Lumpy Investment, Business Cycles, and Stimulus Policy*

Thomas Winberry

University of Chicago, Booth School of Business

November 8, 2016

Abstract

I study the implications of two key facts for aggregate investment dynamics: micro-level investment mainly occurs along the extensive margin and the real interest rate is mildly countercyclical. I build a dynamic general equilibrium model which captures these facts and find two key results. First, the elasticity of aggregate investment with respect to productivity shocks or policy stimulus is procyclical because in expansions more firms are likely to make an extensive margin investment. Second, targeting firms close to the extensive margin can substantially increase the cost effectiveness of stimulus policy.

*I am extremely grateful to Richard Rogerson, Greg Kaplan, and Esteban Rossi-Hansberg for their guidance and support. I am also indebted to Mike Golosov, Jesus Fernandez-Villaverde, Pablo Ottonello, and Christian vom Lehn for many insightful conversations. Finally, I thank Mark Aguiar, Diane Alexander, Gabe Chodorow-Reich, Oleg Itskhoki, Nobu Kiyotaki, Ben Moll, Ezra Oberfield, Chris Sims, Eric Zwick, and numerous seminar participants for useful comments at various stages of this project. Financial support from Chicago Booth is gratefully acknowledged.

1 Introduction

Aggregate investment is one of the most volatile components of GDP over the business cycle, accounting for 38% of the decline in GDP during recessions.[1] Measures to stimulate investment are therefore a key element of countercyclical fiscal policy; for example, the Bonus Depreciation Allowance, which was the main investment stimulus used in the recent crisis, cost an estimated \$100 billion per year in foregone tax revenue.[2] Evaluating these policies, as well as designing new ones going forward, requires a model consistent with the key role of aggregate investment over the business cycle.

The predictions of any such model are determined by two components of the model: how individual investment decisions are made taking prices as given, and how prices are determined in general equilibrium. Existing models in the literature do not jointly match two key facts about these components in the data. First, at the micro level investment decisions are "lumpy," i.e., occur mainly along the extensive margin.[3] Second, the real interest rate – a key component of the cost of capital – is mildly countercyclical, whereas most models predict that it is highly procyclical.[4] Taken together, these two facts point to a gap in the literature since they jointly determine how aggregate investment responds to business cycle shocks and policy stimulus.

[1] Computed as the average contribution to percentage change in GDP from BEA Table 1.1.2 during NBER recession dates, 1953-2012.

[2] Computed as the forgone tax revenue for the fiscal years 2011 and 2012 with respect to 100% bonus depreciation, estimated by the White House in http://www.whitehouse.gov/sites/default/files/fact_sheet_expensing_9-8-10.pdf. The estimated cost over ten years is only \$30 billion, reflecting the fact that the bonus defers tax payments to the future.

[3] Lumpy micro-level investment has been extensively documented in, for example, Doms and Dunne (1998), Cooper and Haltiwanger (2006), or Gourio and Kashyap (2007).

[4] Although less well known than lumpy investment, this fact has also been extensively documented in the literature; see Beaudry and Guay (1996), King and Watson (1996), King and Rebelo (1999), or Cooper and Willis (2014).

I fill this gap in the literature by developing a dynamic general equilibrium model that matches both micro-level lumpiness of investment and the macro-level cyclicality of the real interest rate. The model predicts that the elasticity of aggregate investment with respect to shocks is procyclical; in expansions, more firms than average are close to making an extensive margin investment, so additional shocks induce more total investment. This mechanism also implies that aggregate investment is less responsive to stimulus policies in recessions. However, policymakers can substantially increase the cost effectiveness of these policies by targeting firms close to the extensive margin, and I develop a size-dependent policy which implements this idea in a simple way.

In order to emphasize the role of lumpy investment and realistic interest rate dynamics, the model is a direct extension of the real business cycle framework.[5] To generate lumpy investment, I assume there are heterogeneous firms who invest subject to a fixed capital adjustment cost. To generate realistic interest rate dynamics, I assume the household's preferences feature habit formation over consumption, which makes capital supply more responsive to shocks.[6] I empirically discipline these two ingredients by matching key features of micro investment real interest rate data.

Quantitatively, the calibrated model predicts that aggregate investment is up to 35% more responsive to a productivity shock in a brisk expansion than in a similarly deep recession. This procyclical elasticity is in contrast to nearly

[5]Because of its approximate linearity, the real business cycle model is a natural benchmark against which to compare the nonlinearities generated by my model. These comparisons would apply equally well to any approximately linear business cycle model at the cost of additional complications not central to the analysis.

[6]Throughout the paper I focus on the real interest rate as the key cyclical component of the cost of capital, but in principle the cyclicality of risk premia should also enter into this calculation. Empirically, risk premia are also countercyclical, so including these movements would strengthen my conclusions. However, generating movements in risk premia is a conceptually and computationally challenging particularly in a heterogeneous firm environment, so I leave it to future work.

linear models – including both real business cycle and New Keynesian models – which instead predict a constant elasticity. This additional flexibility allows the model to match the procyclical volatility in the aggregate investment rate time series recently documented by Bachmann, Caballero, and Engel (2013); in expansions, the elasticity is high so underlying shocks generate more volatility in the time series.

Jointly matching both micro-level investment lumpiness and macro-level interest rate dynamics is important in generating this procyclical elasticity. As described above, lumpiness at the micro level implies that aggregate dynamics depend on the distribution of firms relative to their adjustment thresholds. But this force alone is not enough; as Thomas (2002) and Khan and Thomas (2003, 2008) forcefully demonstrate, the quantitative strength of this mechanism is sensitive to general equilibrium movements in the real interest rate. In their model, highly procyclical movements in the real interest rate increase the cost of capital in expansions and choke off investment demand; in my model, the cost of capital falls in expansions and this choking off does not occur. Building on Khan and Thomas' insight that equilibrium price movements are important to take into account, this result shows that the dynamics of the real interest rate place sharp discipline on the exact specification of equilibrium.

The same mechanism also implies that the aggregate effect of stimulus policy falls by more than 15% in severe recessions.[7] More generally, the extent of this decline depends on the severity of the recession. Forecasts based on linear models, which imply a constant policy elasticity, will therefore potentially

[7] The policy analysis in this paper is purely positive; the goal is simply to understand how matching micro-level lumpiness and macro-level interest rate dynamics matters for these policies. Carefully characterizing the normative implications of these policies would require specifying the market failure they address, such as an upward-sloping aggregate supply curve due to nominal rigidities. The results presented here would be an important input into such an exercise.

overstate the effectiveness of stimulus policy in severe recessions.

Finally, I use the model to develop a simple size-dependent policy which increases cost effectiveness up to five times compared to existing policies. The main insight of this alternative policy is to avoid subsidizing inframarginal investment that would have been done even without the policy; because investment is lumpy, most of this inframarginal investment is accounted for by subsidizing firms who make an extensive margin investment even without the policy. In the model, small firms grow faster than average and are therefore more likely to be inframarginal to the policy.

A key challenge throughout the analysis is efficiently computing the equilibrium of the model, which involves the entire cross-sectional distribution of individual firms. The standard approach in the literature, following Krusell and Smith (1998), is to approximate the distribution with a small number of moments. This strategy places sharp restrictions on how the distribution can affect aggregate dynamics, a centerpiece of the analysis. I instead use the methodology developed concurrently in Winberry (2016), which allows for an approximation of the entire distribution.

Related Literature This paper relates to three main strands of literature. First, it contributes to a long-standing question of how micro-level lumpy investment matters for aggregate dynamics. Early papers working in partial equilibrium find that lumpy investment generates a time-varying aggregate elasticity, as in my model.[8] However, Thomas (2002) and Khan and Thomas (2003, 2008) show that when prices are endogenized in an otherwise standard real business cycle framework this time-varying elasticity disappears, rendering

[8]See, for example, Caballero, Engel, and Haltiwanger (1995), Caballero and Engel (1999), or Cooper and Haltiwanger (2006).

lumpy investment irrelevant for aggregate dynamics.[9] I show that a different specification of general equilibrium, disciplined by data on real interest rate dynamics, once again generates a time-varying elasticity.[10]

To match the dynamics of the real interest rate, I follow Beaudry and Guay (1996) in using habit formation and capital adjustment costs. These features have also been used to match the level of the equity premium in production models such as Jermann (1998) and Boldrin, Christiano, and Fisher (2001). All these papers work in a representative agent environment; my results show that many of their lessons carry over to a heterogeneous firm environment as well.

Finally, this paper contributes to a large literature which studies investment stimulus policy. Most recently in House and Shapiro (2008) and Zwick and Mahon (2016), many papers estimate the effect of policy through linear user cost or tax-adjusted q models, ruling out state-dependence by construction. Edge and Rudd (2011) introduce the Bonus Depreciation Allowance into a linearized New Keynesian model, again ruling out aggregate state dependence.[11]

[9] In this spirit, House (2008) argues that even with fixed costs, the intertemporal elasticity of investment timing is infinite, implying that the capital demand curve is flat and therefore irrelevant for aggregate dynamic. His model includes many simplifying assumptions, such as no depreciation and no idisoyncratic shocks, which are not satisfied in my model. Miao and Wang (2014a, 2014b) study the aggregate implications of lumpy investment, and how it shapes the response to stimulus policy, in a model with constant returns to scale. This assumption implies that their model "rules out distributional dynamics and cannot address distributional asymmetry and nonlinearity" which are the focus of my analysis.

[10] Other papers challenge the irrelevance results on other grounds; see, for example, Gourio and Kashyap (2007), Bachmann and Ma (2016), or Bachmann, Caballero, and Engel (2013). Bridging the gap between partial and general equilibrium approaches, Cooper and Willis (2014) parameterize an interest rate process from the data and solve the firms' problems given this process. My paper produces such an interest rate process endogenously in general equilibrium.

[11] Berger and Vavra (2015) analyze a related class of consumer durable stimulus policies in a model of lumpy durable investment. They find that stimulus policies are less effective in recessions for similar reasons as here; however, they focus detailed features of the micro

Road Map The rest of the paper is organized as follows. I describe the model and solution method in Section 2. I then document the empirical properties of micro-level lumpy investment and macro-level interest rate dynamics central to the paper and calibrate the model to match them in Section 3. In Section 4, I show that these two features generate a procyclical elasticity of aggregate investment with respect to shocks and discuss the role of the two key facts in generating this. In Section 5, I introduce stimulus policy into the model, show that the effect of the policy is state-dependent, and develop my micro-targeting proposal. Section 6 concludes.

2 Model

In this section I extend the benchmark real business cycle model to incorporate lumpy investment and realistic interest rate dynamics.

2.1 Environment

The model is a version of the neoclassical growth model in discrete time.

Firms The firm side of the model builds heavily on Khan and Thomas (2008), extended to include the corporate tax code. There is a fixed mass of firms $j \in [0, 1]$ who produce output y_{jt} using the production function

$$y_{jt} = e^{z_t} e^{\varepsilon_{jt}} k_{jt}^{\theta} n_{jt}^{\nu},\ \theta + \nu < 1$$

data while I focus on the role of real interest rate dynamics in aggregation and designing more cost effective policies.

where z_t and ε_{jt} are productivity shocks, k_{jt} is capital, n_{jt} is labor, and θ and ν are parameters. z_t is an aggregate shock which is common to all firms and drives business cycle fluctuations. It follows the AR(1) process

$$z_{t+1} = \rho_z z_t + \omega_{t+1}^z, \text{ where } \omega_{t+1}^z \sim N(0, \sigma_z^2).$$

ε_{jt} is an idiosyncratic shock which generates heterogeneity in investment patterns across firms and time. It is independent across firms but within firm follows the AR(1) process

$$\varepsilon_{jt+1} = \rho_\varepsilon \varepsilon_{jt} + \omega_{jt+1}^\varepsilon, \text{ where } \omega_{jt+1}^\varepsilon \sim N(0, \sigma_\varepsilon^2).$$

Each period, a firm j observes these two shocks, uses its pre-existing capital stock, hires labor from a competitive market, and produces output.

After production, the firm decides how much capital to invest in for the next period. Gross investment of firm j in period t, i_{jt}, yields $k_{jt+1} = (1-\delta)k_{jt} + i_{jt}$ units of capital in period $t+1$. This investment is subject to two capital adjustment costs. First, if $i_{jt} \notin [-ak_{jt}, ak_{jt}]$ the firm must pay a fixed cost ξ_{jt} in units of labor, generating an extensive margin decision for the firm. The parameter a captures the idea that small maintenance investments do not incur the fixed cost. The fixed cost ξ_{jt} is a random variable distributed uniformly over $\left[0, \overline{\xi}\right]$, and is independent across firms and over time. The second adjustment cost is $-\frac{\phi}{2}\left(\frac{i_{jt}}{k_{jt}}\right)^2 k_{jt}$ in units of output, which captures costs which are increasing in the amount of investment.[12]

After production and investment, the firm pays a linear tax τ on its revenue y_{jt} net of two deductions. First, the firm deducts its labor costs $w_t n_{jt}$ where

[12] Abel and Eberly (1994) carefully analyze a related adjustment cost function with both fixed and convex components in partial equilibrium.

w_t is the real wage in period t. Second, it deducts capital depreciation costs according to the following geometric schedule. The firm enters the period with a pre-existing stock of depreciation allowances, d_{jt}. It writes off a constant fraction $\widehat{\delta}$ of this stock d_{jt}, as well as the same fraction $\widehat{\delta}$ of new investment, i_{jt}. The remaining portion is then carried into the next period, so that $d_{j+1} = (1-\widehat{\delta})\,(d_{jt}+i_{jt})$. In total, the tax bill in period t is

$$\tau \times \left(y_{jt} - w_t n_{jt} - \widehat{\delta}\,(d_{jt}+i_{jt})\right).$$

I include the tax code to analyze investment stimulus policy in Section 5.

