \left(\mathbb{E}\left[(x')^{1-\gamma}(U')^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}$ denotes the certainty equivalent of continuation welfare. For more concise notation in the derivations that follow, let $U \equiv U(\bar{k}, \eta_x)$ and $\mathcal{R} \equiv \mathcal{R}(\bar{k}, \eta_x)$. The FOC equals

$$0 = \frac{1}{1-\frac{1}{\psi}} U^{\frac{1}{\psi}} \left((1-\beta)\left(1-\frac{1}{\psi}\right) c^{-\frac{1}{\psi}} + \beta \left(1-\frac{1}{\psi}\right) \mathcal{R}^{-\frac{1}{\psi}} \mathbb{E}\left[\frac{\partial \mathcal{R}}{\partial U'} \frac{\partial U'}{\partial \bar{k}'} \frac{\partial \bar{k}'}{\partial c} \right] \right).$$

Combining the aggregate resource constraint with the law of motion for aggregate capital and re-arranging yields $\bar{k}' = \frac{1-\pi}{x'}\left(\bar{y} - c + (1-\delta)\bar{k} + \frac{\pi\xi}{1-\pi}\bar{k}\right)$, so that $\frac{\partial \bar{k}'}{\partial c} = -\frac{1-\pi}{x'}$. Substituting this back into the FOC and re-arranging yields

$$(1-\beta)c^{-\frac{1}{\psi}} = (1-\pi)\beta \mathbb{E}\left[\frac{\mathcal{R}^{-\frac{1}{\psi}}}{x'} \frac{\partial \mathcal{R}}{\partial U'} \frac{\partial U'}{\partial \bar{k}'} \right].$$

The first partial is given by

$$\begin{aligned} \frac{\partial \mathcal{R}}{\partial U'} &= \frac{1}{1-\gamma} \left(\mathbb{E}\left[(x')^{1-\gamma}(U')^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}-1} (x')^{1-\gamma}(1-\gamma)(U')^{-\gamma} \\ &= \left(\frac{U'}{\mathcal{R}}\right)^{-\gamma} (x')^{1-\gamma} \end{aligned}$$

The second partial is given by

$$\begin{aligned} \frac{\partial U'}{\partial \bar{k}'} &= \left(\frac{1}{1-\frac{1}{\psi}}\right) (U')^{\frac{1}{\psi}} (1-\beta)\left(1-\frac{1}{\psi}\right) (c')^{-\frac{1}{\psi}} \left(\alpha \frac{\bar{y}'}{\bar{k}'} + (1-\delta) + \frac{\pi}{1-\pi} \right) \\ &= (U')^{\frac{1}{\psi}} (1-\beta)(c')^{-\frac{1}{\psi}} \left(\alpha \frac{\bar{y}'}{\bar{k}'} + (1-\delta) + \frac{\pi}{1-\pi} \right) \end{aligned}$$

Once again substituting back into the Euler equation yields

$$
\begin{aligned}
(1-\beta)c^{-\frac{1}{\psi}} =&(1-\pi)\beta\mathbb{E}\left[\frac{\mathcal{R}^{-\frac{1}{\psi}}}{x'}\left(\frac{U'}{\mathcal{R}}\right)^{-\gamma}(x')^{1-\gamma}(U')^{\frac{1}{\psi}}(1-\beta)(c')^{-\frac{1}{\psi}}\left(\alpha\frac{\bar{y}'}{\bar{k}'}+(1-\delta)+\frac{\pi}{1-\pi}\right)\right] \\
\Leftrightarrow 1 =&(1-\pi)\mathbb{E}\left[\beta\left(\frac{c'}{c}\right)^{-\frac{1}{\psi}}\left(\frac{U'}{\mathcal{R}}\right)^{\frac{1}{\psi}-\gamma}(x')^{-\gamma}\left(\alpha\frac{\bar{y}'}{\bar{k}'}+(1-\delta)+\frac{\pi}{1-\pi}\right)\right] \\
=&(1-\pi)\mathbb{E}\left[M'\left(\alpha\mathbb{E}[(\varepsilon')^{1-\alpha}|\eta_x']\left(\frac{\mathcal{T}}{x'}\right)^{\alpha-1}+(1-\delta)+\frac{\pi}{1-\pi}\right)\right],
\end{aligned}
$$

where $M'=\beta\left(\frac{c'}{c}\right)^{-\frac{1}{\psi}}\left(\frac{U'}{\mathcal{R}}\right)^{\frac{1}{\psi}-\gamma}(x')^{-\gamma}$. The social planner thus chooses the capital target

$$
\mathcal{T}(\bar{k},\eta_x)=\left(\frac{1-(1-\pi)\left(1-\delta+\frac{\pi}{1-\pi}\right)\mathbb{E}[M']}{(1-\pi)\alpha\mathbb{E}\left[M'\mathbb{E}[(\varepsilon')^{1-\alpha}|\eta_x'](x')^{1-\alpha}\right]}\right)^{\frac{1}{\alpha-1}},
$$

which is identical to the target for an individual firm in equation 34. The model without adjustment costs can therefore be solved via the equivalent social planer problem. We do so by computing the planers value function and the associated adjustment target via value function iteration. We use a fine grid for $\bar{k}$, discretize η_x as in the full model (see the description in Appendix B), and use a cubic spline interpolation to evaluate the value function for off-grid values of $\bar{k}$.

8 Appendix B: Numerical Method

In this appendix, we detail our numerical solution approach. Numerical values for all computational parameters are summarized in table 2. Our algorithm iterates on the coefficients on the equilibrium forecasting rule for consumption (33) until convergence. The coefficients are initialized based on the model without adjustment costs, the solution to which is described in Appendix A. An iteration consists of the following steps, each of which is described in more detail below.

1. The infinite-dimensional state μ in the firm's problem (16)–(18) is replaced by normalized consumption c and the associated forecasting rule (33). The coefficients of the conditional c-forecasting rule, $(\varphi_0(\eta_x'), \varphi_1(\eta_x'))$, as well as the associates no arbitrage rule $w(c)$ are taken as given in the firm's problem. We compute the solution via value function iteration, which yields a set of adjustment targets $\mathcal{T}_d(c)$ and $\mathcal{T}_u(c)$ that describe the endogenous evolution of firms' detrended capital stocks.

2. Taking the adjustment targets $\mathcal{T}_d(c)$ and $\mathcal{T}_u(c)$ as given, simulate a continuum of firms over a large number of periods. This step is based on an extension of the non-stochastic simulation approach of **?**. In each period of the simulation, solve for the market clearing value of c based on the aggregate resource constraint (24).

