Lumpy Investment, Business Cycles, and Stimulus Policy*

Job Market Paper, Latest Version:

http://scholar.princeton.edu/sites/default/files/winberry/files/winberryjmp.pdf

Thomas Winberry†

January 8th, 2015

Abstract

Aggregate investment accounts for much of the decline in GDP during recessions, making investment stimulus a key element of countercyclical policy. However, existing models used to study these issues are jointly inconsistent with two basic facts about investment over the cycle: micro-level investment largely occurs along the extensive margin, and the real interest rate covaries slightly negatively with aggregate productivity. I build a dynamic general equilibrium model which captures these two facts, and show that they have important implications for business cycles and countercyclical policy. First, aggregate investment is more responsive to productivity shocks in expansions than in recessions, because in expansions more firms are likely to make an extensive margin investment. Second, the policy multiplier is also state dependent, declining substantially in recessions. Third, a simple size-dependent policy, which targets extensive margin investment, is five times more cost effective than existing policies.

*I am extremely grateful to Richard Rogerson, Greg Kaplan, and Esteban Rossi-Hansberg for continual guidance and support. I am also indebted to Mike Golosov, Jesus Fernandez-Villaverde, and Christian vom Lehn for numerous 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 participants of the Princeton Macro Workshop, Princeton Public Finance Working Group, Federal Reserve Bank of Minneapolis Bag Lunch Seminar, Federal Reserve Bank of Chicago Rookie Conference, and Penn Macro Jamboree for useful comments at various stages of this project.

†Department of Economics, Princeton University; email: winberry@princeton.edu

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 policy. For example, following the recent crisis, Congress enacted two investment stimulus policies: raising the Section 179 cutoff, which targeted small firms, and the Bonus Depreciation Allowance, which applied to all firms. At its peak, the Bonus Depreciation Allowance was estimated to cost up to $100 billion per year in foregone tax revenue.[2]

To evaluate these policies and design more cost effective ones, we need a model which is consistent with the central role of aggregate investment in propagating business cycle shocks. The predictions of such a model generically rely on two effects: the direct effect of a shock on individual investment decisions, and the general equilibrium effect on prices. However, existing models in the literature are jointly inconsistent with basic facts about these two effects. First, individual investment decisions are "lumpy," i.e., occur largely along the extensive margin. For example, Gourio & Kashyap (2007) find that most of the variation in manufacturing investment is due to variation in the number of firms investing. Second, the real interest rate comoves slightly negatively with aggregate productivity, whereas most models predict a highly positive comovement. Taken together, this points to a troubling gap in the literature, since the two effects jointly determine how investment responds to both business cycle and policy shocks.

I fill this gap in the literature by building a dynamic general equilibrium model which matches both the lumpiness of investment in the micro data and the real interest rate dynamics in the macro data. I first use the model to analyze business cycle fluctuations, and find that aggregate investment responds more to productivity shocks in expansions than in recessions. This happens because in expansions more firms are likely to make an extensive margin investment. I confirm this prediction in the data by showing that my model matches the procyclical volatility in the aggregate investment rate time series. I then introduce investment stimulus policy into the model and show

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

[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. However, the estimated cost over ten years is only $30 billion, reflecting the fact that the bonus defers tax payments to the future.

that aggregate investment responds less to the policy in recessions, when few firms are close to investing. Predictions which ignore this force are therefore biased upward in recessions. Finally, I develop a simple size-dependent policy which generates five times more investment than existing policies, at the same cost. The source of these cost savings is targeting firms who are likely to make an extensive margin investment in response to the policy.

In order to emphasize the contribution of lumpy investment and realistic interest rate dynamics, my model is a direct extension of a benchmark real business cycle framework.[3] Specifically, to generate lumpy investment, I assume a fixed mass of heterogeneous firms who make investment decisions subject to convex and nonconvex adjustment costs. To generate realistic interest rate dynamics, I also assume habit formation over consumption in the household's preferences. I then calibrate these new ingredients to reproduce key features of lumpy investment and real interest rate dynamics in the data.