Households There is a representative household with preferences represented by the expected utility function

$$\mathbb{E}\sum_{t=0}^{\infty}\beta^t \log\left(C_t - H_t - \chi\frac{N_t^{1+\eta}}{1+\eta}\right),$$

where C_t is consumption, H_t is habit stock, and N_t is labor supplied to the market. I define the habit stock H_t to capture the idea that utility of current consumption is judged relative to past consumption. Specifically, following Campbell and Cochrane (1999), I first define the surplus consumption ratio $S_t = \frac{C_t - H_t}{C_t}$ and then specify the law of motion

$$\log S_{t+1} = (1-\rho_S)\log\overline{S} + \rho_S \log S_t + \lambda \log\frac{C_{t+1}}{C_t}, \tag{1}$$

which implies that current habit is approximately a geometric average of past consumption. I assume that the household does not take into account the fact that their choice of consumption impacts the habit stock. The total time endowment per period is 1, so that $N_t \in [0,1]$. The household owns all firms in the economy and markets are complete.

Government The government collects the corporate profits tax and transfers the proceeds lump sum to the household. In period t, this transfer is

$$T_t = \tau \left(Y_t - w_t N_t - \widehat{\delta} \left(D_t + I_t \right) \right), \tag{2}$$

where Y_t is aggregate output, N_t aggregate labor input, D_t aggregate stock of depreciation allowances, and I_t is aggregate investment.

2.2 Firm Optimization

I characterize the optimization problem of a firm recursively. The firm's individual state variables are ε_{jt}, its current draw of the idiosyncratic productivity shock, k_{jt}, its pre-existing stock of capital, d_{jt}, its pre-existing stock of depreciation allowances, and ξ_{jt}, its current draw of the fixed cost. The aggregate state vector is denoted $\mathbf{s}_t$ and determines prices which firms take as given. I postpone discussion of the elements in $\mathbf{s}_t$ until I define the recursive competitive equilibrium in Section 2.4.

The firm's value function, $v(\varepsilon, k, d, \xi; \mathbf{s})$, solves the Bellman equation

$$\begin{aligned} v(\varepsilon, k, d, \xi; \mathbf{s}) = \tau \widehat{\delta} d + \max_n \left\{ (1 - \tau) \left(e^z e^\varepsilon k^\theta n^\nu - w(\mathbf{s}) n \right) \right\} \\ + \max \left\{ v^a(\varepsilon, k, d; \mathbf{s}) - \xi w(\mathbf{s}), v^n(\varepsilon, k, d; \mathbf{s}) \right\}. \end{aligned} \tag{3}$$

The first max operator represents the optimal choice of labor and the second max operator represents the optimal choice of investment. These two choices are independent because the choice of labor is a purely static problem.

If the firm chooses to pay its fixed cost $-\xi w(\mathbf{s})$, it achieves the choice-specific value function $v^a(\varepsilon, k, d; \mathbf{s})$, defined by the Bellman equation:

$$v^a(\varepsilon, k, d; \mathbf{s}) = \max_{i \in \mathbb{R}} -\left(1 - \tau\widehat{\delta}\right) i - \frac{\phi}{2}\left(\frac{i}{k}\right)^2 k + \mathbb{E}[\Lambda(z'; \mathbf{s}) v(\varepsilon', k', d', \xi'; \mathbf{s}') | \varepsilon, k, d] \tag{4}$$

$$\text{s.t. } k' = (1-\delta)k + i \text{ and } d' = \left(1 - \widehat{\delta}\right)(d + i),$$

where $\Lambda(z'; \mathbf{s})$ is the stochastic discount factor. The implied "target" capital stock $k^a(\varepsilon, k, d; \mathbf{s}) = (1-\delta)k + i^a(\varepsilon, k, d; \mathbf{s})$ is what firms would like to adjust to absent the fixed cost.

If the firm chooses not to pay its fixed cost, it achieves the choice-specific value function $v^n(\varepsilon, k, d; \mathbf{s})$, defined by the Bellman equation:

$$v^n(\varepsilon, k, d; \mathbf{s}) = \max_{i \in [-ak, ak]} -\left(1 - \tau\widehat{\delta}\right) i - \frac{\phi}{2}\left(\frac{i}{k}\right)^2 k + \mathbb{E}[\Lambda(z'; \mathbf{s}) v(\varepsilon', k', d', \xi'; \mathbf{s}') | \varepsilon, k, d] \tag{5}$$

$$\text{s.t. } k' = (1-\delta)k + i \text{ and } d' = \left(1 - \widehat{\delta}\right)(d + i).$$

The only difference from the unconstrained Bellman equation (4) is that investment is constrained to be in the set $[-ak, ak]$. I call the implied capital stock, $k^n(\varepsilon, k, d; \mathbf{s}) = (1-\delta)k + i^n(\varepsilon, k, d; \mathbf{s})$, the constrained capital stock because firms face the constrained choice set $[-ak, ak]$.

The firm will choose to pay the fixed cost if and only if the value from doing so is higher than not paying the fixed cost, i.e., if and only if $v^a(\varepsilon, k, d; \mathbf{s}) - \xi w(\mathbf{s}) \geq v^n(\varepsilon, k, d; \mathbf{s})$. For each $(\varepsilon, k, d; \mathbf{s})$, there is a unique threshold $\widehat{\xi}(\varepsilon, k, d; \mathbf{s})$ which makes the firm indifferent between these two options. This threshold solves

$$\widehat{\xi}(\varepsilon, k, d; \mathbf{s}) = \frac{v^a(\varepsilon, k, d; \mathbf{s}) - v^n(\varepsilon, k, d; \mathbf{s})}{w(\mathbf{s})}. \tag{6}$$

For draws of the fixed cost ξ below $\widehat{\xi}(\varepsilon, k, d; \mathbf{s})$, the firm pays the fixed cost; for draws of the fixed cost above $\widehat{\xi}(\varepsilon, k, d; \mathbf{s})$, it does not. This threshold is increasing in the "capital imbalance" $|k^a(\varepsilon, k, d; \mathbf{s}) - k^n(\varepsilon, k, d; \mathbf{s})|$ since the gain from adjusting is higher when the target capital stock is further away from the constrained capital stock. The firms' optimal choice to only pay the fixed cost infrequently generates lumpy investment patterns as in the data.

2.3 Household Optimization

Since investment is done by firms, there are no dynamic links in the household's choices and the decision problem is equivalent to the following static problem state by state:

$$\max_{C,N} \ \log\left(C - H(\mathbf{s}) - \frac{N^{1+\eta}}{1+\eta}\right) \text{ subject to } C \leq w(\mathbf{s})N + \Pi(\mathbf{s}) + T(\mathbf{s}). \quad (7)$$

The household chooses consumption and labor supply to maximize its period utility, subject to the budget constraint. Total expenditure is consumption C. Total income is labor income, $w(\mathbf{s})N$, profits from owning the firms $\Pi(\mathbf{s})$, and the lump sum transfer from the government, $T(\mathbf{s})$.

Although households do not make investment decisions themselves, market completeness implies that the stochastic discount factor used by firms to price investment is equal to the household's intertemporal marginal rate of substitution state by state:

$$\Lambda(z'; \mathbf{s}) = \frac{C(\mathbf{s}) \times S(\mathbf{s}) - \chi\frac{N(\mathbf{s})^{1+\eta}}{1+\eta}}{C(\mathbf{s}') \times S(\mathbf{s}') - \chi\frac{N(\mathbf{s}')^{1+\eta}}{1+\eta}}. \quad (8)$$

These preferences, along with firms' capital adjustment costs, allow the model to generate a countercyclical real interest rate. Intuitively, the cyclicality of the interest rate is determined by the responsiveness of capital supply and demand to an aggregate productivity shock. Habit formation makes capital supply sensitive to shocks – households want smoother consumption profiles, so they save more of the income generated by the shock – while the adjustment costs make capital demand less responsive – firms find it more costly to accumulate capital. The calibration will formally discipline the strength of these two forces using the empirical behavior of the real interest rate.

2.4 Equilibrium

To define the recursive competitive equilibrium, I use the aggregate state $\mathbf{s} = (z, S_{-1}, C_{-1}, \mu)$, where z is the aggregate productivity shock, C_{-1} is previous period's consumption, and μ is the distribution of firms over their individual state vector (ε, k, ξ).

***Definition* 1.** *A **Recursive Competitive Equilibrium** for this economy is a list of functions $v(\varepsilon, k, d, \xi; \mathbf{s})$, $n(\varepsilon, k, d, \xi; \mathbf{s})$, $i^a(\varepsilon, k, d; \mathbf{s})$, $i^n(\varepsilon, k, d; \mathbf{s})$, $\widehat{\xi}(\varepsilon, k, d; \mathbf{s})$, $C(\mathbf{s})$, $N(\mathbf{s})$, $T(\mathbf{s})$, $w(\mathbf{s})$, $\Pi(\mathbf{s})$, $\Lambda(z'; \mathbf{s})$, $S'_{-1}(\mathbf{s})$, $C'_{-1}(\mathbf{s})$, and $\mu'(\mathbf{s})$ such that*

(i) (Household Optimization) Taking $w(\mathbf{s})$, $\Pi(\mathbf{s})$, and $T(\mathbf{s})$ as given, $C(\mathbf{s})$ and $N(\mathbf{s})$ solve the utility maximization problem (7).

(ii) (Firm Optimization) Taking $w(\mathbf{s})$, $\Lambda(z'; \mathbf{s})$, $C'_{-1}(\mathbf{s})$, and $\mu'(\mathbf{s})$ as given, $v(\varepsilon, k, \xi; \mathbf{s})$, $n(\varepsilon, k, \xi; \mathbf{s})$, $i^a(\varepsilon, k, \xi; \mathbf{s})$, $i^n(\varepsilon, k, \xi; \mathbf{s})$ and $\widehat{\xi}(\varepsilon, k; \mathbf{s})$ solve the firm's maximization problem (3) - (6).

(iii) (Government) For all $\mathbf{s}$, $T(\mathbf{s})$ is given by (2).

(iv) (Consistency) For all $\mathbf{s}$,

(a) $\Pi(\mathbf{s}) = \int [\, (1-\tau)\left(e^z e^\varepsilon k^\theta n\left(\varepsilon,k,d,\xi;\mathbf{s}\right)^\nu - w(\mathbf{s}) n\left(\varepsilon,k,d,\xi;\mathbf{s}\right)\right) + \tau\widehat{\delta} d$
$- \left(1-\tau\widehat{\delta}\right) i(\varepsilon,k,d,\xi;\mathbf{s}) - \frac{\phi}{2}\left(\frac{i(\varepsilon,k,d,\xi;\mathbf{s})}{k}\right)^2 k - \xi w(\mathbf{s}) 1\left\{\frac{i(\varepsilon,k,d,\xi;\mathbf{s})}{k} \notin [-a,a]\right\}]$
$\mu(d\varepsilon, dk, dd, d\xi)$, *where* $i\left(\varepsilon,k,d,\xi;\mathbf{s}\right) = i^a\left(\varepsilon,k,d,\xi;\mathbf{s}\right)$ *if* $\xi \leq \widehat{\xi}\left(\varepsilon,k,d;\mathbf{s}\right)$ *and* $i\left(\varepsilon,k,d,\xi;\mathbf{s}\right) = i^n\left(\varepsilon,k,d,\xi;\mathbf{s}\right)$ *otherwise.*

(b) $\Lambda(z';\mathbf{s})$ *is given by (8).*

(c) $S'_{-1}(\mathbf{s})$ *follows (1).*

(d) $C'_{-1}(\mathbf{s}) = C(\mathbf{s})$.

(e) *For all measurable sets* $\Delta_\varepsilon \times \Delta_k \times \Delta_d \times \Delta_\xi$, $\mu'(\Delta_\varepsilon \times \Delta_k \times \Delta_d \times \Delta_\xi)$ $= \int p(\varepsilon' \in \Delta_\varepsilon|\varepsilon) d\varepsilon' \times 1\{\ i(\varepsilon,k,d,\xi;\mathbf{s}) + (1-\delta)k \in \Delta_k\ \} \times 1\{$ $(1-\widehat{\delta})\ (i(\varepsilon,k,d,\xi;\mathbf{s}) + d) \in \Delta_d\ \} \times G(\Delta_\xi) \times \mu(d\varepsilon, dk, dd, d\xi)$, *where* $G(\xi)$ *is the CDF of* ξ.

(v) *(Market Clearing) For all* $\mathbf{s}$, $N(\mathbf{s}) = \int n\left(\varepsilon,k,d,\xi;\mathbf{s}\right)\mu(d\varepsilon, dk, dd, d\xi)$.

The mapping in Condition iv(e) defines the measure of firms in the set $\Delta_\varepsilon \times \Delta_k \times \Delta_d \times \Delta_\xi$ next period in terms of the distribution of firms and individual decisions in the current period. Intuitively, this mapping counts up the mass of individual states in the current period which leads into the set $\Delta_\varepsilon \times \Delta_k \times \Delta_d \times \Delta_\xi$ next period. The mass of firms in Δ_ε is determined by the mass of firms who had a particular draw of ε and a draw of the innovation ω'_ε such that $\rho_\varepsilon \varepsilon + \omega'_\varepsilon \in \Delta_\varepsilon$. The mass of firms in Δ_k are those firms whose investment policy leads to capital in that set, i.e., $(1-\delta)\,k + i\left(\varepsilon,k,d,\xi;\mathbf{s}\right) \in \Delta_k$. The mass of firms in Δ_d are those for whom $(1-\widehat{\delta})\ (i(\varepsilon,k,d,\xi;\mathbf{s}) + d) \in \Delta_d$. Finally, the mass of firms with fixed cost in Δ_ξ is simply $G(\Delta_\xi)$, since ξ is i.i.d. over firms and time.