3. Based on the simulated time series of equilibrium consumption, estimate new coefficients (φ_0', φ_1') for the c-forecasting rule. Using these coefficients, compute the no arbitrage rule $w(c)$ based on the Euler equation for the return on aggregate wealth. If $\frac{|\eta'-\eta|}{|\eta|}$ is less than Δ_η =1E-5 for the intercept and slope coefficients corresponding to each value of η_x', stop the algorithm. Otherwise, go back to step 1.

Once the algorithm has converged, we use the solution to (stochastically) simulate a panel of 100,000 firms. This simulation is based on the adjustment targets and the equilibrium consumption path resulting from the last iteration of the solution algorithm. Based on the simulated panel, we compute four quarter sales growth and investment rates, and use them to determine he cross-sectional moments that define the objective function of our SMM estimation. In what follows, we describe each step of our numerical solution method in more detail.

8.1 Firm Problem

We discretize the firm's state space (κ, c) using $n_\kappa = 500$ log-linearly spaced points on the interval $[0.01, 100]$ for the capital-to-productivity ratio, and $n_c = 50$ linearly spaced points on the interval $[x, x]$ for normalized consumption. Expectations over $(\eta'_x, \eta'_\varepsilon)$ are evaluated via Gaussian quadrature, using $n_{\eta_x} = 5$ and $n_{\eta_\varepsilon} = 11$ nodes respectively. This implies that the coefficients of the forecasting rule for c only need to be known for a finite set of η_x-values, so that they can be stored in an array. Expectation over Poisson shocks J are evaluated by truncating the set of possible outcomes at 8 since the probability of receiving more jumps is essentially zero for reasonable values of the jump intensity λ.

To evaluate the value function on off-grid values for (κ', c') we rely on a bivariate cubic spline interpolation. An iterated grid search is used to find the adjustment targets $\mathcal{T}_u(c)$ and $\mathcal{T}_d(c)$. We compute the solution by value function iteration, using 5 steps of policy improvement after each actual optimization step, and we iterate until the relative change of both targets is less than $\Delta_\mathcal{T}$ =1E-6 for all values on the c-grid.

8.2 Equilibrium Simulation

To update the coefficients in the c-forecasting rule, we rely on an equilibrium simulation of a continuum of firms over T^{sim} periods. The effect of Monte Carlo noise is minimized by stratifying the realization of the aggregate shock η_x, and by relying on the nonstochastic simulation method of **?** in order to update the measure of firms over time. Specifically, μ is stored on a bivariate grid for $(k, \mathcal{E})$ that contains $n_{\mu,k}$ points for detrended capital and $n_{\mu,\mathcal{E}}$ points for idiosyncratic productivity, with $\mu(k_i, \mathcal{E}_j)$ denoting the value at a particular gridpoint.

As described in more detail below, we update μ from one period to the next based on the exogenous law of motion for idiosyncratic productivity and firms' equilibrium investment decisions. The later depend on detrended consumption via its effect on firms' common adjustment targets $\mathcal{T}_d(c)$ and $\mathcal{T}_u(c)$. We determine the equilibrium value of c in each period by solving the fixed point problem associated with the aggregate resource constraint (24). As a result, the evolution of c in the simulation is based on market clearing rather than on the forecasting rule (33). In computing equilibrium consumption, we rely on the adjustment targets that resulted from the previous solution of the firm problem, and evaluate them on off-grid values based on a cubic spline interpolation. The consumption fixed point is found with a bisection method and a precision of Δ_c =1E-10.

Updating the measure of firms.– The mapping $\Gamma : (\mu, \eta'_x) \mapsto \mu'$ is adjusted to the discretized support of μ as follows. For detrended capital, note that it will typically be the case that $k' \notin \{k_1, ..., k_{n_{\mu,k}}\}$, i.e. the transition law corresponding to $(k_i, \mathcal{E}_j, c, \eta'_x)$ will not fall on a gridpoint for k. We assign probability mass resulting from such outcomes to the k-grid according to a linear weighting scheme. Specifically, for $k' \in (k_m, k_{m+1}]$, we assign weights of $\frac{k_{m+1}-k'}{k_{m+1}-k_m}$ and $\left(1 - \frac{k_{m+1}-k'}{k_{m+1}-k_m}\right)$ to k_m and k_{m+1} respectively. Letting $k'_{i,j} \equiv k'(k_i, \mathcal{E}_j, c, \eta'_x)$, next period's marginal capital distribution evaluated at a gridpoint k_m is given by[23]

$$\mu'(k_m|\eta'_x) = (1-\pi)\left\{\sum_{i=1}^{n_{\mu,k}}\sum_{j=1}^{n_{\mu,\varepsilon}} \mu(k_i, \mathcal{E}_j)\left[\mathbf{1}_{\{k'_{i,j}\in(k_m,k_{m+1}]\}}\frac{k_{m+1}-k'_{i,j}}{k_{m+1}-k_m}\right.\right.$$
$$\left.\left. +\mathbf{1}_{\{k'_{i,j}\in(k_{m-1},k_m]\}}\left(1-\frac{k_{m+1}-k'_{i,j}}{k_{m+1}-k_m}\right)\right]\right\} + \pi\mathbf{1}_{\{k_m=0\}}.$$

The transition function of the marginal productivity distribution is discretized by integrating the normal probability density functions of $\log(\varepsilon')$ and $\log(\mathcal{E}_0)$ over adjacent grid midpoints.

[23] In practice, any probability mass associated with policies that fall outside of the grid for $\log(k)$ is allocated to $\log(k_1)$ and $\log(k_{n_k})$. Specifically, for $k' \in (-\infty, k_1]$, we assign a weight of 1 to k_1 and for $k' \in (k_{n_{\mu,k}}, \infty)$, we assign a weight of 1 to $k_{n_{\mu,k}}$. We choose the boundaries of the support to ensure that the probability mass at these points is always less than 0.01%.