Quantitatively, the calibrated model predicts that the response of aggregate investment to a positive productivity shock is nearly 40% lower in a severe recession than in a similarly extreme expansion. This state dependence is in sharp contrast to linear models – including both the benchmark real business cycle model and its New Keynesian relatives – which instead predict that the impulse response does not vary over the cycle. I argue that my model is a better description of the data because it matches the procyclical volatility of the investment rate time series recently documented by Bachmann, Caballero, & Engel (2013). In my model, this procyclical volatility reflects the procyclical responsiveness of aggregate investment to shocks.

Both lumpy investment and realistic interest rate dynamics are crucial to generating these state dependent impulse responses. Lumpy investment is the source of the state dependence because it implies that the impulse response depends on the mass of firms close to investing. But realistic interest rate dynamics are also important because they overturn an important irrelevance result due to Khan & Thomas (2008). In particular, Khan & Thomas (2008) show that when lumpy investment is embedded into an otherwise standard real business cycle model, highly procyclical movements in the real interest rate eliminate most of the state dependence. However, this mecha-

[3] Although this choice facilitates comparison to the existing literature, the basic economic arguments carry through to other settings as well. As I will show, the key implication of these features is to create aggregate nonlinearities, which distinguishes my framework from both the real business cycle literature and the New Keynesian DSGE literature.

nism is inconsistent with the fact that in the data, the real interest rate is mildly countercyclical. Thus, matching the dynamics of the real interest rate is crucial for understanding the aggregate implications of lumpy investment.

Before analyzing the effect of stimulus policy on aggregate investment, I show that the model matches the effect on firm-level investment measured in the data. In particular, I reproduce results from a recent paper by Zwick & Mahon (2014), who study the Bonus Depreciation Allowance in the last two recessions. Their identification strategy exploits variation in the exposure of firms to the bonus through a difference in differences design. I reproduce this variation in my model and find that the resulting difference in differences coefficient is within one standard error of Zwick & Mahon (2014)'s estimate. Lumpy investment is crucial to the model's fit; without it, the coefficient falls to nearly one half of the data's estimate.

The model predicts that the aggregate effect of stimulus policy is also state dependent, declining by more than 20% in severe recessions.[4] More generally, the extent of this decline is determined by the severity of the recession. Hence, forecasts based on policy multipliers which do not take this state dependence into account will overstate the effect of stimulus policy in recessions, with the size of the bias increasing in the severity of the recession. Such multipliers are routinely used to forecast the effect of investment stimulus policies, for example by Romer & Bernstein (2009) to predict the effect of the American Reinvestment and Recovery Act.

Finally, I use the model to learn how to improve the cost effectiveness of investment stimulus policies. Most of the cost of existing policies comes from subsidizing inframarginal investment that would have been done even without the policy. In my model, the total amount of this inframarginal investment is mainly determined by the number of firms who would make an extensive margin adjustment without the policy. Therefore, a promising way to increase cost effectiveness is to avoid subsidizing such firms, and instead concentrate on firms who would not make an adjustment without the policy. I implement this idea with a simple size-dependent subsidy which generates five times more investment than existing policies, at the same cost. My proposal avoids subsidizing small firms because they grow faster than average, and therefore invest more without the subsidy.

[4] The policy analysis in this paper is purely positive; I assume that policymakers want to stimulate investment, and study the effect of policies designed to do so. More fully characterizing the normative implications of these policies would require specifying the market failure they address, such as an aggregate demand externality. The results presented here are an important input into this exercise.

More generally, I argue that targeting extensive margin adjustment is a powerful way to increase the cost effectiveness of investment stimulus policies.

A key challenge in my analysis is efficiently solving the model, which is difficult because the distribution of firms is part of the aggregate state vector. The standard approach, following Krusell & Smith (1998), approximates this distribution with a small number of moments, typically just the mean. However, this approach breaks down in my model because the response to an aggregate shock depends on the distribution of firms relative to their adjustment thresholds; two distributions may have the same mean, but a different number of firms close to adjusting. Adding more moments to capture this feature is not feasible due to the curse of dimensionality.

To overcome this challenge, I develop a solution method which includes the entire distribution in the state vector by approximating it with a flexible parametric family. In practice, the number of required parameters may be large and still subject to the curse of dimensionality. To get around this, I follow Reiter (2009) and implement a mixed globally and locally accurate approximation. In particular, I only solve for locally accurate dynamics of the aggregate state vector, and can therefore handle many parameters describing the distribution. Implementing this mixed approach relies crucially on parameterizing the distribution, and cannot be used to overcome the curse of dimensionality in Krusell & Smith (1998).