2.5 Solution Method

The fact that the aggregate state vector **s** contains the entire cross-sectional distribution of firms is a key challenge in solving the model. The standard approach, following Krusell and Smith (1998), is to approximate the distribution with a small number of moments. This places sharp restrictions on how the distribution can impact aggregate dynamics which are at the center of my analysis.[13]

I solve the model using a method, developed concurrently in Winberry (2016), which includes the entire distribution in the aggregate state vector. To do that, I approximate the cross-sectional distribution of firms at any point in time using a flexible, finite-dimensional parametric family. A good approximation of the distribution may require many parameters, leaving globally accurate approximation techniques infeasible due to the curse of dimensionality. I therefore solve for the dynamics of the distribution using locally accurate perturbation methods. See Appendix A.3 for details on the implementation.

3 Empirical Targets and Model Calibration

The goal of the model is to show that jointly matching micro-level lumpy investment and macro-level interest rate dynamics is important to understanding aggregate investment dynamics. In this section, I document those two features of the data and calibrate key model parameters to match those features.

[13] One can in principle add enough moments of the distribution to capture its relevant features for aggregate dynamics, but this is subject to the curse of dimensionality.

3.1 Empirical Targets

Micro-Level Lumpy Investment The lumpiness of investment at the micro level is a well-known fact documented in the Longitudinal Research Database (LRD) sample of Census manufacturing firms. For example, Doms and Dunne (1998) show that up 40% of an average plant's total investment over a sixteen year period can be accounted for by one large investment project, and outside this large spike plants invest relatively little. Cooper and Haltiwanger (2006), among others, estimate that nonconvex adjustment costs are necessary to match micro investment behavior.

Instead of using LRD data, I calibrate my model to match moments drawn from a dataset drawn from annual IRS corporate income tax returns, reported in Zwick and Mahon (2016) Appendix B. The key advantage of this dataset is that it covers all sectors of the economy, not just manufacturing, and so allows for a more representative sample than previous studies.[14]

Table I show that the IRS sample features significant micro-level lumpiness, in line with the previous literature's findings in Census data. Following Cooper and Haltiwanger (2006), the table focuses on the distribution of investment rates pooled over firms and time.[15] A significant share of observations, about one fourth of the sample, have essentially zero investment; at the same time, about one-fifth of the sample have large investment rate spikes greater than

[14]A disadvantage of the IRS data for studying lumpy investment is that measured investment includes mainly equipment goods while measured capital includes both equipment and structures. This mismatch potentially biases measured investment rates down to the extent that the denominator includes more goods than the numerator. Despite these limitations, I prefer to work with the IRS data for the main text since it provides a comprehensive sample of firms.

[15]Another difference between the IRS and Census datasets is that the IRS data is recorded at the firm level while Census is at the plant level. The firm is the appropriate unit of analysis for this study given my focus on investment stimulus policy, which operates by changing a firm's tax incentives.

Table I
Micro-Level Lumpy Investment

Statistic	**Value**
Inaction rate (%)	23.7%
Spike rate (%)	14.4%
Positive investment rates (%)	61.9%
Average investment rate (%)	10.4%
Standard deviation of investment rates	0.160

Notes: Micro investment moments from annual firm-level IRS data, 1998 - 2010, as reported in Zwick and Mahon (2016) Appendix Table B.1. Statistics drawn from distribution of investment rates pooled over firms and time. Inaction rate is fraction of observations with investment rate less than 1%. Spike rate is fraction of observations with investment rate greater than 20%. Positive investment is fraction of observations between 1% and 20%.

20% annually. Cooper and Haltiwanger (2006) show the coexistence of inaction and spikes is consistent with nonconvex adjustment costs.

Real Interest Rate Dynamics Starting with Summers (1986), a common early criticism of the real business cycle framework is that it predicts a highly procyclical real interest rate, while in the data interest rates are typically countercyclical. King and Rebelo (1999) note this fact in their survey of real business cycle models and Beaudry and Guay (1996) show that it is also true conditional on fluctuations driven by productivity shocks.

The counterfactual performance of the RBC model for real interest rates is potentially problematic for studying aggregate investment given the tight link between the interest rate and the stochastic discount factor firms use to price their investment decisions:

$$\mathbb{E}_t[\Lambda_{t+1} \times R_{t+1}] = 1, \tag{9}$$

TABLE II
MACRO-LEVEL REAL INTEREST RATE DYNAMICS

	$\sigma(r_t)$	$\rho(r_t, y_{t-2})$	$\rho(r_t, y_{t-1})$	$\rho(r_t, y_t)$	$\rho(r_t, y_{t+1})$	$\rho(r_t, y_{t+2})$
T-bill	2.18%	-0.01	-0.08	-0.17	-0.25	-0.31
AAA	2.34%	-0.21	-0.29	–0.37	-0.40	-0.38
BAA	2.43%	-0.22	-0.32	-0.41	-0.45	-0.42
Stock	24.7%	-0.22	-0.24	-0.14	0.02	0.16
RBC	0.16%	0.32	0.61	0.97	0.74	0.54

Notes: Cyclical properties of various measures of the real interest rate. "T-bill" refers to 90-day treasury bill rate, "AAA" to Moody's Seasoned AAA Corporate Bond Yield, "BAA" to Moody's Seasoned BAA Corporate Bond Yield, and "Stock" to the return on the Russell 3000 stock index, each expressed in annual percentage points. I construct real rates by subtracting realized CPI inflation. r_t refers to the particular real interest rate and y_t to the log of real GDP, and both have been HP-filtered with smoothing parameter 1600. Data is quarterly, 1947q1 - 2016q2. "RBC" refers to risk-free rate in a benchmark RBC model.

where Λ_{t+1} is the stochastic discount factor between periods t and $t+1$ and R_{t+1} is the rate of return. Equation (9) shows that when interest rates rise, the stochastic discount factor falls in expectation, implying that firms value future profits – and therefore current investment – less.

Table II documents the countercyclicality of four key real interest rates: the 90-day Treasury bill, the yield on AAA rated corporate bonds, the yield on BAA rated corporate bonds, and the return on the aggregate stock market. Each of these rates provides a different measure of the return to capital priced by households through (9). Consistent with the early criticisms, each of these interest rates is negatively correlated with output contemporaneously. Furthermore, also consistent with that literature, interest rates tend to be a leading indicator of the business cycle in the sense that they are negative correlated with future output. In contrast, a benchmark real business cycle model predicts that the interest rate is highly positively correlated with output at each horizon, reaching up to 0.97 contemporaneously. In the calibration, I will focus on the contemporaneous correlation of the real interest rate and output.

Appendix A.1 shows that the empirical countercyclicality of these interest rates is a robust fact. First, it survives using alternative methods of filtering the data (a linear trend, a bandpass filter, or simply comparing the raw rates to log GDP growth). Second, it is robust to using different price indices to compute the inflation rate (the GDP deflator, PCE deflator, or nonresidential fixed investment goods deflator). Finally, it is robust to using ex-ante inflation to compute expectations rather than ex-post realizations.

3.2 Model Calibration

I calibrate the model in two steps. First, I exogenously fix a set of parameters to match standard macroeconomic targets. Given those parameters, I then choose the remaining parameters to match moments of micro-level investment and macro level real interest rate dynamics reported in Section 3.1.

Fixed Parameters Table III lists the parameters I fix exogenously. A model period is one quarter. I set the discount factor $\beta = 0.99$ so that the steady state annual real interest rate is 4%. I set the Frisch elasticity of labor supply to 2, within the range of macro elasticities identified by Chetty et al. (2011). I set the labor share $\theta = 0.64$ and choose the capital share so that the total returns to scale is 85%. The returns to scale lies within the range considered in the current literature, from 60% in Gourio and Kashyap (2007) to 92% in Khan and Thomas (2008). I set $\delta = 0.025$ so that the steady state aggregate investment rate is 10%, roughly in line with the average in the postwar data. I set the stochastic process for TFP to $\rho_z = 0.95$ and $\sigma_z = 0.007$ as in King

Table III
Fixed Parameter Values

Parameter	**Description**	**Value**
β	Discount factor	.99
η	Inverse Frisch	$\frac{1}{2}$
θ	Labor share	.64
ν	Capital share	.21
δ	Capital depreciation	.025
ρ_z	Aggregate TFP AR(1)	.95
σ_z	Aggregate TFP AR(1)	.007
τ	Tax rate	.35
$\widehat{\delta}$	Tax depreciation	.119

Notes: Parameters fixed exogenously in calibration.

and Rebelo (1999).

I set the tax rate $\tau = .35$ to match the top marginal tax rate in the federal corporate income tax code. I then choose the tax depreciation schedule $\widehat{\delta}$ to reproduce the average present value of depreciation allowances per unit of investment documented by Zwick and Mahon (2016). I show in Appendix A.2 that this present value completely summarizes the impact of the tax depreciation schedule on firms' investment decisions.

Fitted Parameters I choose the remaining parameters – governing micro heterogeneity on the firm side and habit formation on the household side – to match the moments in Table IV. The micro investment targets are drawn from Zwick and Mahon (2016) Appendix B as described in Section 3.1. The interest rate targets are drawn from the dynamics of the 90-day Treasury bill adjusted for realized inflation.[16] There are presumably many shocks driving

[16] I use the 90-day Treasury bill rate, which I identify as the risk-free rate in the model, for two reasons. First, it cleanly maps into movements in the conditional expectation of the

fluctuations in the interest rates reported in Table II while in the model the only aggregate shock is productivity. I project the interest rate on ten lags of the level and square of measured TFP as a simple way to extract the portion of empirical interest rate fluctuations driven by productivity shocks. I then extract a Hodrick-Prescott filter of these fitted values to focus on the business cycle component. I target the overall volatility of the interest rate and its correlation with output, which has also been projected on measured TFP and HP filtered. The resulting cyclicality of the real interest rate is similar to the raw series but the overall volatility is one fourth the size.

The calibrated model fits the moments targeted in Table IV well. Importantly, it captures the fraction of observations with nearly zero investment as well as the fraction of observation with large investment rates. The fit to the interest rate dynamics is nearly exact. However, the dispersion of investment rates is low relative to the data.

Table V shows that the calibrated parameter values are broadly comparable to previous findings in the literature. The upper bound on the fixed cost $\overline{\xi}$ is within the wide range of 0.0083 in Khan and Thomas (2008) and 4.4 in Bachmann, Caballero, and Engel (2013). The calibrated value implies that the average fixed cost paid conditional on adjusting is 3.5% of the average output the firm. The average size of the surplus consumption ratio $\overline{S}$ is smaller than in Campbell and Cochrane (1999) because they target the level of the equity premium, which requires a more volatile intertemporal marginal rate of substitution.

The identification of these parameters can be understood in two broad

stochastic discount factor through $1 + r_t = \frac{1}{\mathbb{E}_t[\Lambda_{t+1}]}$. Second, in my model the risk premium is negligible so that the risk-free rate captures almost all variation in the stochastic discount factor. Given the robustness of the empirical facts documented in Section 3.1, I conjecture the results are ultimately robust to this choice.

Table IV
Moments Targeted in Calibration

Micro Investment		
Target	**Data**	**Model**
Inaction rate (%)	23.7%	23.9%
Spike rate (%)	14.4%	15.9%
Positive investment rates (%)	61.9%	60.2%
Average investment rate (%)	10.4%	10.6%
Standard deviation of investment rates	0.160	0.121
Interest Rate Dynamics		
Target	**Data**	**Model**
Standard deviation of interest rate (%)	0.48%	0.48%
Correlation of interest rate and output	-0.21	-0.20

Notes: Micro investment moments from annual firm-level IRS data, 1998 - 2010, as reported in Zwick and Mahon (2016) Appendix Table B.1. Statistics drawn from distribution of investment rates pooled over firms and time. Inaction rate is fraction of observations with investment rate less than 1%. Spike rate is fraction of observations with investment rate greater than 20%. Positive investment is fraction of observations between 1% and 20%. Interest rate dynamics correspond to the 90-day Treasury bill rate, adjusted for realized inflation, projected on ten lags of the level and square of measured TFP, and HP filtered. Output is log of real GDP, also projected on lags of TFP and HP filtered.

Table V
Fitted Parameter Values

Micro Heterogeneity		
Parameter	**Description**	**Value**
$\overline{\xi}$	Upper bound on fixed costs	0.44
a	Size of no fixed cost region	0.003
ϕ	Quadratic adjustment cost	2.69
ρ_ε	Idiosyncratic productivity AR(1)	0.94
σ_ε	Idiosyncratic productivity AR(1)	0.026
Habit Formation		
Parameter	**Description**	**Value**
$\overline{S}$	Average surplus consumption	0.65
ρ_S	Autocorrelation of surplus consumption	0.95

Notes: Parameters chosen to match moments in Table IV.

steps. First, the dynamics of the real interest rate pins down the overall strength of habit formation and adjustment costs in determining capital supply and demand in the model. As discussed in Section 2.3, the countercyclicality of the interest rate indicates that capital supply is more responsive to shocks than demand. Second, given the overall responsiveness of capital demand to shocks, the micro investment data then pins down the importance of idiosyncatic shocks and particular types of capital adjustment frictions.

4 Business Cycle Analysis

Appendix A.4 shows that my model and a benchmark real business cycle model have quantitatively similar predictions for the unconditional second moments of aggregate output, consumption, investment, and hours.[17] However, they differ in their predictions conditional on stage in the business cycle: my model predicts that aggregate investment is more responsive to shocks in expansions than recessions while the benchmark model does not. This state dependence allows my model to match the procyclical volatility in aggregate investment rate series documented by Bachmann, Caballero, and Engel (2013).