Letting $\mathcal{E}_{n,n+1} \equiv \frac{\mathcal{E}_n + \mathcal{E}_{n+1}}{2}$ for $n \in \{1, ..., n_{\mu,\mathcal{E}} - 1\}$, $\mathcal{E}_{0,1} \equiv -\infty$, and $\mathcal{E}_{n_{\mu,\mathcal{E}}, n_{\mu,\mathcal{E}}+1} \equiv \infty$, we get

$$\mu'(\mathcal{E}_n|\eta'_x) = (1-\pi)\Bigg\{\sum_{j=1}^{n_{\mu,\mathcal{E}}}\sum_{l=0}^{\infty}\mu(\mathcal{E}_j)p_l\Bigg[\Phi\left(\frac{\log\left(\frac{\mathcal{E}_{n,n+1}}{\log(\mathcal{E}_j)}\right) - \left(g_\varepsilon - \frac{\sigma_\varepsilon^2}{2} + \chi' l - \lambda\left(e^{\chi'} - 1\right)\right)}{\sigma_\varepsilon}\right)$$
$$- \Phi\left(\frac{\log\left(\frac{\mathcal{E}_{n-1,n}}{\log(\mathcal{E}_j)}\right) - \left(g_\varepsilon - \frac{\sigma_\varepsilon^2}{2} + \chi' l - \lambda\left(e^{\chi'} - 1\right)\right)}{\sigma_\varepsilon}\right)\Bigg]\Bigg\}$$
$$+\pi\Bigg\{\Phi\left(\frac{\log(\mathcal{E}_{n,n+1}) - g_0 + \sigma_0^2/2}{\sigma_0}\right) - \Phi\left(\frac{\log(\mathcal{E}_{n-1,n}) - g_0 + \sigma_0^2/2}{\sigma_0}\right)\Bigg\},$$

where $\mu(\mathcal{E}_j) \equiv \sum_{i=1}^{n_{\mu,k}} \mu(k_i, \mathcal{E}_j)$ denotes the marginal distribution of idiosyncratic productivity and p_l is the probability mass function of a Poisson random variable evaluated at l, i.e. the probability of observing l jumps. Combining the two marginal transition functions yields the mapping Γ as

$$\mu'(k_m, \mathcal{E}_n|\eta'_x) = \mu'(\mathcal{E}_n|\eta'_x) \times \mu'(k_m|\eta'_x).$$

Grid and initial distribution.– Given the power law in firm size, the numerical accuracy of the histogram-based simulation approach relies crucially on appropriately chosen gridpoints for k and $\mathcal{E}$. What matters is the approximation error that results from replacing the integral $\int \mathcal{E}^{1-\alpha}k^\alpha \, \mathrm{d}\mu$ with the sum $\sum_{i=1}^{n_{\mu,k}}\sum_{j=1}^{n_{\mu,\mathcal{E}}} \mathcal{E}_j^{1-\alpha}k_i^\alpha \mu(k_i, \mathcal{E}_j)$, and equivalent approximation errors for the integrals over investment and capital that appear in the resource constraint. A gridpoint allocation that equalizes the contribution of each summand to the overall sum is therefore a sensible choice. Because we work with orthogonal grids, we determine such an output-weighted grid separately for each dimension of μ. Specifically, we set the values for $\{k_1, ..., k_{n_{\mu,k}}\}$ such that $\sum_{j=1}^{n_{\mu,\mathcal{E}}} \mathcal{E}_j^{1-\alpha}k_i^\alpha \mu(k_i, \mathcal{E}_j)$ is equal for each k_i. To compute these terms, we rely on the ergodic distribution of $(k, \mathcal{E})$ for $\mu(k_i, \mathcal{E}_j)$.[24] The $\mathcal{E}$-grid is chosen analogously. We find that this allocation method performs very favorably in terms of efficiency and numerical accuracy relative to both equally-spaced and log linearly-spaced grids. We initialize the equilibrium simulation at the ergodic distribution and discard the initial T^{burn} simulation periods to ensure that this choice is inconsequential.

[24]The ergodic distribution is found from on a long equilibrium simulation based on very fine log linearly-spaced grids. We use the adjustment targets that result from the solution of the firm problem in the first iteration. While the simulation based on these fine grids is very accurate, it is considerably too slow for the purpose of estimating the model. Nevertheless, it is well-suited for choosing gridpoints because we do so only once at the beginning of the estimation.

8.3 Updating forecasting and no arbitrage rules

Based on the simulated data for equilibrium consumption, we re-estimate the coefficients $(\varphi_0(\eta'_x), \varphi_1(\eta'_x))$ of the forecasting rule (33) via OLS. A major weakness of this approach as implemented in the previous literature is that the coefficients for less-likely values of the (discretized) aggregate shock have to be estimated based on very few observations.

We overcome this small sample problem by solving for equilibrium consumption for each possible value of the (discretized) aggregate shock in each simulation period. To see how this works, note that the transition of μ from the previous to the current period depends on the current aggregate shock η_x. Given μ_{-1}, it is therefore possible to determine μ and the associated equilibrium consumption for each hypothetical realization of η_x. We then rely on the distribution that corresponds to the shock that actually materialized to update the distribution in the subsequent period. As in previous papers, the simulated path of the economy is therefore based on a particular realized aggregate shock sequence. However, because we determine equilibrium consumption for both realized and non-realized shock values, we are able to update all forecasting rule coefficients based on T^{sim} observations. As a result, our approach allows us to achieve the same accuracy as in previous papers based on a far shorter simulation.

Given new coefficients for the c-forecasting rule (33), we update the no arbitrage rule for detrended wealth $w(c)$. First, we specify an equally-spaced grid of $n_w = 1,000$ values for c, along with initial values $w(c)$. Next, we compute new values for $w(c)$ based on the Euler equation (15). In particular, for each grid value for c, we use the forecasting rule (33) to determine c' and the no arbitrage rule to determine $w'(c')$. A cubic spline interpolation is used to evaluate $w(c)$ on off grid values for c'. Equation (15) then gives new values for $w(c)$. Lastly, we iterate on this procedure until the relative change in $w(c)$ falls below Δ_w =1E-12 for all grid values of c. Note that the grid for c employed in this step is different from the grid used to solve the firm problem.

Table 2: Computational Parameters

Parameter	Value	Description	Notes
Panel A: Value Function Iteration			
n_k	500	Gridpoints for κ	Equally-spaced in logs between 0.01 and 100
n_c	50	Gridpoints for c	Equally-spaced between x and x
n_{η_x}	5	# of numerical integration nodes for η_x	Gaussian quadrature
n_η	11	# of numerical integration nodes for η_ε	Gaussian quadrature
$\Delta_\mathcal{T}$	1E-6	Precision of investment targets	∞-norm of relative change
Panel B: Nonstochastic Equilibrium Simulation			
$n_{\mu,k}$	200	Gridpoints for k-dimension of μ	
$n_{\mu,\mathcal{E}}$	400	Gridpoints for $\mathcal{E}$-dimension of μ	
T^{sim}	3840	# of simulation periods	20 times data length plus T^{burn}
T^{burn}	800	# of discarted initial periods	Eliminates dependence on initial conditions
Δ_c	1E-10	Precision of consumption fixed point	
Panel C: Updating the Approximate Aggregation Rule			
n_w	1000	Gridpoints for $w(c)$	Equally-spaced between x and x
Δ_w	1E-12	Precision of no arbitrage wealth rule	∞-norm of relative change
Δ_η	1E-5	Precision of c-forecasting rule	∞-norm of relative change

The table reports parameters used in our numerical model solution.