Related Literature My paper relates to four main strands of literature. First, it contributes to a long-standing debate about the implications of micro-level lumpy investment for aggregate dynamics. On one side of the debate, papers working in partial equilibrium find that lumpy investment generates state dependence, as in my model; see, for example Bertola & Caballero (1994), Caballero & Engel (1993, 1999), Caballero, Engel, & Haltiwanger (1995), or Cooper, Haltiwanger, & Power (1999). On the other side of the debate, papers which endogenize prices in an otherwise standard real business cycle framework find that this state dependence disappears; see, for example, Veracierto (2002), Thomas (2002), or Khan & Thomas (2003, 2008). This important series of irrelevance results highlights the central role of general equilibrium in determining the impact of lumpy investment on aggregate dynamics.[5] However, the particular general equilibrium model

[5] In this spirit, House (2008) argues that, even with fixed costs, the limiting case of no capital depreciation implies the intertemporal elasticity of investment timing is infinite. In this case, general equilibrium price movements imply that aggregate investment is independent of the distribution of investment across firms.

they use predicts a highly procyclical real interest rate, which is at odds with the data. I show that when these interest rate dynamics are corrected, the irrelevance results no longer hold.[6]

Second, to match the dynamics of the interest rate, I follow Beaudry & Guay (1996) in including habit formation and convex capital adjustment costs. These features have also been successful in matching asset prices in models with production, as in Jermann (1998) or Boldrin, Christiano, & Fisher (2001). However, all of these papers assume a representative firm, and calibrate the amount of adjustment costs faced by this representative firm to match prices. It is ex ante unclear whether these calibrated adjustment costs are consistent with the micro investment data. I show that the adjustment costs necessary to match real interest rate dynamics are indeed consistent with the micro data.

Third, my paper also contributes to a large literature which studies investment stimulus policy. On the empirical side, many papers estimate the effect of policy through traditional user cost or tax-adjusted q models; see for example, Hall & Jorgensen (1967) or Cummins et. al. (1994, 1995). Zwick & Mahon (2014) take a slightly different approach and instead implement a simple difference in differences regression. I connect with this empirical literature by replicating Zwick & Mahon (2014)'s regression on data simulated from my model. On the theoretical side, Edge & Rudd (2011) incorporate investment stimulus policy into a nearly linear New Keynesian model. Hence, they predict a constant multiplier and rule out micro-targeted policies by construction. House & Shaprio (2008) present an empirical and theoretical study of the Bonus Depreciation Allowance, but also do not focus on the implications of lumpy investment for the aggregate effects.[7]

Finally, my solution method builds heavily on two key papers in the computational literature. First, I follow Reiter (2009) in using a mixed globally and locally accurate approximation scheme. However, Reiter (2009) approximates the distribution with a fine histogram, which requires many

[6] Other papers challenge these irrelevance results on other grounds; see, for example, Gourio & Kashyap (2007), Bachmann, Caballero, & Engel (2013), or Bachmann & Ma (2012). However, none of these papers emphasize the importance of interest rate dynamics.

Bridging the gap between the partial and general equilibrium approaches, Cooper & Willis (2014) parameterize an interest rate process from the data, feed it into a firm-level model of lumpy investment, and show that this also overturns the irrelevance results. My paper produces such an interest rate process endogenously in general equilibrium.

[7] Berger & Vavra (2014) analyze a related class of policies meant to stimulate consumer durable investment. In a model of lumpy durable investment, they find that stimulus policies are less effective in recessions, for the same reason that I argue here. However, they focus on more detailed features of the micro data, whereas I emphasize the importance of interest rate dynamics in aggregation. Furthermore, they do not focus on the design of more cost effective policies.

parameters. My method instead approximates the distribution with a parametric family, which requires far less parameters and is more easily extended to higher dimensional problems. Second, my method builds on Algan et al (2008), who approximate the distribution with a parametric family. However, Algan et al (2008) implement a fully global approximation, which is slow for small problems and infeasible for larger ones.