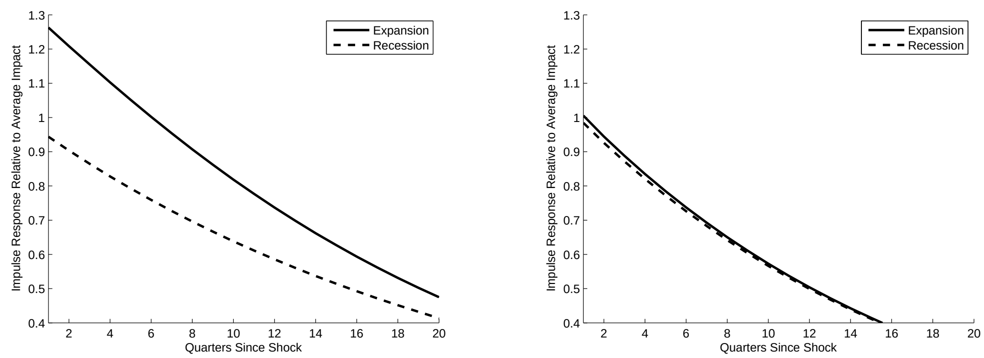
4.1 State Dependent Impulse Responses

I first illustrate the model's state dependence by comparing the response of aggregate investment to a productivity shock starting from different points in the business cycle. As an example, I feed in two different histories of aggregate shocks; one history pushes the economy into an expansion and the other into

[17]The benchmark real business cycle model can be seen as a special case of the full model in which there are no capital adjustment frictions, which implies that thee firm side aggregates to a representative firm, and no habit formation, which implies the household has standard GHH preferences.

FIGURE 1: State Dependent Impulse Response to Productivity Shock

(a) Model (b) Benchmark



Notes: Impulse response of aggregate investment to one standard deviation productivity shock as a function of the previous three quarter's shocks. Normalized so that response upon impact in steady state is 1. (a) Model from text. (b) Benchmark model without adjustment costs or habit formation.

a recession.[18] Starting from these different points I then compute the response to an additional one standard deviation positive shock.

The left panel of Figure 1 shows that the impulse response to this additional shock differs substantially starting from the expansion and recession. Starting from the expansion, the shock generates 26% more investment upon impact than it would starting from steady state; summed over the life of the shock, the total effect is 17% higher, reflecting intertemporal substitution as firms initially pull forward investment projects. In contrast, starting from the recession, the shock generates 6% less investment than it would starting from steady state; summed over the life of the shock, the total effect is also 6% less,

[18]To generate these histories of shocks, I first compute the history of aggregate productivity shocks required to reproduce the aggregate investment rate time series observed in the data 1953 - 2012. These shocks do not share exactly the same time series properties of aggregate TFP used by firms to form expectations, so I rescale to reflect the assumed stochastic process. The example expansion then corresponds to the late 1990s expansion and the recession to the early 2000s recession.

indicating intertemporal plays a smaller role. As I explain in Section 4.2, this state dependence is due to the fact that more firms find it optimal to pay their fixed cost ξ and make an extensive margin investment in the expansion than in the recession.[19]

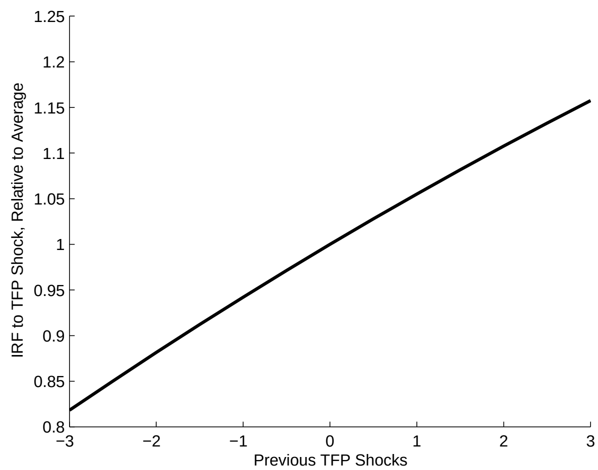
The right panel of Figure 1 shows that the benchmark real business cycle model generates virtually none of this state dependence starting from the same history of shocks. In the expansion, the effect of the additional productivity shock is only 1% larger than starting from steady state; in the recession, the effect is 1% smaller. The small degree of state dependence is due to the fact that capital is higher starting in the expansion and is complementary with productivity. However, this effect is quantitatively small, reflecting the approximate linear of the real business cycle model.

Figure 2 confirms that state dependence is a general property of my model. A one standard deviation productivity shock generates 20% less investment than average starting from a severe recession and 15% more investment than average starting from a similarly extreme expansion.

Evidence in Aggregate Data Appendix A.5 shows that my model is consistent with the conditional heteroskedasticity estimated in the aggregate investment rate time series by Bachmann, Caballero, and Engel (2013). They find the residuals of a reduced-form time series model are more volatile in expansions than recessions; my model generates this reduced-form because investment is more responsive to underlying homoskedastic shocks in expansions than recessions while the benchmark model does not. The calibrated model quantitiatively generates between 50%-100% of the heteroskedasticity in the

[19] The asymmetry starting from the expansion and recession reflects the fact that the histories of shocks are themselves state dependent; Figure 2 shows the underlying mechanism is more symmetric.

Figure 2: State Dependent Impulse Responses Over the Cycle



Notes: Impulse response upon impact to one standard deviation productivity shock, after different histories of shocks. Normalized so that response upon impact in steady state is 1.

data depending on the specification of the time series model.

4.2 Role of Two Key Facts

Jointly capturing both micro-level lumpy investment and the mild countercyclicality of the real interest rate is important in generating these state-dependent impulse responses.

Lumpy Investment The presence of fixed adjustment costs implies that the impulse response of aggregate investment to a productivity shock depends on the number of firms who make an extensive margin investment. In an expansion, more firms are close to making an extensive margin investment so aggregate investment is more responsive to additional shocks; in recessions, the opposite occurs. The mechanism underlying this asymmetry over the cycle has been discussed extensively in for example Caballero and Engel (2007) or

Bachmann, Caballero, and Engel (2013), so I keep my description brief. For simplicity, consider a version of the model without taxes ($\tau = 0$) and quadratic adjustment costs ($\phi = 0$).

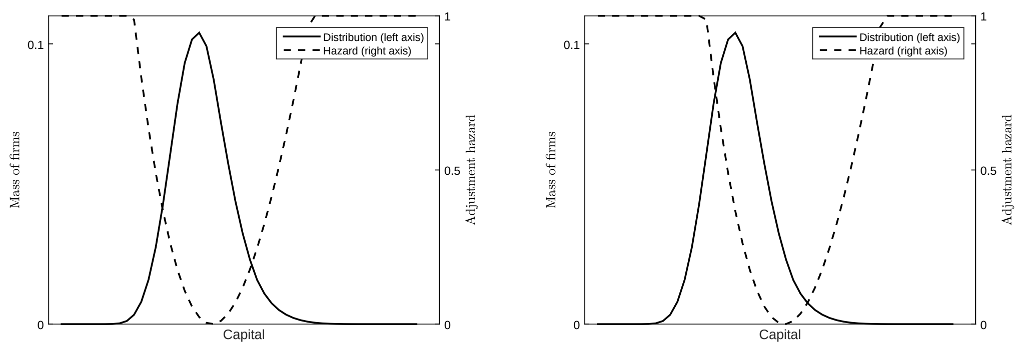
As discussed in Section 2.2, firms' investment decisions are characterized by the target capital stock $k^a(\varepsilon, k; \mathbf{s})$ it adjusts to conditional on paying the fixed cost, the constrained capital stock $k^n(\varepsilon, k; \mathbf{s})$ it adjusts to conditional on not paying the fixed cost, and the fixed cost threshold $\widehat{\xi}(\varepsilon, k; \mathbf{s})$ below which it finds it optimal to pay the fixed cost. Since ξ is i.i.d., for each value of productivity and capital (ε, k) a fraction $\widehat{\xi}(\varepsilon, k; \mathbf{s})$ of firms pay their fixed costs and adjust while the remaining fraction $1 - \widehat{\xi}(\varepsilon, k; \mathbf{s})$ do not. The function $\widehat{\xi}(\varepsilon, k; \mathbf{s})$ is referred to as the adjustment hazard because it controls the proportion of firms who adjust their capital.

The key source of state dependence in the impulse responses is that on average firms hold less capital than their target, implying $k^n(\varepsilon, k; \mathbf{s}) < k^a(\varepsilon, k; \mathbf{s})$; once firms adjust to their target, capital depreciates. A history of negative shocks, which generates a recession, will decrease the target capital stock and bring the average firm more in line with its target $k^n(\varepsilon, k; \mathbf{s}) \approx k^a(\varepsilon, k; \mathbf{s})$. In this case, firms will be close to their optimum and be relatively unwilling to adjust to additional shocks. On the other hand, a history of positive shocks which generates an expansion will increase the target capital stock implying $k^n(\varepsilon, k; \mathbf{s}) << k^a(\varepsilon, k; \mathbf{s})$. In this case, firms will be even further from their optimum and be more willing to adjust to additional shocks. Caballero and Engel (2007) emphasize the "increasing hazard" property that adjustment probabilities increase in $|k^n(\varepsilon, k; \mathbf{s}) - k^a(\varepsilon, k; \mathbf{s})|$ is the key channel through which fixed cost models potentially generate nonlinear dynamics.

Figure 3 plots this mechanism in a stylized example. The left panel plots the distribution of firms and their adjustment hazard, conditional on the mean

FIGURE 3: Illustrating the Role of Lumpy Investment

(a) Stationary Distribution (b) Expanssion



Notes: Stylized illustration of lumpy investment mechanism. “Distribution” refers to distribution of firms over capital. “Hazard” refers to probability firm draws a fixed cost below its adjustment threshold. Panel (a) plots these objects in a stationary environment, while Panel (b) plots them after a positive shock which shifts the adjustment threshold rightward.

value of idiosyncratic productivity, in steady state. The adjustment hazard is flat where $k^n(\varepsilon, k; \mathbf{s}) = k^a(\varepsilon, k; \mathbf{s})$ and increasing in $|k^n(\varepsilon, k; \mathbf{s}) - k^a(\varepsilon, k; \mathbf{s})|$. The mass of the stationary distribution is concentrated on the left side of the hazard because firms hold less capital than their target. A positive productivity shock increases the target capital stock and shifts the adjustment hazard rightward, plotted in the right panel. In this case, the distribution is concentrated in a region of the state space where the adjustment hazard is increasing.

Real Interest Rate Dynamics Although the lumpy investment mechanism described above is well-established in partial equilibrium models, Thomas (2002) and Khan and Thomas (2003, 2008) show that its quantitative strength is sensitive to general equilibrium price movements. Specifically, they endogenize prices in an otherwise standard real business cycle framework featuring a highly procyclical real interest rate and, therefore, a highly procyclical cost of capital. After a positive productivity shock, when in partial equilibrium many firms would like to undertake an extensive margin investment, the real interest rate increases and chokes off capital demand.[20] This force restrains movements in the adjustment hazards illustrated in Figure 3 and eliminates state dependence in aggregate impulse responses.

Building on Khan and Thomas' critical insight, my results illustrate how the exact specification of general equilibrium also matters for the quantitative strength of the lumpy investment mechanism. In my model, a positive productivity shock is accompanied by a fall in the real interest rate, consistent with the empirical evidence documented in Section 3.1. In this case general equilibrium does not choke off investment demand and the distributional dynamics re-emerge as a quantitatively relevant force generating state dependent impulse responses. To ensure that the overall volatility of investment rates are in line with the data, my model requires larger calibrated adjustment costs.[21]

In general equilibrium, the calibration of adjustment costs and the dynamics of the real interest rate are tightly linked through the intersection of capital supply and demand.[22] To be consistent with a countercyclical real interest rate, the capital demand curve must be less responsive to productivity shocks than the capital supply curve. The calibrated model achieves this by higher values for the fixed and convex adjustment costs than in Khan and Thomas' (2008)

[20]Recall that a higher real interest rate implies a lower discount factor through the asset pricing equation $1 + r_t = \frac{1}{\mathbb{E}_t[\Lambda_{t,t+1}]}$.

[21]Ultimately my general equilibrium model returns aggregate dynamics qualitatively consistent with previous partial equilibrium studies. However, the endogenous price process leads to quantitatively different dynamics and allows me to study policy implications in Section 5.

[22]Cooper and Willis (2014) separate these forces by estimating a price process from the data and solving firms' decision problems given that process. In this case, the fixed costs change aggregate dynamics, consistent with the results presented here.

parameterization, giving stronger quantitative kick to the lumpy investment mechanism.[23]

5 Policy Analysis

Having shown the model accounts for the dynamics of aggregate investment in recessions, I now introduce countercyclical stimulus policy into the model and analyze the model's policy implications. As with business cycle shocks, the effectiveness of policy is also state dependent and falls in recessions. However, the fact that investment is lumpy at the micro level implies that a micro-targeted policy can increase cost effectiveness up to five times compared to existing policies.

5.1 Introducing Stimulus Policy Into the Model

Historical stimulus policies in the U.S. incentivize firms to investment by increasing the effective tax writeoff for new investment. To illustrate these policies, Table VI reproduces the tax writeoff associated with a $1000 computer purchase under three tax schedules: the standard schedule, known as MACRS; the schedule under a Bonus Depreciation Allowance, the main policy used following the 2001 and 2007 recessions; and the schedule under an investment tax credit. The Bonus Depreciation Allowance allows firms to immediately write off a fraction of their investment rather than following the baseline MACRS schedule, raising the present value of tax writeoffs. Similarly, the investment tax credit reduces the firm's overall tax bill by a fraction of the investment

[23]In the language of Bachmann, Caballero, and Engel (2013), the countercyclical real interest rate eliminates a large portion of "price response smoothing" and instead loads more onto "adjustment cost" smoothing, which leads to state dependence.