References

ANDREW B. ABEL, J. C. E. (1996): "Optimal Investment with Costly Reversibility," *Review of Economic Studies*, 63, 581–593.

AXTELL, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293, 1818–1820.

BACHMANN, R. AND C. BAYER (2014): "Investment Dispersion and the Business Cycle," *American Economic Review*, 104, 1392–1416.

BANSAL, R. AND A. YARON (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

BARRO, R. J. (2006): "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 121, 823–866.

BARTELSMAN, E., J. HALTIWANGER, AND S. SCARPETTA (2013): "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review*, 103, 305–334.

BERGER, D., I. DEW-BECKER, AND S. GIGLIO (2016): "Contractionary volatility or volatile contractions?" Working paper.

BERK, J., R. C. GREEN, AND V. NAIK (1999): "Optimal Investment, Growth Options, and Security Returns," *Journal of Finance*, 54, 1553–1607.

BLOOM, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623–685.

BLOOM, N., M. FLOETOTTO, N. JAIMOVICH, I. SAPORTA-EKSTEN, AND S. J. TERRY (2014): "Really Uncertain Business Cycles," Working paper.

BOLDRIN, M., L. J. CHRISTIANO, AND J. D. M. FISHER (2001): "Habit Persistence, Asset Returns and the Business Cycle," *American Economic Review*, 91, 149–166.

CABALLERO, R. J. AND E. M. ENGEL (1999): "Explaining investment dynamics in US manufacturing: A generalized (S, s) approach," *Econometrica*, 67, 783–826.

CARLSON, M., A. FISHER, AND R. GIAMMARINO (2004): "Corporate Investment and Asset Price Dynamics: Implications for the Cross-section of Returns," *Journal of Finance*, 59, 2577–2603.

CONSTANTINIDES, G. M. AND A. GHOSH (2015): "Asset Pricing with Countercyclical Household Consumption Risk," Working Paper.

COOPER, I. (2006): "Asset Pricing Implications of Non-Convex Adjustment Costs of Investment," *Journal of Finance*, 61, 139–170.

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

Dixit, A. K. and R. S. Pindyck (1994): *Investment under Uncertainty*, Princeton, New Jersey: Princeton University Press.

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

Eisfeldt, A. and T. Muir (2016): "Aggregate External Financing and Savings Waves," *Journal of Monetary Economics*.

Eisfeldt, A. L. and A. Rampini (2006): "Capital Reallocation and Liquidity," *Journal of Monetary Economics*, 53, 269–399.

Epstein, L. and S. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Gabaix, X. (1999): "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, 114, 739–767.

——— (2009): "Power Laws in Economics and Finance," *Annual Review of Economics*, 1, 255–293.

——— (2011): "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 79, 733–772.

——— (2012): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance." *Quarterly Journal of Economics*, 127.

Gomes, J. F., L. Kogan, and L. Zhang (2003): "Equilibrium Cross Section of Returns," *Journal of Political Economy*, 111, 693–732.

Gomes, J. F. and L. Schmid (2010): "Levered Returns," *Journal of Finance*, 65, 467–494.

Gourio, F. (2012): "Disaster Risk and Business Cycles," *American Economic Review*, 102, 2734–2766.

Guvenen, F., S. Ozkan, and J. Song (2014): "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 122, 621–660.

Hsieh, C.-T. and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403–1448.

Jermann, U. J. (1998): "Asset Pricing in Production Economies," *Journal of Monetary Economics*, 41, 257–275.

Kaltenbrunner, G. and L. Lochstoer (2010): "Long-Run Risk through Consumption Smoothing," *Review of Financial Studies*, 23, 3141–3189.

Kelly, B., H. Lustig, and S. V. Nieuwerburgh (2013): "Firm Volatility in Granular Networks," Working Paper.

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

——— (2013): "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*, 121, 1055–1107.

Kogan, L. (2001): "An Equilibrium Model of Irreversible Investment," *Journal of Financial Economics*, 62, 201–245.

Krusell, P. and J. Smith, Anthony A. (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–96.

Luttmer, E. G. (2007): "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 1103–1144.

Olley, G. S. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1297.

Salgado, S., F. Guvenen, and N. Bloom (2015): "Skewed Business Cyles," Working Paper.

Schmidt, L. (2015): "Climbing and Falling Off the Ladder: Asset Pricing Implications of Labor Market Event Risk," Working Paper.

Veracierto, M. L. (2002): "Plant-Level Irreversible Investment and Equilibrium Business Cycles," *American Economic Review*, 92, 181–97.

Wachter, J. A. (2013): "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *Journal of Finance*, 68, 987–1035.