Road Map The rest of my paper is organized as follows. In Section 2, I describe the model and the solution method. I then calibrate the model in Section 3. In Section 4, I document the state dependent impulse responses and show that my model matches the procyclical volatility of aggregate investment rates in the data. In Section 5, I introduce stimulus policy into the model, show that the policy multiplier is state dependent, and develop my micro-targeting proposal. I then conclude in Section 6.

2 Model

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

2.1 Environment

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

Firms There is a fixed mass of firms $j \in [0,1]$ who produce output y_{jt} using the production function

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

where z_t and ε_{jt} are productivity shocks, k_{jt} is capital, n_{jt} is labor, and θ and ν are parameters. Production is decreasing returns to scale to ensure that the size distribution of firms is not generate, even without adjustment costs. 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

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

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, which generates 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.

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 w_t is the 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 later analyze investment stimulus policy, which works through increasing the writeoff for new investment.

Households There is a representative household with the utility function

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

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 over current consumption is judged relative to an average of past consumption. Specifically, I first define the surplus consumption ratio as $S_t = \frac{C_t - H_t}{C_t}$, and then define the law of motion for S_t as

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

which implies that current habit is a geometric average of past consumption.[8] 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 the firms in the economy, receives a lump sum transfer payment from the government, and markets are complete.

These preferences allow the model to match key features of real interest rate dynamics in the data. With standard CRRA preferences, the interest rate would be too highly correlated with output and not volatile enough. Habit formation overcomes these problems through the implied dynamics of the intertemporal marginal rate of substitution. However, with standard Cobb-Douglas or additively separable preferences over labor supply, this countercyclical interest rate would imply countercyclical hours worked. The form of labor supply here overcomes this by eliminating the wealth effect on labor supply. The form of the labor supply term used here eliminates this wealth effect, as introduced by Greenwood, Hurcowitz, & Huffman (1988).[9]

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.

[8] This approach was introduced by Campbell & Cochrane (1999), and subsequently used in the asset pricing literature with production by Lettau & Uhlig (2001) and Chen (2014), among others.

[9] These preferences are a limiting case of a more general class of preferences introduced by Jaimovich and Rebelo (2006).

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 Section 2.4, where I define the recursive competitive equilibrium.

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

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

In (3) I have written the firm's decision problem in two pieces. The first max operator represents the optimal choice of labor, and the second max operator represents the optimal choice of investment. Because the choice of labor is a purely static problem, these two choices are independent.

It is convenient to formulate the optimal investment problem conditional on whether the firm pays the fixed cost. 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:

$$\begin{aligned} v^a(\varepsilon, k, d; \mathbf{s}) &= \max_{i \in \mathbb{R}} -\left(1-\tau\widehat{\delta}\right)i - \frac{\phi}{2}\left(\frac{i}{k}\right)^2 k + E[\Lambda(z'; \mathbf{s})v(\varepsilon', k', d', \xi'; \mathbf{s}')|\varepsilon, k, d; \mathbf{s}] \\ \text{s.t. } k' &= (1-\delta)k + i \text{ and } d' = \left(1-\widehat{\delta}\right)(d+i), \end{aligned} \tag{4}$$

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

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:

$$\begin{aligned} v^n(\varepsilon, k, d; \mathbf{s}) &= \max_{i \in [-ak, ak]} -\left(1-\tau\widehat{\delta}\right)i - \frac{\phi}{2}\left(\frac{i}{k}\right)^2 k + E[\Lambda(z'; \mathbf{s})v(\varepsilon', k', d', \xi'; \mathbf{s}')|\varepsilon, k, d; \mathbf{s}] \\ \text{s.t. } k' &= (1-\delta)k + i \text{ and } d' = \left(1-\widehat{\delta}\right)(d+i). \end{aligned} \tag{5}$$

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.

2.3 Household Optimization

Since there are no dynamic links in the household's choices, the decision problem is equivalent to the following static problem state by state:[10]

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

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, $T(\mathbf{s})$.