Table VI
Tax Depreciation Schedule Under Different Investment Stimulus Policies

Standard MACRS Schedule (No policy)								
Year	0	1	2	3	4	5	Total	PV, 7%
Deductions	200	320	192	115	115	58	1000	890
50% Bonus Depreciation								
Year	0	1	2	3	4	5	Total	PV, 7%
Deductions	500+100	160	96	57.5	57.5	29	1000	945
5% Investment Tax Credit								
Year	0	1	2	3	4	5	Total	PV, 7%
Deductions	$\frac{50}{35\%}$+190	304	182.4	109.3	109.3	55	1093	1093

Notes: Tax depreciation schedule for purchase of $1000 computer. Top panel: standard schedule absent stimulus policy. Middle panel: 50% bonus depreciation allowance. Bottom panel: 5% investment tax credit. Present value computed using 7% discount rate.

cost.

The present value of tax depreciation schedule in these two examples is a useful summary of their impact on incentives; Proposition 1 shows that, in the model, the present value is the *only* way the tax schedule affects the incentive to invest. Firms behave as if there is a tax-adjusted price of investment $q_t = 1 - \tau \times PV_t$ where PV_t is the present value of tax writeoffs. Intuitively, once a firm has invested the tax writeoffs are simply a stream of risk-free dividends which the firm values at $\tau \times PV_t$.

***Proposition* 1.** *The only way the tax depreciation schedule affects a firm's investment policy rule is through the effective price* q_t*:*

$$\begin{aligned} q_t &= 1 - \tau \times PV_t, \text{ where} \\ PV_t &= \mathbb{E}_t \sum_{s=0}^{\infty} \underbrace{\left(\Pi_{j=0}^{s} \frac{1}{R_{t+j}} \right)}_{\text{discount rate for period } t+s} \times \underbrace{\left(1 - \widehat{\delta}\right)^s \widehat{\delta}}_{\text{deduction in period } t+s} . \end{aligned}$$

Proof. In Appendix A.2. ■

A wide class of investment stimulus policies simply map into a change in PV_t; I therefore model investment stimulus policy as an exogenous shock to this present value:

$$q_t = 1 - \tau \times (PV_t \times sub_t)$$

where sub_t is an implicit subsidy representing a general stimulus policy. For example, the Bonus Depreciation Allowance maps into an implicit subsidy $sub_t = .5 \times (1 - PV_t)$ and the investment tax credit into $sub_t = .05 \times (\frac{1}{\tau} - PV_t)$. I assume the subsidy sub_t follows the simple stochastic process

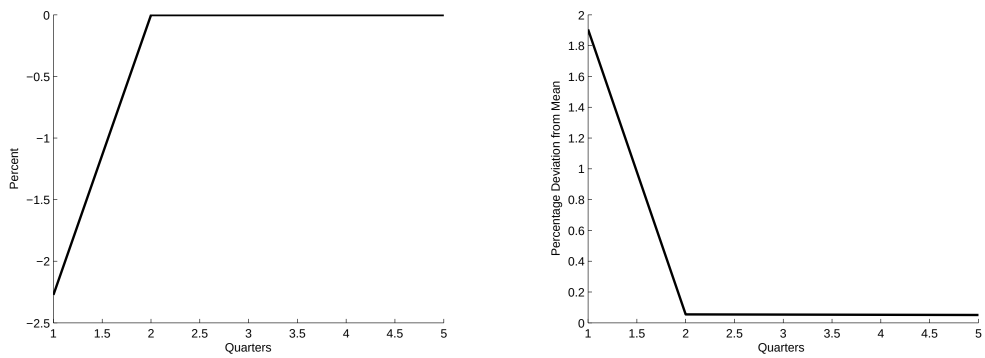
$$\log sub_t = \log \overline{sub} + \varepsilon_t. \tag{10}$$

The policy shock ε_t is distributed $N(0, \sigma_{sub}^2)$. I choose $\overline{sub} = 0.01$, and given that choose the variance of the policy shock σ_{sub} so that a one standard deviation shock roughly corresponds to a 50% bonus depreciation allowance.

Average Impulse Response to a Policy Shock Figure 4 illustrates the transmission mechanism of stimulus policy starting from steady state. Taken together, the two panels imply that the average semi-elasticity of aggregate investment with respect to stimulus is nearly one. In what follows, I will not focus on this average elasticity and instead focus on two key implications of my model: how the elasticity varies over the business cycle and how it can be increased by a more cost effective micro-targeted policy.

FIGURE 4: Impulse Response to Policy Shock in Steady State

(a) Effective Price q_t (b) Aggregate Investment



Notes: Impulse response to one standard deviation policy shock, starting from steady state. (a) Effective price of investment. (b) Aggregate investment.

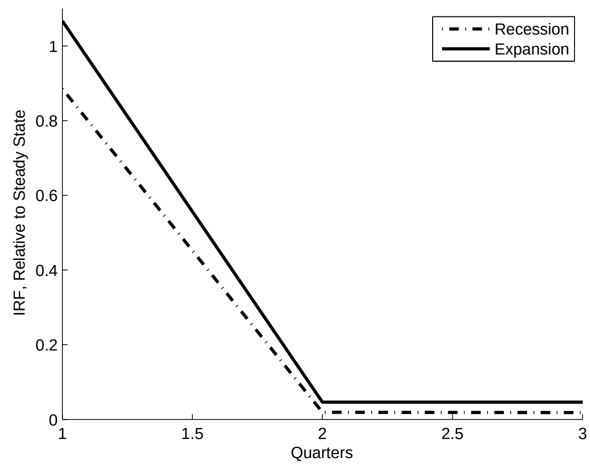
5.2 State Dependent Effective of Policy

As in Section 4.1, I first illustrate the state dependent effects of policy by comparing the impulse response of a one standard deviation positive policy shock starting from two different points in the cycle.[24] I generate these two different starting points with two different histories of productivity shocks: a history of no shocks, which leaves the economy in steady state, and a history of negative shocks, which pushes the economy into a severe recession comparable to the 2007 recession in terms of the fall in GDP. Starting from the recession, stimulus policy is generates nearly 20% less investment than it would on average. A linear forecasting model – such as user cost or tax-adjusted q models – would abstract from this state dependence and therefore be biased up in recessions.

A natural way to overcome this negative state dependence is to offer a larger

[24]State dependence in the effect of policy does not follow directly from the result sin Section 4.1 because policy shocks have different general equilibrium implications than productivity shocks.

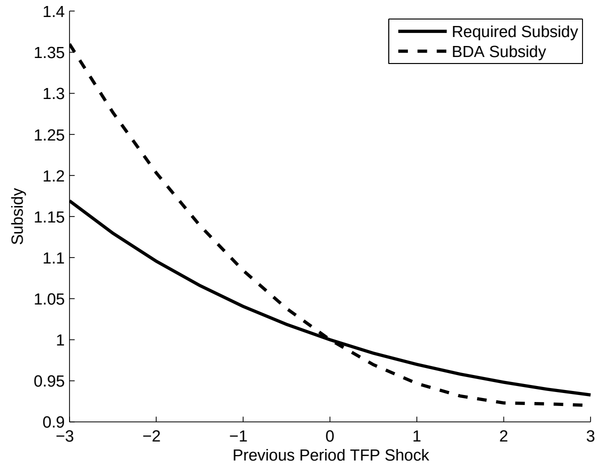
FIGURE 5: State Dependent Impulse Response to Policy Shock



Notes: Impulse response upon impact to one standard deviation productivity shock, after different histories of shocks. Normalized so that response upon impact in steady state is 1.

subsidy in recessions. The solid line in Figure 6 plots the subsidy necessary to raise the same amount of investment as in steady state as a function of the last year's productivity shocks. Consistent with the impulse response in Figure 5, the required subsidy is larger in recessions, when any given policy is less powerful. The dashed line in Figure 6 shows that the subsidy implicitly offered by the Bonus Depreciation Allowance, $sub_t = 1 - PV_t$, qualitatively follows this countercyclical pattern and is quantitatively large enough to overcome the negative state dependence. This happens because $1 - PV_t$ is countercyclical; in recessions, high interest rates imply that firms discount future writeoffs more heavily and therefore value pulling the writeoffs into the present more highly.

FIGURE 6: State Dependence of Policy Multiplier



Notes: Solid line is subsidy required to generate same amount of investment as one standard deviation policy shock in steady state. Dotted line is subsidy implied by a 50% bonus depreciation allowance. Horizontal axis is size of last year's productivity shocks. Normalized so that subsidy in steady state is 1.

5.3 Increasing Cost Effectiveness with Micro Targeting

A general issue with investment stimulus policies is that most of their cost is due to subsidizing investment which would have been done even without the policy. To increase cost effectiveness, we would like to avoid paying for this inframarginal investment and instead focus incentives on subsidizing marginal investment that is done because of the policy. In general, it is difficult to identify the marginal from inframarginal investment. However, because micro-level investment is lumpy in my model, most of the inframarginal investment is accounted for by subsidizing inframarginal *firms* who would have made an *extensive* margin investment even without the policy. This simplifies the policy problem to identifying these inframarginal firms and not subsidizing them.

To illustrate the power of this insight, I propose a simple micro-targeted policy which conditions on the size of the firm. In the model, small firms grow faster than average due to mean reversion in idiosyncratic shocks, and are therefore more likely to invest even without the policy and should therefore be avoided. I avoid subsidizing these firms with the following firm-level subsidy per unit of investment:

$$sub_{jt} = \alpha_1 \times n_{jt}^{\alpha_2}$$

where α_1 controls the baseline slope of the subsidy and α_2 controls how much the subsidy favors avoids subsidizing small firms. I assume this policy is implemented for one period and is completely unexpected.[25]

Figure 7 shows that this size-dependent subsidy can generate up to five time more investment than existing policies. For each value of α_2 on the horizontal axis, I choose α_1 to ensure budget equivalence and plot the total amount of investment generated by the policy. Larger values of α_2 place a smaller weight on small firms and generate substantially more investment.[26]

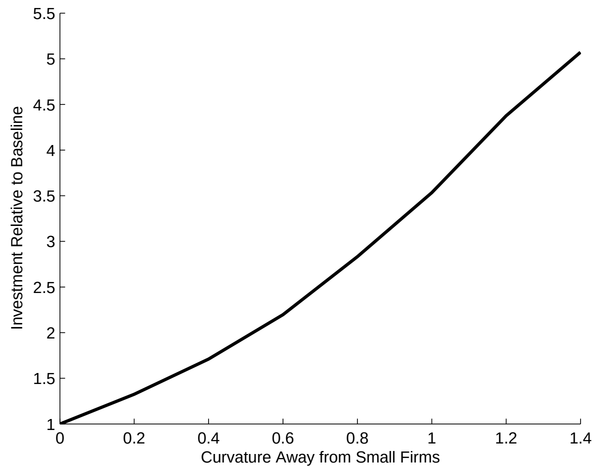
6 Conclusion

In this paper, I have argued that jointly accounting for the lumpiness of investment at the micro level and the mild countercyclicality of the real interest

[25] The goal of this exercise is to illustrate the potential cost savings associated with micro-targeting firms along the extensive margin rather than advocate this particular size-dependent specification. In reality, other factors affect firm's investment decisions that are potentially important for policy design. Furthermore, to the extent that the subsidy is expected firms may alter their employment decisions to take advantage of it. However, these complications are only important to consider in light of the result that micro-targeted policies are in principle a powerful policy tool.

[26] I do not consider higher values of α_2 because they would lead to a negative effective price of investment, as the same amount of resources are targeted to a smaller group of firms.

Figure 7: Amount of Investment From Micro-Targeted Policy



Notes: Total amount of investment generated by per-unit subsidy $\alpha_1 \times n^{\alpha_2}$, where n is employment. Normalized so that amount generated by $\alpha_2 = 0$ is 1. Horizontal axis is α_2; given α_2, I choose α_1 which is budget equivalent to the case $\alpha_2 = 0$, and plot the total amount of investment generated on the vertical axis.

rate at the macro level has important implications for aggregate investment dynamics. Together, these two facts imply that the elasticity of aggregate investment with respect to productivity and policy shocks is procyclical. I also explored how to exploit the lumpiness of investment to significantly improve the cost effectiveness of investment stimulus policy.

In order to emphasize the key role played by micro-level lumpiness and real interest rate dynamics, I have abstracted from other forces which are also potentially important in accounting for aggregate investment dynamics. First, in adopting a real business cycle framework, I exclude nominal rigidities which likely motivates the use of investment stimulus policy in the first place. Incorporating this channel would likely raise the average size of the policy multiplier through the aggregate demand channel but not significantly alter the state-dependence generated by lumpy investment. Second, I have abstracted from financial frictions. Although such frictions are not required to match the features of the data I target in this paper, it is likely that the estimated adjustment costs partly capture the effect of financial frictions in these data. Determining the relative magnitudes of adjustment costs and financial frictions is an important task for future research.

References

Bachmann, Rudiger, Ricardo Caballero, and Eduardo Engel (2013), "Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model," *American Economic Journal: Macroeconomics*, 5(4), 29-67.

Bachmann, Rudiger and Lin Ma (2016), "Lumpy Investment, Lumpy Inventories," *Journal of Money, Credit and Banking*, 48(5), 821-855.

Basu, Susanto, Miles Kimball, and John Fernald (2006), "Are Technology Improvements Contractionary?," *American Economic Review*, 96(5), 1418-1448.

Beaudry, Paul and Alain Guay (1996), "What do Interest Rates Reveal About the Functioning of Real Business Cycle Models?," *Journal of Economic Dynamics and Control*, 20(9), 1661 - 1682.

Berger, David and Joseph Vavra (2015), "Consumption Dynamics During Recessions," *Econometrica*, 83(1), 101-154.

Boldrin, Michele, Lawrence Christiano, and Jonas Fisher (2001), "Habit Persistence, Asset Returns, and the Business Cycle," *American Economic Review*, 149 - 166.