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

Figure 1: Optimal capital and investment rate policies

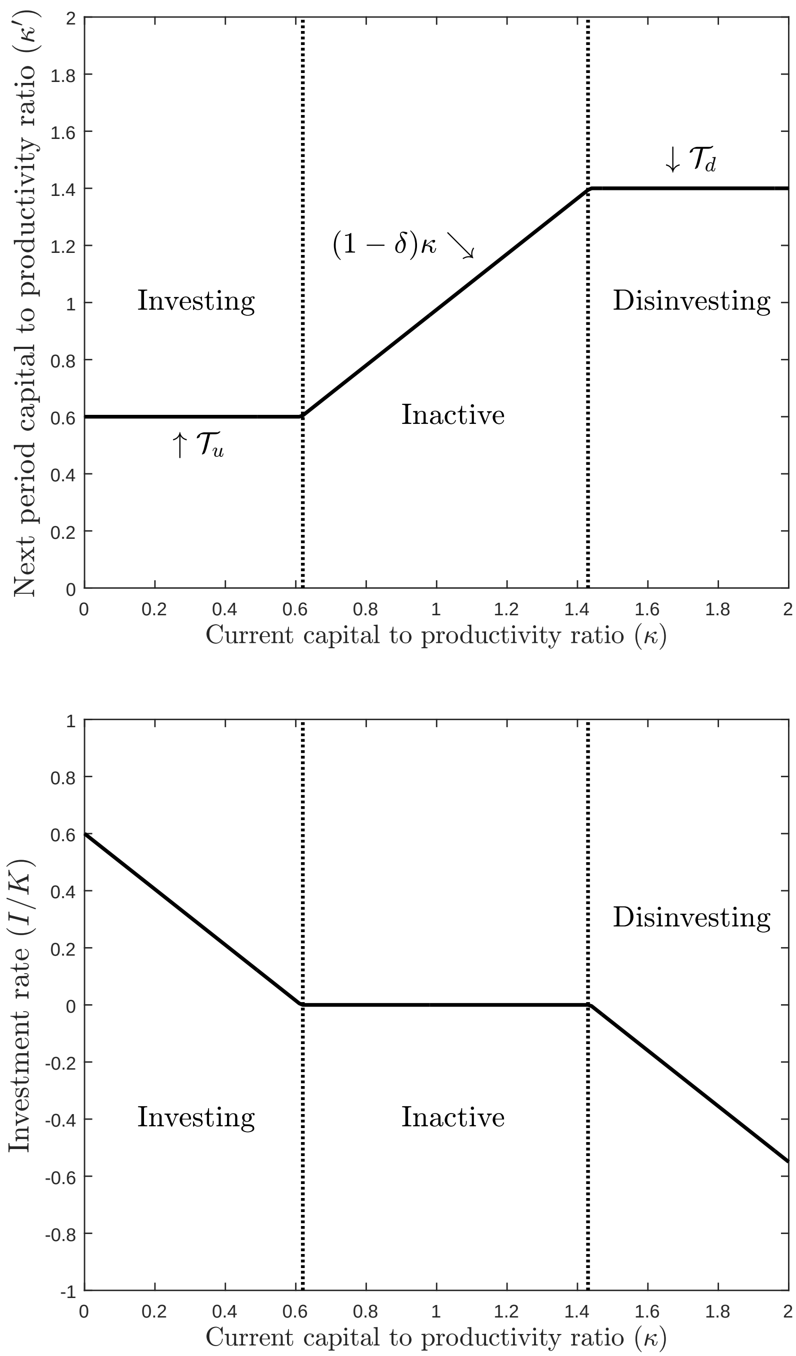


Figure 2: Optimal capital policies in the μ-distribution

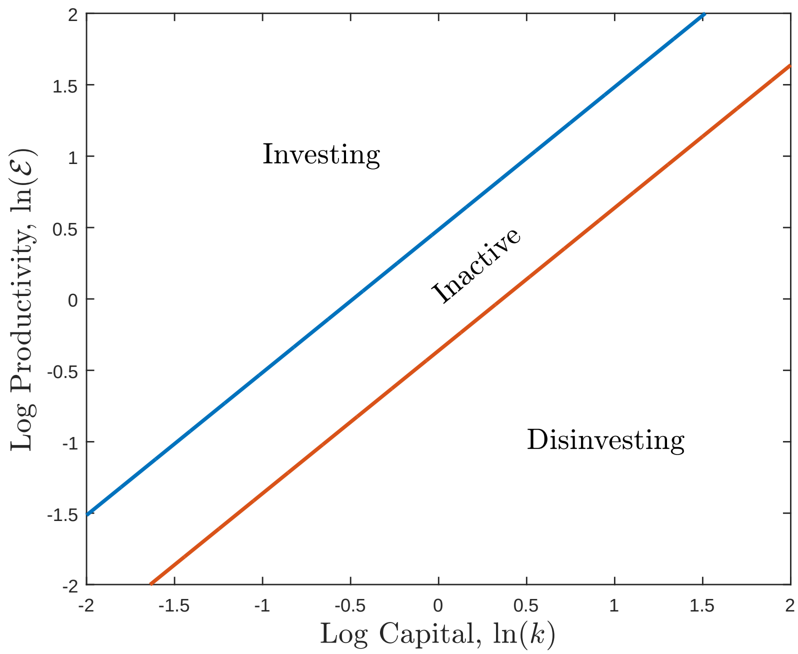
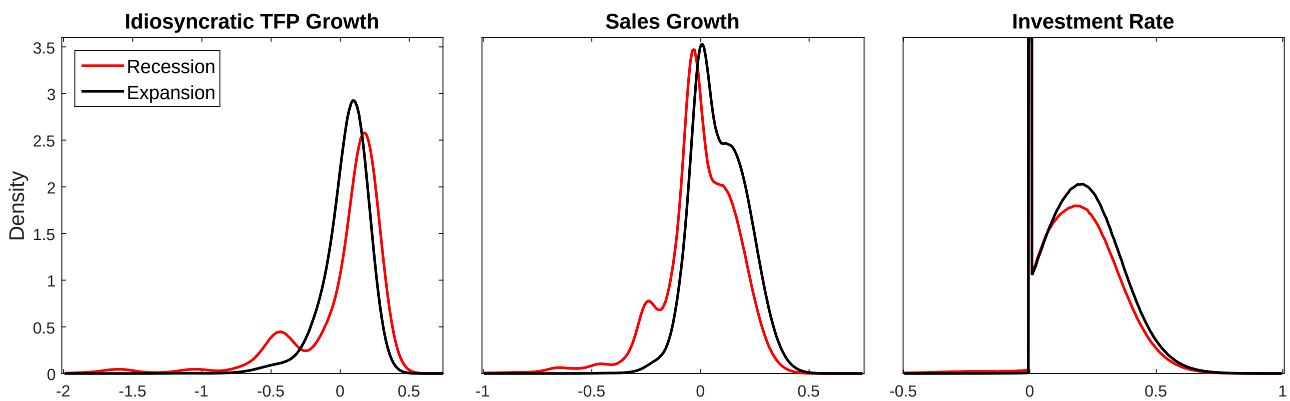


Figure 3: Robust cross-sectional moments of sales growth and investment rates

Notes: The figure shows robust cross-sectional moments of firms' four quarter sales growth and investment rates. The sample comes from CRSP-COMPUSTAT and spans 1976 to 2015.

Figure 4: Model-implied cross-sectional densities



Notes: The figure shows model-implied cross-sectional densities of four quarter moments for both expansion and recession periods. Recessions are defined as periods during which four quarter aggregate output growth falls below its unconditional 25th percentile, and expansions are periods where it falls above its 75th percentile. Productivity growth equals log growth rate in $X\mathcal{E}$, sales growth equals the log growth rate in Y, and the investment rate equals the sum of four quarterly investments divided by the capital stock in the initial quarter. To generate the figure, we simulate a large panel of firms based on the benchmark parameter estimates in Table 4.