Because markets are complete, the stochastic discount factor used by firms equals the intertemporal marginal rate of substitution state by state:

$$\Lambda(z'; \mathbf{s}) = \frac{\left(C(\mathbf{s}') \times S(\mathbf{s}') - \chi \frac{N(\mathbf{s}')^{1+\alpha}}{1+\alpha}\right)^{-\sigma}}{\left(C(\mathbf{s}) \times S(\mathbf{s}) - \chi \frac{N(\mathbf{s})^{1+\alpha}}{1+\alpha}\right)^{-\sigma}}. \tag{8}$$

[10] There are no dynamic links because investment is done by firms and the evolution of the habit stock is taken as given.

2.4 Recursive Competitive 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, S_{-1} the previous period's surplus consumption ratio, C_{-1} is previous period's consumption, and μ is the distribution of firms over their individual state vector (ε, k, d, 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*

1. *(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).*

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

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

4. *(Consistency) For all* $\mathbf{s}$,

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

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

The mapping in Condition 4(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. Specifically, 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., $(1-\delta)\, k + i\, (\varepsilon, k, d, \xi; \mathbf{s}) \in \Delta_k$. Similarly, 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 main challenge in solving for the recursive competitive equilibrium is the fact that the aggregate state vector $\mathbf{s}_t$ contains the distribution of firms, which is an infinite-dimensional object. The standard approach, following Krusell & Smith (1998), is to approximate this distribution with a small number of moments, typically only the mean. This approach breaks down in my model because, as I will show in Section 4, the response of aggregate investment to a shock depends on the mass of firms relative to their adjustment thresholds.[11] Adding more moments to capture this effect is infeasible because of the curse of dimensionality.

To overcome this challenge, I employ a method which does not rely on approximating the distribution with moments, and is nonetheless feasible to compute. The solution method has three main steps; I sketch the idea here, and provide further details in Appendix B. In the first step, I approximate the distribution of firms with a parametric family, and include the parameters of this family in the aggregate state vector. This approach can provide an arbitrarily good approximation of the distribution with a flexible enough parametric family.[12]

In the second step, I solve for the steady state of the model without aggregate shocks. To compute the parameters of the steady state distribution, I define a mapping from the current period's

[11] Using only the mean in my model yields forecast accuracy less than 80%, far below the standards of the literature. In contrast, using only the mean yields very accurate forecasts in Khan & Thomas (2008)'s model, exactly because their model does not feature the state dependent impulse responses I find in Section 3.

[12] In my particular implementation, I use a mixture of log-normals. This class of distributions works well because the idiosyncratic productivity shocks are themselves log-normal, and in the frictionless limit the capital stock would be exactly proportional to the productivity shock.

distribution into the next period's distribution as follows: given the current period's distribution and decision rules, I exactly compute moments of the next period's distribution. I then choose the parameters approximating next period's distribution to match these exact moments. The parameters of the invariant distribution are then the fixed point of this mapping.

In the third step, I solve for the dynamics of the model using a globally accurate approximation with respect to the individual state vector, but a locally accurate approximation with respect to the aggregate state vector. A globally accurate approximation is necessary for the individual state because individual shocks are large, but a locally accurate approximation is acceptable with respect to the aggregate state because aggregate shocks are small. Crucially, the locally accurate approximation is fast to compute, and can therefore accommodate many parameters to describe the distribution. However, this mix cannot be used to break the curse of dimensionality in a standard Krusell & Smith (1998) algorithm, because it relies on differentiating the mapping from the current period's distribution into the next period's distribution. Krusell & Smith (1998) never directly approximate this mapping, instead leaving it implicit in the simulation of the model.

3 Calibration

A key objective of the calibration strategy is to ensure that the two key features of the model – lumpy investment and realistic interest rate dynamics – are quantitatively reasonable. To do so, I choose a subset of parameters to reproduce these features of the data, and exogenously fix the remaining parameters.

3.1 Fixed Parameters

The parameters I fix are listed in Table 1. Most of these parameters – β, σ, η, θ, ν, δ, ρ_z, and σ_z – are common to the business cycle literature, so to ensure comparability I set them to common values. Given these parameters, I then set the tax parameters τ and $\widehat{\delta}$ to reproduce the tax burden faced by firms in the data.