Caballero, Ricardo, Eduardo Engel, and John Haltiwanger (1995), "Plant-level Adjustment and Aggregate Investment Dynamics," *Brookings Papers on Economic Activity*, 1-54.

Caballero, Ricardo and Eduardo Engel (1999), "Explaining Investment Dynamics in US Manufacturing: A Generalized (S,s) Approach," *Econometrica* 67(4), 783-826.

Caballero, Ricardo and Eduardo Engel (2007), "Price Stickiness in Ss Models: New Interpretations of Old Results," *Journal of Monetary Economics*, 54, 100-121.

Campbell, John and John Cochrane (1999), "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market," *The Journal of Political Economy*, 107(2), 205-251.

Cooper, Russell and John Haltiwanger (2006), "On the Nature of Capital Adjustment Costs," *The Review of Economic Studies*, 73(3), 611-633.

Cooper, Russell and Jonathan Willis (2014), "Discounting: Investment Sensitivity and Aggregate Implications," *Unpublished manuscript.*

Doms, Mark and Timothy Dunne (1998), "Capital Adjustment Patterns in Manufacturing Plants," *Review of Economic Dynamics*, 1(2), 409-429.

Edge, Rochelle and Jeremy Rudd (2011), "General-Equilibrium Effects of Investment Tax Incentives," *Journal of Monetary Economics*, 58(6), 564-577.

Gourio, Francois and Anil Kashyap (2007), "Investment Spikes: New Facts and a General Equilibrium Exploration," *Journal of Monetary Economics*, 54, 1-22.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory Huffman (1998), "Investment, Capacity Utilization, and the Real Business Cycle," *The American Economic Review*, 402-417.

House, Christopher and Matthew Shapiro (2008), "Temporary Investment Tax Incentives: Theory with Evidence from Bonus Depreciation," *The American Economic Review*, 737-768.

Jermann, Urban (1998), "Asset Pricing in Production Economies," *Journal of Monetary Economics*, 41(2), 257-275.

Khan, Aubhik and Julia Thomas (2004), "Nonconvex Factor Adjustments in Equilibrium Business Cycle Models: Do Nonlinearities Matter?," *Journal of Monetary Economics,* 50(2), 331-360.

Khan, Aubhik and Julia Thomas (2008), "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, 76(2), 395-436.

King, Robert and Sergio Rebelo (1999), "Resuscitating Real Business Cycles," *Handbook of Macroeconomics*, 927-1007.

King, Robert and Mark Watson (1996), "Money, Prices, Interest Rates, and the Business Cycle," *The Review of Economics and Statistics*, 35-53.

Krusell, Per and Anthony Smith (1998), "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867-896.

Miao, Jianjun and Pengfei Wang (2014), "A Q-Theory Model with Lumpy Investment," *Economic Theory* 57(1), 133-159.

Miao, Jianjun and Pengfei Wang (2014), "Lumpy Investment and Corporate Tax Policy," *Journal of Money, Credit and Banking*, 46(6), 1171 - 1203.

Reiter, Michael (2009), "Solving Heterogeneous Agent Models by Projections and Perturbation," *Journal of Economic Dynamics and Control*, 33(3), 649-665.

Summers, Lawrence (1986), "Some Skeptical Observations on Real Business Cycle Theory," *Quarterly Review,* Fall, 23-27.

Thomas, Julia (2002), "Is Lumpy Investment Relevant for the Business Cycle?," *Journal of Political Economy*, 110(3), 508-534.

Winberry, Thomas (2016), "A Toolbox for Solving and Estimating Heterogeneous Agent Macro Models," *Unpublished manuscript.*

Zwick, Eric and James Mahon (2016), "Do Financial Frictions Amplify Fiscal Policy? Evidence from Business Investment Stimulus," *The American Economic Review* (forthcoming).

A Appendix (For Online Publication Only)

A.1 Robustness of Countercyclical Real Interest Rate Dynamics

Data Sources I obtain rate of return data from the St. Louis `FRED` database. The 90-day treasury bill rate is the average secondary market rate over the quarter, not seasonally adjusted. The AAA corporate bond yield is Moody's seasoned AAA corporate bond yield averaged over the quarter and not seasonally adjusted. The BAA corporate bond yield is Moody's seasoned BAA corporate bond yield averaged over the quarter and not seasonally adjusted. The return on the stock market is computed from the Russell 3000 total market index, averaged over the quarter and not seasonally adjusted.

I obtain quantity data from the BEA tables. Output is real gross domestic product. I construct TFP as measured TFP at a quarterly frequency using total hours in the nonfarm business sector from BLS productivity and costs release and a quarterly capital series constructed as in Bachmann, Caballero, and Engel (2013).

Filtering Table VII shows that the countercyclicality of the real interest rates reported in Table II is robust to different filtering methods. Relative to the Hodrick-Prescott filter used in the main text, a linear trend leaves less overall variation but the cyclicality is largely unaffected. The bandpass filter leaves less variation but again the cyclicality is unaffected. The correlation of raw interest rates with raw GDP growth is substantially smaller than for the other filters, though still negative for the T-bill, AAA, and BAA rates.

TABLE VII
INTEREST RATE DYNAMICS USING DIFFERENT FILTERS

	Linear trend		**Bandpass**		**First diff.**	
	$\sigma(r_t)$	$\rho(r_t, y_t)$	$\sigma(r_t)$	$\rho(r_t, y_t)$	$\sigma(r_t)$	$\rho(r_t, \Delta y_t)$
T-bill	2.53%	-0.13	1.40%	-0.12	3.03%	-0.06
AAA	2.70%	-0.31	1.64%	-0.36	3.31%	-0.09
BAA	2.83%	-0.35	1.77%	-0.40	3.96%	-0.12
Stock	25.68%	-0.15	16.76%	-0.17	25.97%	0.27

Notes: Cyclical properties of various measures of the real interest rate from Table II and different filtering methods. r_t refers to particular rate of return and y_t to the particular cyclical component of log real GDP. "HP filter" takes out an HP trend with smoothing parameter 1600 of both r_t and y_t. Linear trend takes out a linear trend, approximated as an HP filter with smoothing parameter 100,000, of both r_t and y_t. "Bandpass" takes out a Baxter and King (1996) bandpass filter with minimum periodicity 6 quarters and maximum periodicity 32 quarters of both r_t and y_t. "First diff." uses the level of the real interest rate r_t and the first difference of log real GDP Δy_t.

Measuring Inflation Table VIII shows that the conclusions in the main text are robust to different measures of inflation. Table IX shows that conclusions are robust to using either ex-ante or ex-post inflation to measure expectations.

A.2 Characterizing Equilibrium

In this Appendix I characterize the recursive competitive equilibrium defined in Section 2.4. I use this characterization to numerically compute the equilibrium in Appendix A.3 and to model investment stimulus policy in Proposition 1 of the main text.

Firm's Decision Problem I begin by simplifying the firm's decision problem in a series of three propositions. These propositions eliminate two individual state variables, which greatly simplifies the numerical approximation.

Table VIII
Interest Rate Dynamics For Different Measures of Inflation

	CPI		**GDP Deflator**		**PCE Defl.**		**Inves. Defl.**	
	$\sigma(r_t)$	$\rho(r_t, y_t)$	$\sigma(r_t)$	$\rho(r_t, y_t)$	$\sigma(r_t)$	$\rho(r_t, y_t)$	$\sigma(r_t)$	$\rho(r_t, y_t)$
T-bill	2.18%	-0.17	2.75%	-0.01	1.81%	-0.08	2.27%	-0.20
AAA	2.34%	-0.37	3.33%	-0.30	1.85%	-0.34	2.37%	-0.40
BAA	2.43%	-0.41	4.29%	-0.36	1.94%	-0.40	2.46%	-0.44
Stock	24.7%	-0.14	9.04%	-0.12	24.60%	-0.13	25.05%	-0.15

Notes: Cyclical properties of various measures of the real interest rate from Table II and different measures of inflation. "CPI" refers to the Consumer Price Index for all urban consumers. "GDP Deflator" is the implicit price deflator for GDP from BEA Table 1.1.9. "PCE Defl." is the implicit price deflator for personal consumption expenditure from BEA Table 1.1.9. "Invest. Defl." is the implicit price deflator for nonresidential fixed investment from BEA Table 1.1.9.

Table IX
Interest Rate Dynamics Using Different Measures of Expectations

	Realized Inflation		**Ex-Ante Inflation**	
	$\sigma(r_t)$	$\rho(r_t, y_t)$	$\sigma(r_t)$	$\rho(r_t, y_t)$
T-bill	2.18%	-0.17	2.12%	-0.17
AAA	2.34%	-0.37	2.26%	-0.38
BAA	2.43%	-0.41	2.36%	-0.42
Stock	24.7%	-0.14	24.78%	-0.15

Notes: Cyclical properties of various measures of the real interest rate from Table II and different methods of computing expected inflation. "Realized inflation" refers to using realized value of inflation in $t+1$ to compute return in $t+1$. "Ex-Ante Inflation" refers to using realized value of inflation in t to compute return in $t+1$; this is expected inflation if inflation follows a random walk.

For ease of notation, define after-tax revenue net of tax writeoffs:

$$\pi\left(\varepsilon,k;\mathbf{s}\right)=\max_{n}\left\{\left(1-\tau\right)\left(e^{z}e^{\varepsilon}k^{\theta}n^{\nu}-w(\mathbf{s})n\right)\right\}$$

By construction, this does not depend on current depreciation allowances d or the fixed adjustment cost ξ.

The first proposition shows that the firm's value function $v(\varepsilon,k,d,\xi;\mathbf{s})$ is linear in the pre-existing stock of depreciation allowances d. I exploit this property in the other propositions to simplify the decision rules.

***Proposition* 2.** *The firm's value function is of the form $v(\varepsilon,k,d,\xi;\mathbf{s})=v^{1}(\varepsilon,k,\xi;\mathbf{s})+\tau PV(\mathbf{s})d$ where $PV(\mathbf{s})$ is defined by the recursion $PV(\mathbf{s})=\widehat{\delta}+\left(1-\widehat{\delta}\right)E\left[\Lambda(z';\mathbf{s})PV(\mathbf{s}')\right]$. Furthermore, $v^{1}(\varepsilon,k,\xi;\mathbf{s})$ is defined by the Bellman equation*

$$v^{1}(\varepsilon,k,\xi;\mathbf{s})=\pi\left(\varepsilon,k;\mathbf{s}\right)+\max_{i}\left\{\begin{array}{c}-\left(1-\tau PV(\mathbf{s})\right)i-\frac{\phi}{2}\left(\frac{i}{k}\right)^{2}k-\xi w(\mathbf{s})1\left\{i\notin[-ak,ak]\right\}\\+E[\Lambda(z';\mathbf{s})v^{1}\left(\varepsilon',\left(1-\delta\right)k+i,\xi';\mathbf{s}'\right)]\end{array}\right\}\tag{11}$$

Proof. First, I show that the value function is of the form $v(\varepsilon,k,d,\xi;\mathbf{s})=v^{1}(\varepsilon,k,\xi;\mathbf{s})+\tau PV(\mathbf{s})d$ for some function $v^{1}(\varepsilon,k,\xi;\mathbf{s})$. I begin by showing that the operator T defined by the right hand side of the Bellman equation maps function of the form $f(\varepsilon,k,\xi;\mathbf{s})+\tau PV(\mathbf{s})d$ into functions of the form $g(\varepsilon,k,\xi;\mathbf{s})+\tau PV(\mathbf{s})d$. Applying T to f, we get:

$$\begin{aligned}T(f)(\varepsilon,k,\xi;\mathbf{s})&=\pi(\varepsilon,k;\mathbf{s})+\tau\widehat{\delta}d\\&+\max_{i}\left\{\begin{array}{c}-\left(1-\tau\widehat{\delta}\right)i-\frac{\phi}{2}\left(\frac{i}{k}\right)^{2}k-\xi w(\mathbf{s})1\left\{i\notin[-ak,ak]\right\}\\+E[\Lambda(z';\mathbf{s})(f(\varepsilon',(1-\delta)k+i,\xi';\mathbf{s}')+\tau PV(\mathbf{s})(1-\widehat{\delta})(d+i))]\end{array}\right\}\end{aligned}$$

Collecting terms,

$$T(f)(\varepsilon, k, \xi; \mathbf{s}) = \pi(\varepsilon, k; \mathbf{s}) + \tau \left(\widehat{\delta} + (1-\widehat{\delta})E[\Lambda(z'; \mathbf{s})PV(\mathbf{s}')]\right) d \qquad (12)$$
$$+ \max_i \left\{ \begin{array}{c} -\left(1 - \tau\widehat{\delta} - \tau(1-\widehat{\delta})E[\Lambda(z'; \mathbf{s})PV(\mathbf{s}')]\right) i - \frac{\phi}{2}\left(\frac{i}{k}\right)^2 k \\ -\xi w(\mathbf{s}) 1\{i \notin [-ak, ak]\} + E[\Lambda(z'; \mathbf{s}) f(\varepsilon', (1-\delta)k + i, \xi'; \mathbf{s}')] \end{array} \right\}$$

By the definition of $PV(\mathbf{s})$, we have that

$$\tau \left(\widehat{\delta} + (1-\widehat{\delta})E[\Lambda(z'; \mathbf{s})PV(\mathbf{s}')]\right) d = \tau PV(\mathbf{s})$$
$$-\left(1 - \tau\widehat{\delta} - \tau(1-\widehat{\delta})E[\Lambda(z'; \mathbf{s})PV(\mathbf{s}')]\right) i = -\left(1 - \tau PV(\mathbf{s})\right) i$$

Plugging this back into (12) and rearranging gives

$$T(f)(\varepsilon, k, \xi; \mathbf{s}) = \tau PV(\mathbf{s}) d +$$
$$\underbrace{\pi(\varepsilon, k; \mathbf{s}) + \max_i \left\{ \begin{array}{c} -\left(1 - \tau PV(\mathbf{s})\right) i - \frac{\phi}{2}\left(\frac{i}{k}\right)^2 k - \xi w(\mathbf{s}) 1\{i \notin [-ak, ak]\} \\ + E[\Lambda(z'; \mathbf{s}) f\left(\varepsilon', (1-\delta)k + i, \xi'; \mathbf{s}'\right)] \end{array} \right\}}_{g(\varepsilon, k, \xi; \mathbf{s})}$$

which is of the form $\tau PV(\mathbf{s})d + g(\varepsilon, k, \xi; \mathbf{s})$. Hence, T maps functions of the form $\tau PV(\mathbf{s})d + f(\varepsilon, k, \xi; \mathbf{s})$ into functions of the form $\tau PV(\mathbf{s})d + g(\varepsilon, k, \xi; \mathbf{s})$. This is a closed set of functions, so by the contraction mapping theorem, the fixed point of T must lie in this set as well. Since the fixed point of T is the value function, this establishes that $v(\varepsilon, k, d, \xi; \mathbf{s}) = v^1(\varepsilon, k, \xi; \mathbf{s}) + \tau PV(\mathbf{s})d$.