Figure 5: Estimated output gap in the data

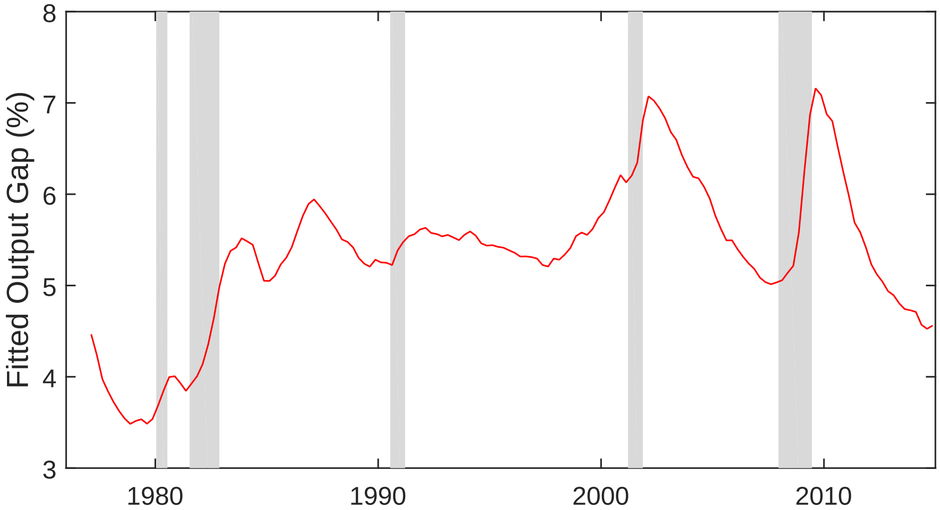


Figure 6: Life cycle of the average firm

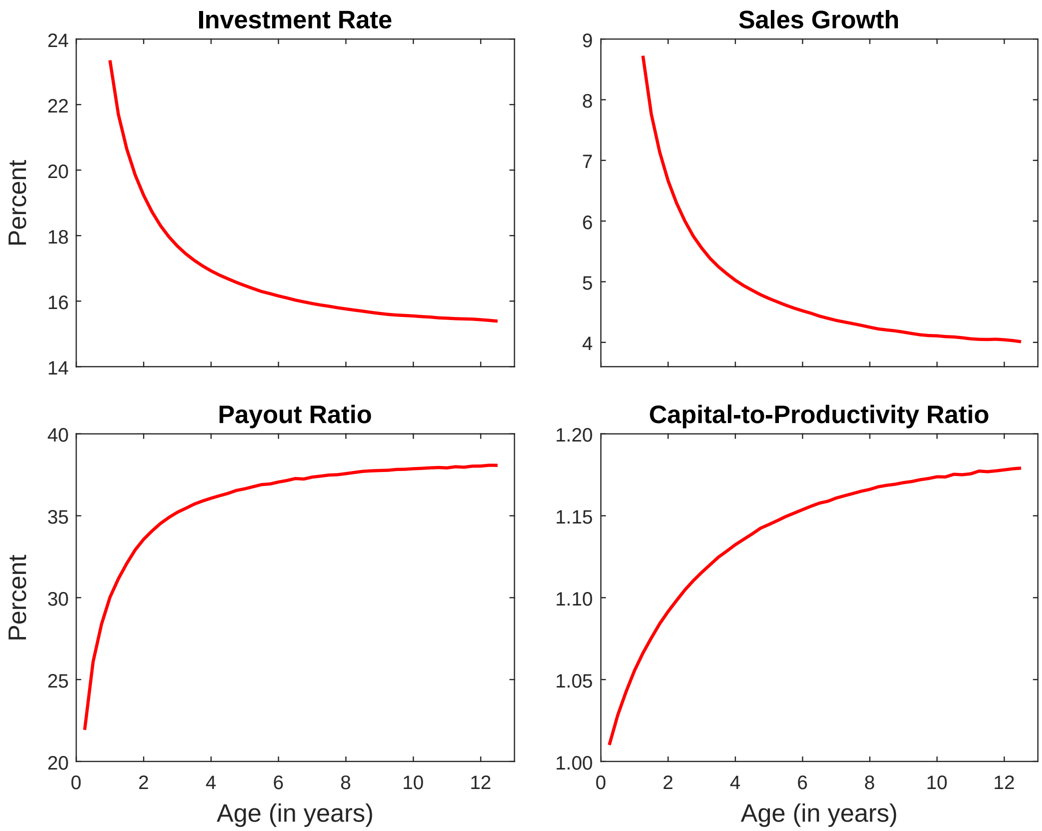
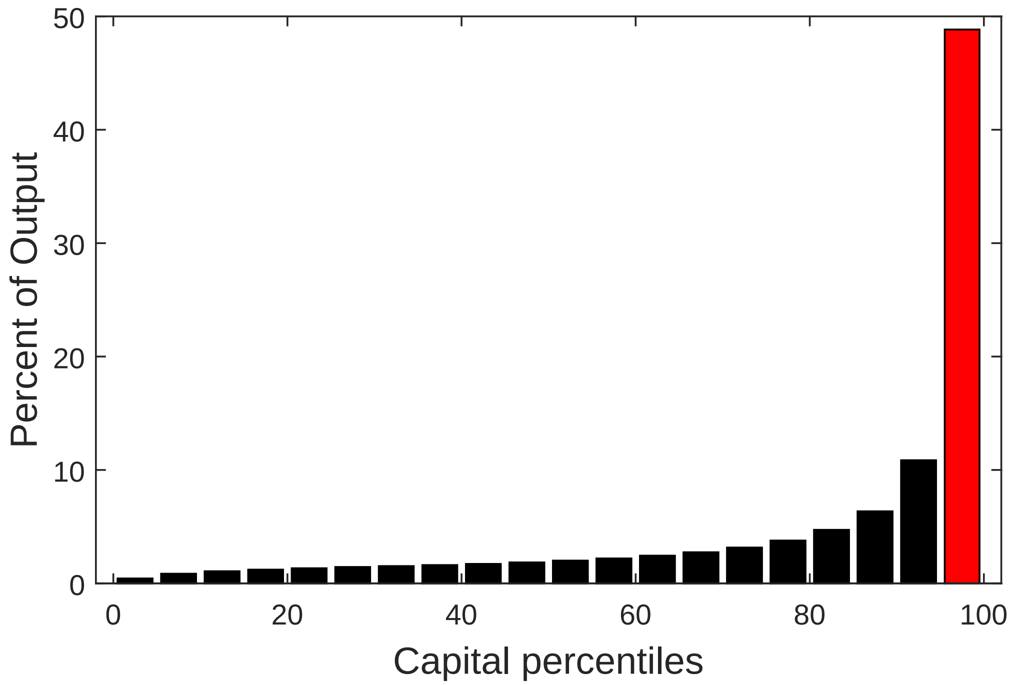


Figure 7: Output concentration



Notes: The figure shows the fraction of aggregate output that is produced by firms in different parts of the capital distribution.