A model period is one quarter. I set $\beta = .99$ so that the steady state interest rate is 4% annually, as in Prescott (1986). I set the utility function's curvature parameter to 1, which corresponds to logarithmic utility, as in Jaimovich & Rebelo (2006). I set the Frisch elasticity of labor supply

Table 1: Fixed Parameter Values

Parameter	Description	Value
β	Discount factor	.99
σ	Curvature in utility	1
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

$\frac{1}{\eta} = 2$, consistent with the range of macro elasticities discussed in Chetty et. al. (2011). I set the steady state labor share $\theta = .64$ as in Prescott (1986), and given this choose the capital share $\nu = .21$ so that the total returns to scale is .85. This value lies well within the range considered by the current literature, from .6 in Gourio & Kashyap (2007) to .92 in Khan & Thomas (2008). I set $\delta = .025$ so that the steady state aggregate investment to capital ratio is 10%, roughly in line with the average in the postwar data. I set the stochastic process for TFP to $\rho_z = .95$ and $\sigma_z = .007$, as in King & Rebelo (1999).

I set the tax rate to $\tau = .35$ to match the top marginal tax rate in the federal corporate income tax code. I then choose the tax depreciation of capital $\widehat{\delta} = .112$ to reproduce the average present value of depreciation allowances documented by Zwick & Mahon (2014). I show in Appendix A that this present value completely summarizes the impact of the tax depreciation schedule on firms' decisions.

3.2 Fitted Parameters

The parameters I choose to match firm-level investment behavior and interest rate dynamics in the data are listed in Table 3. On the firm side, they are $\overline{\xi}$, the upper bound of fixed cost draws, a, the region around zero investment which is not subject to the fixed cost, ϕ, the coefficient on the quadratic cost, and ρ_ε and σ_ε, which govern the AR(1) process for idiosyncratic productivity shocks. On the household side, they are $\overline{S}$, the mean of surplus consumption, and $\rho_{\overline{S}}$, the autocorrelation of surplus consumption. At each step of the moment-matching process, I set the leisure parameter χ to ensure that steady state hours worked is $\frac{1}{3}$.

3.2.1 Data Targets

Firm-Level Investment The firm-level investment targets are drawn from IRS tax data between 1997 and 2010, as reported in Zwick & Mahon (2014). The data are recorded annually at the firm level.[13] I focus on the distribution of investment rates, which is investment divided by capital stock. Investment is measured as expenditure on capital goods eligible for the Bonus Depreciation Allowance (essentially equipment goods). Capital is measured as the value of all tax-depreciable assets (essentially equipment and structures; I return to the mismatch of measured investment and capital below). Because there is no entry or exit in the model, I use targets drawn from a balanced panel. In what follows, I focus on the distribution of investment rates pooled over all firms and years.

This IRS data is new to the literature, which mainly uses either Compustat or the Longitudinal Research Database (LRD). The main advantage of the IRS data is that it is representative of all corporate firms in the US, while the alternatives are subject to severe sample selection. In particular, the LRD only contains firms in the manufacturing sector, which is problematic because manufacturing accounts for a small and decreasing share of total output.[14] Compustat only contains large, publicly traded firms, which is problematic because large firms behave differently than average. Overall, the IRS data used here presents a more comprehensive sample firms in the economy.

However, the main disadvantage of using the IRS data is that the measure of investment only includes equipment goods, while the measure of the capital stock includes both equipment and structures. This mismatch biases measured investment rates down, because the measure of investment excludes some goods included in the measure of capital. I show in Appendix C that the main

[13] I time-aggregate the quarterly observations in my model to match the annual frequency in the data.

[14] The LRD is also at the establishment level rather than the firm level. This is a disadvantage for studying investment stimulus policy, which operates through changes in the firm level tax bill.

Table 2: Data 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	.160	.121
Interest Rate Dynamics		
Target	**Data**	**Model**
Standard deviation of interest rates (%)	.12%	.12%
Correlation of interest rate and output	$-.21$	$-.20$

Notes: Moments targeted in calibration. Micro investment targets from firm-level IRS data, 1997 - 2010. Inaction rate is amount investment rate less than 1%. Spike rate is amount of investment rate greater than 20%. Positive investment is remainder. Interest rate is return on 90-day treasury bill adjusted for realized inflation. Output is real GDP. Both series projected on lags of TFP, logged, and HP filtered.

results in this paper are robust to matching moments from Compustat, which does not have this measurement issue. Given that Compustat has the selection issues described above, I use the IRS data for the main text of the paper.