To derive the form of $v^1(\varepsilon, k, \xi; \mathbf{s})$, plug $v(\varepsilon, k, d, \xi; \mathbf{s}) = v^1(\varepsilon, k, \xi; \mathbf{s}) + \tau PV(\mathbf{s})d$ into both sides of the Bellman equation to get

$$v^1(\varepsilon, k, \xi; \mathbf{s}) + \tau PV(\mathbf{s})d = \pi(\varepsilon, k; \mathbf{s}) + \tau\widehat{\delta} d +$$
$$\max_i \left\{ \begin{array}{c} -\left(1 - \tau\widehat{\delta}\right) i - \frac{\phi}{2}\left(\frac{i}{k}\right)^2 k - \xi w(\mathbf{s}) 1\{i \notin [-ak, ak]\} \\ + E[\Lambda(z'; \mathbf{s})(v^1(\varepsilon', (1-\delta)k + i, \xi'; \mathbf{s}') + \tau PV(\mathbf{s})(1-\widehat{\delta})(d+i))] \end{array} \right\}$$

Rearranging terms as before shows that

$$v^1(\varepsilon,k,\xi;\mathbf{s})+\tau PV(\mathbf{s})d=\pi\left(\varepsilon,k;\mathbf{s}\right)+\tau PV(\mathbf{s})d+$$
$$\max_i\left\{\begin{array}{c}-\left(1-\tau PV(\mathbf{s})\right)i-\frac{\phi}{2}\left(\frac{i}{k}\right)^2k-\xi w(\mathbf{s})1\left\{i\notin[-ak,ak]\right\}\\+E[\Lambda(z';\mathbf{s})v^1\left(\varepsilon',(1-\delta)\,k+i,\xi';\mathbf{s}'\right)]\end{array}\right\}$$

Subtracting $\tau PV(\mathbf{s})d$ from both sides establishes (11). ■

The above proposition shows that the depreciation allowances d do not interact with the other state variables of the firm. The next proposition shows that this implies that investment decisions do not depend on d. To ease notation, I first define the ex ante value function:

$$v^0\left(\varepsilon,k;\mathbf{s}\right)=\int_0^{\bar{\xi}}v^1(\varepsilon,k,\xi;\mathbf{s})\frac{1}{\bar{\xi}}d\xi.$$

Proposition 3. *The investment decision rule is independent of d and given by*

$$i(\varepsilon,k,\xi;\mathbf{s})=\left\{\begin{array}{c}i^a\left(\varepsilon,k;\mathbf{s}\right)\ \text{ if } \xi\leq\widehat{\xi}\left(\varepsilon,k;\mathbf{s}\right)\\ i^n\left(\varepsilon,k;\mathbf{s}\right)\ \text{ if } \xi>\widehat{\xi}\left(\varepsilon,k;\mathbf{s}\right)\end{array}\right\}$$

where

$$i^a\left(\varepsilon,k;\mathbf{s}\right)=\arg\max_i-\left(1-\tau PV(\mathbf{s})\right)i-\frac{\phi}{2}\left(\frac{i}{k}\right)^2k+E[\Lambda(z';s)v^0\left(\varepsilon',(1-\delta)\,k+i;\mathbf{s}'\right)$$

$$i^n\left(\varepsilon,k;\mathbf{s}\right)=\left\{\begin{array}{c}ak\ \text{ if } i^a\left(\varepsilon,k;\mathbf{s}\right)>ak\\ i^a\left(\varepsilon,k;\mathbf{s}\right)\ \text{ if } i^a\left(\varepsilon,k;\mathbf{s}\right)\in[-ak,ak]\\ -ak\ \text{ if } i^a\left(\varepsilon,k;\mathbf{s}\right)<-ak\end{array}\right\}$$

$$\widehat{\xi}(\varepsilon,k;\mathbf{s})=\frac{1}{w(\mathbf{s})}\times\left\{\begin{array}{c}-(1-\tau PV(\mathbf{s}))(i^{a}(\varepsilon,k;\mathbf{s})-i^{n}(\varepsilon,k;\mathbf{s}))\\ -\frac{\phi}{2}\left(\left(\frac{i^{a}(\varepsilon,k;\mathbf{s})}{k}\right)^{2}-\left(\frac{i^{n}(\varepsilon,k;\mathbf{s})}{k}\right)\right)k\\ +E[\Lambda(z';\mathbf{s})(v^{0}(\varepsilon',(1-\delta)k+i^{a}(\varepsilon,k;\mathbf{s});\mathbf{s}')\\ -v^{0}(\varepsilon',(1-\delta)k+i^{n}(\varepsilon,k;\mathbf{s});\mathbf{s}'))]\end{array}\right\}$$

Proof. The form of $i^a(\varepsilon,k;\mathbf{s})$ follows directly from the Bellman equation, using the law of iterated expectations and the fact that ξ' is i.i.d. The form of $i^n(\varepsilon,k;\mathbf{s})$ also follows from the Bellman equation, which shows that the objective function in the no-adjust problem is the same as the adjust problem and the choice set is restricted. The form of $i(\varepsilon,k,\xi;\mathbf{s})$ comes from the following argument. At $\xi=0$, the objective function of adjusting must be weakly greater than the no-adjust problem, because the no-adjust problem has a constrained choice set. Further, the payoff of adjusting is strictly decreasing in ξ. Therefore, there must be a cutoff rule. Setting the adjust and no adjust payoffs equal gives the form of the threshold $\widehat{\xi}(\varepsilon,k;\mathbf{s})$. ∎

The above proposition shows that knowing $v^0(\varepsilon,k;\mathbf{s})$ is enough to derive the decision rules. The next and final proposition defines the Bellman equation which determines $v^0(\varepsilon,k;\mathbf{s})$.

***Proposition* 4.** $v^0(\varepsilon,k;\mathbf{s})$ *solves the Bellman equation*

$$\begin{aligned}v(\varepsilon,k;\mathbf{s})&=\pi(\varepsilon,k;\mathbf{s})\\&+\frac{\widehat{\xi}(\varepsilon,k;\mathbf{s})}{\overline{\xi}}\left\{\begin{array}{c}-(1-\tau PV(\mathbf{s}))\,i^{a}(\varepsilon,k;\mathbf{s})-\frac{\phi}{2}\left(\frac{i^{a}(\varepsilon,k;\mathbf{s})}{k}\right)^{2}k\\ -\frac{\widehat{\xi}(\varepsilon,k;\mathbf{s})}{2}w(\mathbf{s})+E[\Lambda(z';\mathbf{s})v^{0}(\varepsilon',(1-\delta)k+i^{a}(\varepsilon,k;\mathbf{s});\mathbf{s}')]\end{array}\right\}\\&+\left(1-\frac{\widehat{\xi}(\varepsilon,k;\mathbf{s})}{\overline{\xi}}\right)\left\{\begin{array}{c}-(1-\tau PV(\mathbf{s}))\,i^{a}(\varepsilon,k;\mathbf{s})-\frac{\phi}{2}\left(\frac{i^{n}(\varepsilon,k;\mathbf{s})}{k}\right)^{2}k\\ +E[\Lambda(z';\mathbf{s})v^{0}(\varepsilon',(1-\delta)k+i^{n}(\varepsilon,k;\mathbf{s});\mathbf{s}')]\end{array}\right\}\end{aligned}$$

Proof. This follows from integrating $v^0(\varepsilon, k; \mathbf{s}) = \int v^1(\varepsilon, k, \xi; \mathbf{s}) \frac{1}{\bar{\xi}} d\xi$, using the expression for $v^1(\varepsilon, k, \xi; \mathbf{s})$ from Proposition 2 and the form of the policy function from Proposition 3. ∎

A Characterization of the Equilibrium The series of propositions above show that firms' decision rules are determined by the alternative value function $v^0(\varepsilon, k; \mathbf{s})$. I now embed this alternative value function into a simplified characterization of the recursive competitive equilibrium. In addition to simplifying firms' decisions, this characterization also eliminates household optimization by directly imposing the implications of optimization on firm behavior through prices as in Khan and Thomas (2008). To do so, define the marginal utility of consumption in state $\mathbf{s}$ as $p(\mathbf{s})$. Abusing notation, I then renormalize the value function through

$$v(\varepsilon, k; \mathbf{s}) = p(\mathbf{s}) v^0(\varepsilon, k; \mathbf{s})$$

This renormalization leaves the decision rules unchanged and I continue to denote them $i^a(\varepsilon, k; \mathbf{s})$, etc. In a final abuse of notation, I denote the distribution of firms over measurable sets $\Delta_\varepsilon \times \Delta_k$ as μ.

***Proposition* 5.** *The recursive competitive equilibrium from Definition 1 is characterized by a list of functions $v(\varepsilon, k; \mathbf{s})$, $w(\mathbf{s})$, $p(\mathbf{s})$, $S'_{-1}(\mathbf{s})$, $C'_{-1}(\mathbf{s})$, and $\mu'(\mathbf{s})$ such that*

(i) (Firm optimization) $v(\varepsilon, k; \mathbf{s})$ solves the Bellman equation

$$\begin{aligned} v(\varepsilon, k; \mathbf{s}) &= p(\mathbf{s})\pi(\varepsilon, k; \mathbf{s}) \\ &+ \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\bar{\xi}} \left\{ \begin{array}{c} -p(\mathbf{s})(1 - \tau PV(\mathbf{s}))\, i^a(\varepsilon, k; \mathbf{s}) - p(\mathbf{s}) \frac{\phi}{2} \left(\frac{i^a(\varepsilon, k; \mathbf{s})}{k} \right)^2 k \\ -p(\mathbf{s}) \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{2} w(\mathbf{s}) + \beta E[v(\varepsilon', (1-\delta)k + i^a(\varepsilon, k; \mathbf{s}); \mathbf{s}')] \end{array} \right\} \\ &+ \left(1 - \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\bar{\xi}}\right) \left\{ \begin{array}{c} -p(\mathbf{s})(1 - \tau PV(\mathbf{s}))\, i^n(\varepsilon, k; \mathbf{s}) - p(\mathbf{s}) \frac{\phi}{2} \left(\frac{i^n(\varepsilon, k; \mathbf{s})}{k} \right)^2 k \\ +\beta E[v(\varepsilon', (1-\delta)k + i^n(\varepsilon, k; \mathbf{s}); \mathbf{s}')] \end{array} \right\} \end{aligned}$$

where $i^a(\varepsilon,k;\mathbf{s})$, $i^n(\varepsilon,k;\mathbf{s})$, *and* $\widehat{\xi}(\varepsilon,k;\mathbf{s})$ *are derived from* $v(\varepsilon,k;\mathbf{s})$ *using*

$$i^a(\varepsilon,k;\mathbf{s}) = \arg\max_i -p(\mathbf{s})\left(1-\tau PV(\mathbf{s})\right)i - p(\mathbf{s})\frac{\phi}{2}\left(\frac{i}{k}\right)^2 k + \beta E[v(\varepsilon',(1-\delta)k+i;\mathbf{s}'))$$

$$i^n(\varepsilon,k;\mathbf{s}) = \left\{\begin{array}{c} ak \;\; \textit{if } i^a(\varepsilon,k;\mathbf{s}) > ak \\ i^a(\varepsilon,k;\mathbf{s}) \;\; \textit{if } i^a(\varepsilon,k;\mathbf{s}) \in [-ak,ak] \\ -ak \;\; \textit{if } i^a(\varepsilon,k;\mathbf{s}) < -ak \end{array}\right\}$$

$$\widehat{\xi}(\varepsilon,k;\mathbf{s}) = \frac{1}{p(\mathbf{s})w(\mathbf{s})} \times \left\{\begin{array}{c} -p(\mathbf{s})\left(1-\tau PV(\mathbf{s})\right)(i^a(\varepsilon,k;\mathbf{s}) - i^n(\varepsilon,k;\mathbf{s})) \\ -p(\mathbf{s})\frac{\phi}{2}\left(\left(\frac{i^a(\varepsilon,k;\mathbf{s})}{k}\right)^2 - \left(\frac{i^n(\varepsilon,k;\mathbf{s})}{k}\right)\right)k \\ +\beta E[(v(\varepsilon',(1-\delta)k+i^a(\varepsilon,k;\mathbf{s});\mathbf{s}') \\ -v(\varepsilon',(1-\delta)k+i^n(\varepsilon,k;\mathbf{s});\mathbf{s}'))] \end{array}\right\}$$

and $PV(\mathbf{s})$ *is defined by the recursion*

$$p(\mathbf{s})PV(\mathbf{s}) = p(\mathbf{s})\widehat{\delta} + \left(1-\widehat{\delta}\right)\beta E\left[p(\mathbf{s}')PV(\mathbf{s}')|\mathbf{s}\right].$$

(ii) *(Labor market clearing)*

$$\left(\frac{w(\mathbf{s})}{\chi}\right)^{\frac{1}{\eta}} = \int \left(n(\varepsilon,k;\mathbf{s}) + \frac{\widehat{\xi}(\varepsilon,k;\mathbf{s})^2}{2\overline{\xi}}\right)\mu(d\varepsilon,dk)$$

where $n(\varepsilon,k;\mathbf{s}) = \left(\frac{e^z e^\varepsilon k^\theta \nu}{w(\mathbf{s})}\right)^{\frac{1}{1-\nu}}$.