Figure 8: Comparative statics for the exit rate

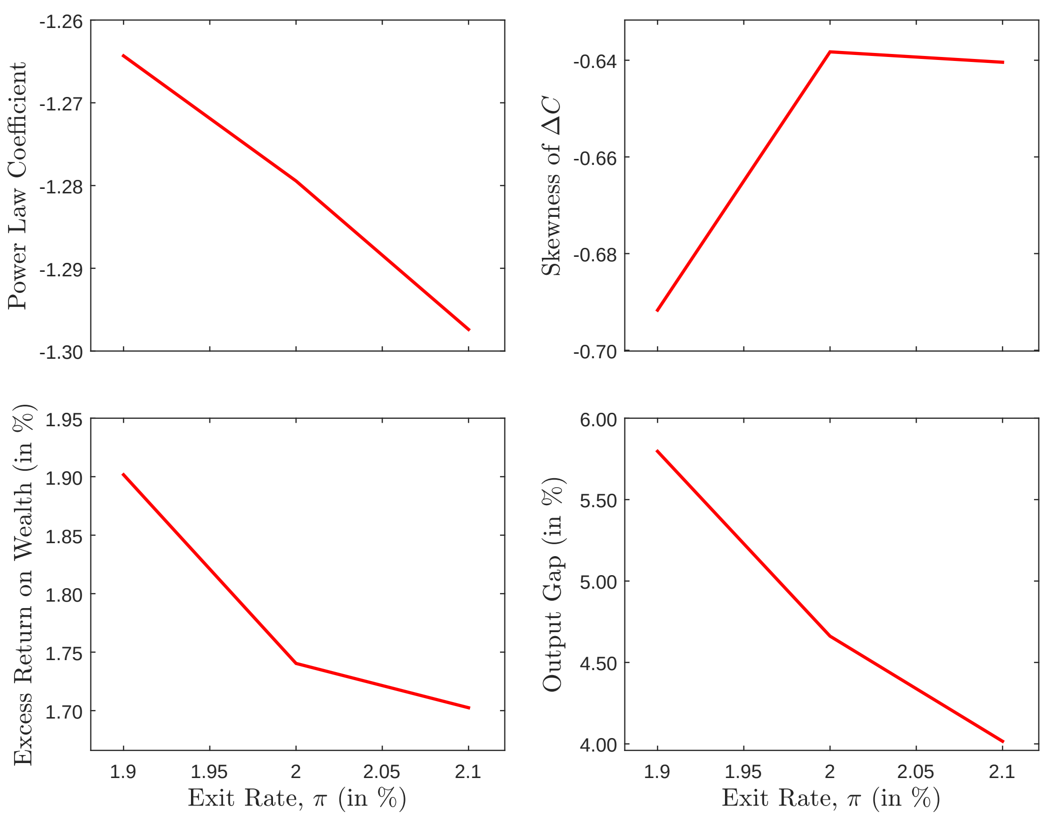
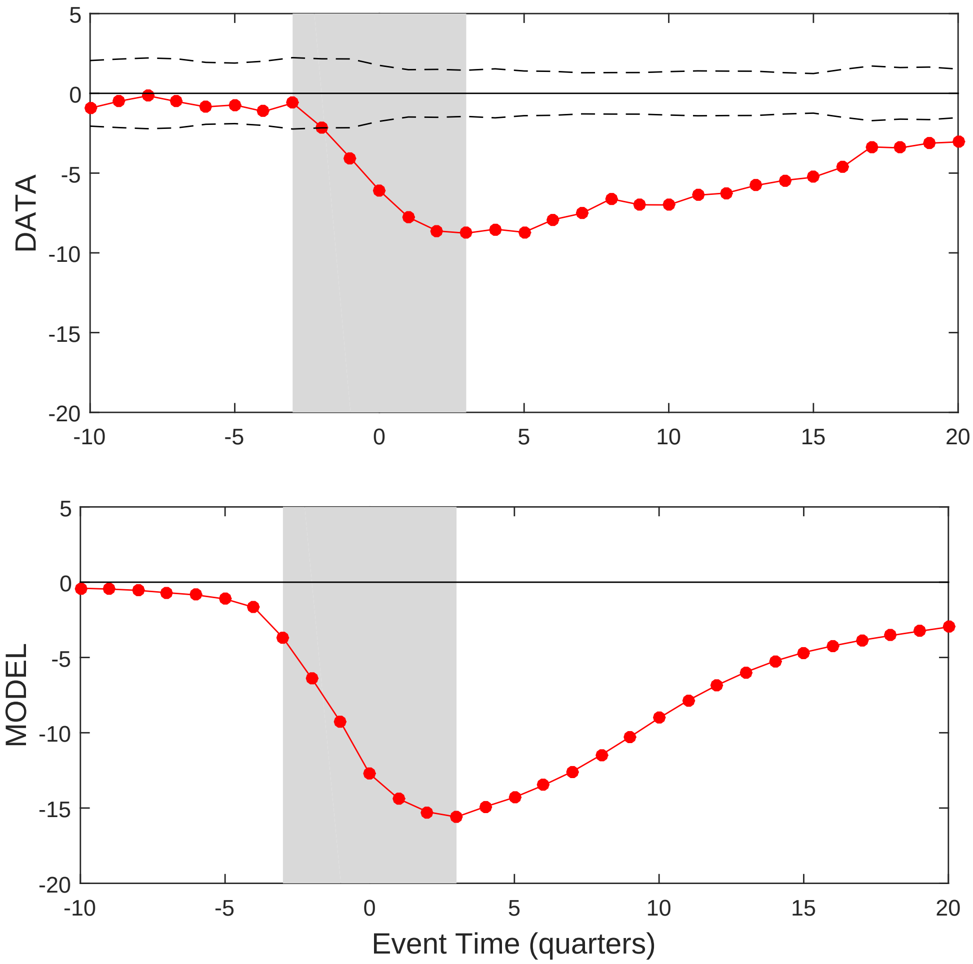