I target the five moments from the IRS data listed in the top panel of Table 2. The first three – the "inaction rate", or fraction of the sample with investment rates less than 1%, the "spike rate," the fraction of investment rates greater than 20%, and positive investment, the fraction in between – summarize a coarse histogram of the data. The large amount of observations in both tails reflects lumpiness of micro-level decisions; almost 25% of the observations are inactive, while at the same time 15% of the observations are spikes. I also match the mean and standard deviation to ensure the model fits basic features of the distribution.

Interest Rate Dynamics I target features of the time series of the risk-free interest rate and its comovement with aggregate output. I measure the risk-free interest rate using the return on

90 day Treasury bills, adjusted for realized inflation.[15] I measure aggregate output as real GDP from the National Income and Product Accounts. Both series are 1953-2012.

I perform two transformations to the raw data to make it comparable with the model. First, I project each series on ten lags of measured TFP and squared TFP, computed as the Solow residual. This projection is a simple way to extract the portion of fluctuations due to TFP shocks, the only shock in the model.[16] Second, I apply the Hodrick-Prescott filter with parameter 1600 to the logged, fitted values. I then target the standard deviation of the interest rate series and the contemporaneous correlation of the interest rate with output.

3.2.2 Calibration Results

Table 3 lists the calibrated parameter values, which are broadly comparable to previous findings in the literature. The upper bound on the fixed cost, $\overline{\xi}$, is well between the wide range of .0083 in Khan & Thomas (2008) and 4.4 in Bachmann, Caballero, & Engel (2013). This value implies that the average cost paid is 3.5% of the total output of the firm, which is above the 1% of Khan & Thomas (2008) but below the 7% of Bachmann, Caballero, & Engel (2013). The average surplus consumption ratio, $\overline{S}$, is smaller than in Campbell & Cochrane (1999). Campbell & Cochrane (1999) target the equity premium rather than interest rate dynamics, and need a more volatile intertemporal marginal rate of substitution in order to do so.

As shown in Table 2, the calibrated model fits both the firm-level investment moments and interest rate moments well. Importantly, the model captures the coarse histogram of investment rates which is indicative of firm-level lumpy investment, and the fit to interest rate dynamics is almost exact. However, the dispersion of investment rates is too low relative to the data.

[15] I choose the risk-free rate because in my model it captures almost all of the variation in the stochastic discount factor. This is a well-known issue with most business cycle frameworks. Like these frameworks, my model does not generate a sizeable risk premium compared to the data. This is also important to correct, but I do not attempt to do so in this study given widespread disagreement in the literature. Instead, I focus on the risk-free rate, which this class of models most directly speaks to.

[16] The results are robust to two other methods of extracting the portion of fluctuations due to TFP. First, I could match the impulse response of the real interest rate to a TFP shock in the data. However, this impulse response encodes the same basic information as in the data I present below: negative upon impact and larger than a benchmark real business cycle model. Second, I could measure productivity shocks using long-run restrictions rather than Solow residuals. Beaudry & Guay (1996) perform this exercise and show that it gives similar results.

Table 3: Fitted Parameter Values

Micro Heterogeneity		
Parameter	**Description**	**Value**
$\overline{\xi}$	Upper bound on fixed costs	.44
a	Size of no fixed cost region	.003
ϕ	Quadratic adjustment cost	2.69
ρ_ε	Idiosyncratic productivity AR(1)	.94
σ_ε	Idiosyncratic productivity AR(1)	.026
Interest Rate Dynamics		
Parameter	**Description**	**Value**
$\overline{S}$	Average surplus consumption	.65
$\rho_{\overline{S}}$	Autocorrelation surplus consumption	.95

Notes: Parameters chosen to match moments in Table 2.

4 Business Cycle Analysis

I begin the business cycle analysis by showing that my model and a benchmark real business cycle model have the same predictions for unconditional second moments of aggregate output, consumption, and hours. However, they differ substantially in their predictions conditional on different points in the business cycle: in contrast to a benchmark real business cycle model, my model predicts that the response of aggregate investment to productivity shocks is state dependent.[17] I show that this state dependence allows my model to match the procyclical volatility of aggregate investment rates in the data, which the benchmark real business cycle framework does not.