(iii) *(Consistency)*

$$p(\mathbf{s}) = \left(C(\mathbf{s}) \times S(\mathbf{s}) - \chi\frac{\left(\left(\frac{w(\mathbf{s})}{\chi}\right)^{\frac{1}{\eta}}\right)^{1+\eta}}{1+\eta}\right)^{-\sigma}$$

where $C(\mathbf{s})$ *is derived from the decision rules by* $C(\mathbf{s}) = \int (e^z e^{\varepsilon} k^{\theta} n(\varepsilon, k; \mathbf{s})^{\nu} - i(\varepsilon, k; \mathbf{s}) - AC(\varepsilon, k; \mathbf{s}))\mu(d\varepsilon, dk)$ *using* $i(\varepsilon, k; \mathbf{s}) = \frac{\widehat{\xi}(\varepsilon,k;\mathbf{s})}{\bar{\xi}} i^a(\varepsilon, k; \mathbf{s}) + \left(1 - \frac{\widehat{\xi}(\varepsilon,k;\mathbf{s})}{\bar{\xi}}\right) i^n(\varepsilon, k; \mathbf{s})$ *and*

$$AC(\varepsilon, k; \mathbf{s}) = \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\bar{\xi}} \left(\frac{\phi}{2}\left(\frac{i^a(\varepsilon, k; \mathbf{s})}{k}\right)^2 k\right) + \left(1 - \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\bar{\xi}}\right)\left(\frac{\phi}{2}\left(\frac{i^n(\varepsilon, k; \mathbf{s})}{k}\right)^2 k\right).$$

$S(\mathbf{s})$ *is derived from* $C(\mathbf{s})$ *using* $S(\mathbf{s}) = \overline{S}^{1-\rho_s} S_{-1}^{\rho_s} \left(\frac{C(\mathbf{s})}{C_{-1}}\right)^{\lambda}$.

(iv) (Laws of motion)

$$S'_{-1}(\mathbf{s}) = S(\mathbf{s})$$
$$C'_{-1}(\mathbf{s}) = C(\mathbf{s})$$

(v) (Law of motion for measure) For all measurable sets $\Delta_{\varepsilon} \times \Delta_k$,

$$\mu'(\mathbf{s})(\Delta_{\varepsilon} \times \Delta_k) = \int p(\varepsilon' \in \Delta_{\varepsilon} | \varepsilon)\Big(\frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\bar{\xi}} 1\{(1-\delta)k + i^a(\varepsilon, k; \mathbf{s}) \in \Delta_k\} + \left(1 - \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\bar{\xi}}\right) 1\{(1-\delta)k + i^n(\varepsilon, k; \mathbf{s}) \in \Delta_k\} d\varepsilon' \mu(d\varepsilon, dk)$$

Proof. Condition (i) follows from Propositions 2-4, using the definition $v(\varepsilon, k; \mathbf{s}) = p(\mathbf{s})v^0(\varepsilon, k; \mathbf{s})$ and noting that $\Lambda(z'; \mathbf{s}) = \frac{\beta p(\mathbf{s}')}{p(\mathbf{s})}$. Condition (ii) follows from the household's FOC, the firms' FOC, and labor market clearing. Condition (iii) follows from output market clearing and the definition of $p(\mathbf{s})$. Condition (iv) directly reproduces conditions iv(c) and iv(d) from Section 2.4 in the main text. Condition (v) follows from the original law of motion in condition iv(e) in the main text, eliminating d as an individual state variable and integrating out ξ. ■

A.3 Solution Algorithm

I solve the model using the Winberry (2016) method which approximates the entire cross-sectional distribution of firms. I provide a brief overview of the method in this appendix and refer to the interested reader to Winberry (2016), which describes in detail how to use the method to solve Khan and Thomas' (2008) model. The method proceeds in three broad steps. First, in each period t I discretize the model's equilibrium conditions – including the cross-sectional distribution – using finite-dimensional approximations. Second, I solve for the steady state of the discretized model in which there are no aggregate shocks. Third, I solve for the dynamics of the discretized model by perturbing around the steady state.

The key step of the method is finding an appropriate approximation for the value function in period t, $v_t(\varepsilon, k)$, and the distribution $\mu_t(\varepsilon, k)$; the remaining variables are scalars so no approximation is necessary. I approximate the value function using a weighted sum of Chebyshev polynomials indexed by a vector of weights θ_t.[27] I approximate the density function of the distribution, denoted $g(\varepsilon, \log(k))$, using the parametric family

$$\begin{aligned} g\left(\varepsilon, \log(k)\right) \cong g_0 \exp\{g_1^1 \left(\varepsilon - m_1^1\right) + g_1^2 \left(\log(k) - m_1^2\right) + \\ \sum_{i=2}^{n_g} \sum_{j=0}^{i} g_i^j \left[\left(\varepsilon - m_1^1\right)^{i-j} \left(\log(k) - m_1^2\right)^j - m_i^j\right]\}, \end{aligned} \tag{13}$$

where n_g indexes the degree of approximation, $\left\{g_i^j\right\}_{i,j=(1,0)}^{(n_g,i)}$ are parameters, and $\left\{m_i^j\right\}_{i,j=(1,0)}^{(n_g,i)}$ are centralized moments of the distribution. Winberry (2016) shows that the fact that the parameters and moments must be consistent pro-

[27] The notation in this discussion follows the exposition of Winberry (2016), which provides further details.

vides a convenient method for approximating the law of motion of the distribution. With all of these approximations, the discretized equilibrium of the model is characterized by a sequence of state vectors $\mathbf{x}_t = (\mathbf{m}_t, C_{t-1}, S_{t-1}, z_t)$ and control vectors $\mathbf{y}_t = (\theta_{\mathbf{t}}, \mathbf{g}_t, p_t, w_t)$ which satisfy

$$\mathbb{E}_t[f(\mathbf{x}_t, \mathbf{x}_{t+1}, \mathbf{y}_t, \mathbf{y}_{t+1})] = \mathbf{0},$$

where f is a function returning equilibrium condition residuals. This is a standard canonical form in the perturbation literature and Winberry (2016) shows how it can be solved using `Dynare`.

I use a $n_g = 2$ approximation of the distribution, in which case the parametric family reduces to a log-normal distribution.[28] To focus on the nonlinearities emphasized in the main text, I do not directly use the locally accurate solution to compute the dynamics of aggregate variables; instead, following Krusell and Smith (1997), at each point in time I use the distribution's law of motion to forecast future prices but compute current prices and aggregate quantities to solve the nonlinear equilibrium conditions exactly. Using this approach, the locally accurate dynamics have the same interpretation as a Krusell and Smith (1997) forecasting rule. Table X shows that this implied forecasting rule has a very high R^2, a typical measure of accuracy in this literature.

A.4 Unconditional Second Moments

In this Appendix I show that the model performs equally well as a benchmark real business cycle model with respect to the unconditional second moments of

[28] In a model simulation, higher-order moments implied by directly aggregating individual decisions were close to those implied by the log-normal family. Furthermore, the parametric family implies very similar aggregates to those generated by a fully nonparametric histogram.

Table X
Forecast Accuracy

Forecast	$\mathbf{R}^2$
Wage	.9998
Marginal utility	.9997

Notes: Accuracy of forecasting rules for w_t and p_t in model simulation, comparing the prices implied by the forecasting rule to those which exactly solve market clearing conditions.

Table XI
Unconditional Second Moments

(a) Volatility				(b) Autocorrelation			
Statistic	**Data**	**Model**	**RBC**	**Statistic**	**Data**	**Model**	**RBC**
$\sigma(Y)$	1.57%	1.61%	1.59%	$\rho(Y, Y_{-1})$	.85	.72	.72
$\sigma(C)/\sigma(Y)$	.53	.66	.66	$\rho(C, C_{-1})$	.88	.72	.74
$\sigma(I)/\sigma(Y)$	2.98	3.31	2.76	$\rho(I, I_{-1})$	.91	.71	.71
$\sigma(H)/\sigma(Y)$	1.21	.68	.66	$\rho(H, H_{-1})$	.91	.72	.72

Notes: All aggregate series 1954 - 2012, logged, and HP filtered. Model: full model from main text. Benchmark: no adjustment costs or habit formation.

aggregate output, consumption, investment, and hours, and compare the model predictions to the data. In the data, I measure output as real GDP from BEA `Table 1.1.6`, consumption as the sum of nondurables and durable services from BEA `Table 1.1.6`, investment as nonresidential fixed investment from BEA `Table 1.1.6`, and hours as total hours in the nonfarm business sector from the BLS productivity and costs release. All series are quarterly 1953 - 2012, logged, and HP filtered with smoothing parameter $\lambda = 1600$.

Table XI shows that the model matches these standard aggregate second moments as well as the benchmark real business cycle model. Both models reproduce the overall volatility of output and the relative ranking of the volatility of consumption and investment. Both models underpredict the volatility of hours, a well-known problem with real business cycle models.

A.5 Conditional Heteroskedasticity

In this Appendix I argue that the key prediction of my model – that the elasticity of aggregate investment with respect to shocks is procyclical – has evidence in the aggregate investment rate time series. Intuitively, the procyclical elasticity implies that the volatility of aggregate investment is itself procyclical because aggregate shocks will generate more variation. Bachmann, Caballero, and Engel (2013) formalize this idea with a simple reduced-form time series model of the aggregate investment rate:

$$\frac{I_t}{K_t} = \phi_0 + \sum_{s=1}^{p} \phi_j \frac{I_{t-j}}{K_{t-j}} + \sigma_t e_t,\ e_t \sim N(0,1) \tag{14}$$

$$\sigma_t^2 = \beta_0 + \beta_1 \left(\frac{1}{p} \sum_{j=1}^{p} \frac{I_{t-j}}{K_{t-j}} \right) + u_t$$

The first equation is a standard autoregression except that the variance of the residuals is allowed to vary over time.[29] The second equation specifies that this residual variance depends linearly on an average of past investment rates. The procyclical elasticity in my model predicts that $\beta_1 > 0$; in expansions, when past investment is high, investment is more responsive to shocks.

Table XII confirms the procyclical volatility predicted by the model. With a lag length of $p = 1$, the 90^{th} percentile of the residual standard deviation is approximately 16% higher than the 10^{th} percentile. My model quantitatively

[29] This specification assumes that the propagation of shocks is constant over time. Figure 1 suggests that propagation also changes over time through intertemporal substitution. In the interest of power I do not focus on that prediction here.

Table XII
Conditional Heteroskedasticity

p = 1			
Statistic	**Data**	**Model**	**Bench.**
$\log\left(\frac{\widehat{\sigma}_{90}}{\widehat{\sigma}_{10}}\right)$	0.032*** (.009)	0.159	0.017
$\log\left(\frac{\widehat{\sigma}_{75}}{\widehat{\sigma}_{25}}\right)$	0.015** (0.007)	0.082	0.008
p = 6			
Statistic	**Data**	**Model**	**Bench.**
$\log\left(\frac{\widehat{\sigma}_{90}}{\widehat{\sigma}_{10}}\right)$	0.263*** (0.041)	0.112	0.006
$\log\left(\frac{\widehat{\sigma}_{75}}{\widehat{\sigma}_{25}}\right)$	0.134*** (0.025)	0.057	0.003

Notes: Results from estimating equation 14 in the text. Standard errors computed using a bootstrapping procedure.

generates nearly the same amount of heteroskedasticity as in the data. With the data's preferred lag length (according to the Akaike Information Criterion) $p = 6$, the 90^{th} percentile of the residual standard deviation is nearly 30% larger than the 10^{th}. However, the model can only generate about 50% of this amount of heteroskedasticity due to the fact that it is driven by AR(1) shocks.

I have interpreted the procyclical volatility in Table XII in terms of procyclical responsiveness to underlying homoskedastic shocks, but an alternative explanation is that the respsonsiveness to shocks is constant but that shocks themselves are heteroskedastic. Table XIII present two pieces of evidence against this interpretation.[30] First, there is virtually no heteroskedasticity in measured TFP, a key shock driving a substantial fraction of fluctuations in the data. Second, there is countercyclical, not procyclical, heteroskedasticity in aggregate output. This is consistent with the evidence presented by Bloom

[30]Berger and Vavra (2015) perform similar exercises as I do here and reach the same conclusions.

Table XIII

Conditional Heteroskedasticity in Other Series

Measured TFP, p = 1		**Real GDP, p = 1**	
Statistic	**Data**	**Statistic**	**Data**
$\log\left(\frac{\widehat{\sigma}_{90}}{\widehat{\sigma}_{10}}\right)$	0.032*** (0.009)	$\log\left(\frac{\widehat{\sigma}_{90}}{\widehat{\sigma}_{10}}\right)$	-0.124*** (0.045)
$\log\left(\frac{\widehat{\sigma}_{75}}{\widehat{\sigma}_{25}}\right)$	0.015** (0.007)	$\log\left(\frac{\widehat{\sigma}_{75}}{\widehat{\sigma}_{25}}\right)$	-0.065** (0.032)

Notes: Results from estimating equation 14 in the text. Standard errors computed using a bootstrapping procedure.

(2009) that uncertainty increases in recessions. The fact that aggregate investment does not share this feature is evidence that the procyclical elasticity documented in this paper is a quantitatively powerful mechanism.