Figure 9: Investment rates for firms with extreme shocks to sales growth



Notes: The figure shows firms' average four quarter investment rates around severe idiosyncratic shocks to their sales. At event time zero, we pool all firm/quarter observations for which (i) the individual firm's sales growth rate lies between the 5-th and 15-th cross-sectional percentiles and (ii) the cross-sectional Kelly's skewness of sales growth is negative. The first criterion ensures that we select firms with extreme relative shocks while remaining robust to outliers. The second criterion ensures that the selected firm/quarter observations are also extreme in an absolute sense. When Kelly's skewness remains negative for multiple consecutive quarters, we only include the initial quarter of each such sequence to focus on the response to the initial shock. This results in 8 quarters being included in our pooled sample, for a total of 765 firm/quarter observations. We then average the same firms' sales growth and investment rates for quarters pre- and post- sample formation. The panel is unbalanced. Both sales growth and investment rates are demeaned with the cross-sectional average within the quarter. This isolates the pure cross-sectional effect from the aggregate shocks that hit all firms during the selected quarters. The dashed lines in the two data panels represent 95% confidence intervals. The gray event window region extends from event time -3 to +3 because we work with four quarter moments, so that an extreme annual sales growth rate at time 0 can be the result of an extreme quarterly observation anywhere between times -3 and +3. The sample comes from CRSP-COMPUSTAT and spans 1976 to 2015. The model panel is constructed equivalently based on simulated model data for a very large panel of firms over a very long sample.

Table 3: Predefined and Estimated Parameter Values

Parameter	Spec-1	Spec-2	Spec-3	Description
A: Predefined Parameters				
δ	0.025	0.025	0.025	Depreciation rate
α	0.65	0.65	0.65	Curvature in production function
σ_0	0.1	0.1	0.1	Volatility of productivity of new entrants
π	0.02	0.02	0.03	Exit probability
β	0.995	0.995	0.995	Time discount factor
ξ	0.7	0.7	0.7	Proportional resale value of capital
γ	10	2	10	Risk aversion
ψ	2	0.5	2	Elasticity of intertemporal substitution
B: Estimated Parameters				
χ_0	0.2384 [0.0000]	0.2915	0.2421	Parameter of jump size function
χ_1	0.7027 [0.0000]	0.4189	0.7300	Parameter of jump size function
λ	0.0941 [0.0000]	0.0896	0.0900	Parameter of jump intensity function
g_ε	0.0146 [0.0003]	0.0149	0.0139	Mean idio. productivity growth rate
σ_ε	0.0496 [0.0001]	0.0416	0.0541	Parameter of idiosyncratic volatility function
g_x	0.0030 [0.0012]	0.0011	0.0029	Mean aggregate productivity growth rate
σ_x	0.0356 [0.0001]	0.0369	0.0334	Volatility of aggregate productivity growth rate

Notes: Panel A shows calibrated parameters and Panel B shows parameters estimated via SMM with standard errors in brackets. The model is solved at a quarterly frequency. Spec-1 equals the benchmark specification. Spec-2 replaces the recursive utility function with a time-separable power utility function with a low value for the relative risk aversion parameter. Spec-3 allows for time-variation not only in the jump size, but also in the jump intensity and the volatility of Gaussian idiosyncratic shocks.

Table 4: Moments Targeted in SMM Estimation

		Data	Spec-1	Spec-2	Spec-3
A: Cross-Sectional Sales Growth Moments					
Median	*mean*	0.045	0.035	0.030	0.034
	std	0.046	0.031	0.030	0.029
IQR	*mean*	0.225	0.186	0.183	0.184
	std	0.042	0.016	0.011	0.016
	corr	-0.332	-0.332	-0.351	-0.331
Kelly	*mean*	0.046	0.075	0.077	0.070
	std	0.104	0.128	0.082	0.133
	corr	0.586	0.588	0.597	0.594
B: Cross-Sectional Investment Rate Moments					
Median	*mean*	0.142	0.126	0.125	0.125
	std	0.032	0.043	0.029	0.041
IQR	*mean*	0.207	0.256	0.255	0.253
	std	0.043	0.024	0.019	0.023
	corr	0.244	0.249	0.266	0.244
Kelly	*mean*	0.352	0.337	0.310	0.335
	std	0.104	0.200	0.135	0.191
C: Aggregate Quantity Moments					
Output Growth	*mean*	0.015	0.009	0.002	0.009
	std	0.030	0.033	0.031	0.032
Consumption Growth	*std*	0.026	0.023	0.039	0.023
Investment Growth	*std*	0.066	0.046	0.027	0.047

Notes: The table summarizes the moments used in the SMM estimation. Panels A and B contain time series statistics of cross-sectional moments. For example, the row for IQR *mean* in Panel A contains the time series mean of the cross-sectional sales growth IQR. Panel C contains time series moments of aggregate quantity growth rates. All statistics refer to annual moments, i.e. annual sales growth rates, annual investment rates, as well as annual aggregate quantity growth rates. The model parameters related to each specification are shown in Table 3.

Table 5: Consumption Growth and Asset Prices

	Spec-1	Spec-2	Spec-3
Panel A: Consumption Growth			
Autocorrelation	0.330	0.058	0.462
Skewness	-0.600	-0.440	-0.395
Kurtosis	3.541	3.323	3.411
Panel B: Returns			
Excess return on wealth	1.79%	0.07%	1.60%
Risk-free Rate	1.48%	2.06%	1.62%
Sharpe Ratio	0.573	0.062	0.543
Panel C: Misallocation			
Mean of power law coef. ζ	-1.241	-1.273	-1.614
Mean of misallocation $\mathcal{M}$	0.046	0.037	0.063
Std of misallocation $\mathcal{M}$	0.024	0.014	0.030
AC4 of misallocation $\mathcal{M}$	0.701	0.750	0.635
corr[g$\mathcal{M}$,gY]	-0.589	-0.704	-0.559
Mean of output gap $\mathcal{D}$	0.045	0.064	0.036
Std of output gap $\mathcal{D}$	0.004	0.002	0.003
AC4 of output gap $\mathcal{D}$	0.744	0.749	0.707
corr[g$\mathcal{D}$,gY]	-0.511	-0.561	-0.504

Notes: The table summarizes moments related to consumption risks and risk premia. These moments were not targeted in the SMM estimation. The model parameters related to each specification are shown in Table 3. g$\mathcal{M}$, g$\mathcal{D}$, and gY denote the four quarter log changes in the misallocation measure, the output gap, and aggregate output.