[17] Throughout, I compare to the benchmark real business cycle framework as a concrete example of a model which does not generate state dependence. However, the comparison could be made more broad to include most DSGE models. The key is that the state dependence in my model reflects a nonlinearity in the response of aggregate investment to shocks, while most alternative models are linear.

Given the real business cycle framework I adopt, I focus exclusively on productivity shocks. However, the response of aggregate investment to other shocks is likely to be state dependent as well, because the state dependence reflects time variation in the mass of firms close to undergoing an extensive margin adjustment.

4.1 Unconditional Second Moments

Appendix D shows that my model and a benchmark real business cycle model have similar predictions for aggregate second moments. The particular benchmark I use is the special case of my model without adjustment frictions (ie, $\overline{\xi} = 0$ and $\phi = 0$) or habit formation (ie, $\rho_{\overline{S}} = 1$ and $\overline{S} = 1$). Without adjustment frictions, the firm side of the model collapses to an aggregate Cobb-Douglas production function, which inherits the slight decreasing returns to scale of firm-level production. Without habit formation, the household side collapses to standard GHH preferences. For the remainder of the paper, I refer to this simplified model as the benchmark model.

4.2 Model Generates State Dependent Impulse Responses

I now show that my model predicts that the response of aggregate investment to productivity shocks varies over the business cycle. To generate these different points in the cycle, I feed in two different histories of aggregate TFP shocks; one which pushes the economy into an expansion, and the other into a recession.[18] Starting from these different points, I then compute the impulse response to a one standard deviation productivity shock.

In my model, the impulse response varies substantially between the expansion and recession, as can be seen in the left panel of Figure 1. In the expansion, the response to the shock upon impact is 26% higher than it is on average; summed over time, the total effect of the shock is 17% higher than average. The fact that the total effect is smaller than the initial effect reflects intertemporal substitution, as firms pull forward investment projects they would have done in the future. In the recession, the response upon impact is 6% lower than average. Summed over time, the total effect is again 6%. This indicates that intertemporal substitution plays a smaller role in the recession.[19]

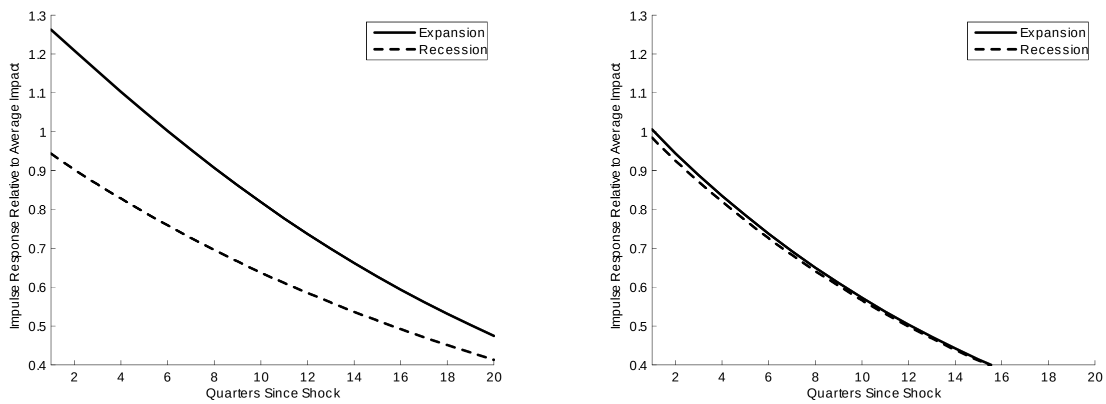
Conversely, the benchmark model generates virtually no state dependence, as shown in the right panel of Figure 1. In the expansion, the effect of the shock is 1% higher than average, and in the recession 1% less. This lack of state dependence reflects the fact that the benchmark model is

[18] To generate these histories of shocks, I back out the series aggregate productivity which generate the observed investment rate time series, and rescale the resulting shocks to match the assumed σ_z. I then generate the expansion using shocks corresponding to the late 1990s expansion and the recession using shocks corresponding to the early 2000s recession.

[19] The asymmetry in the state dependence between expansion and recession happens because the history of shocks themselves are asymmetric.

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, after different histories of 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.

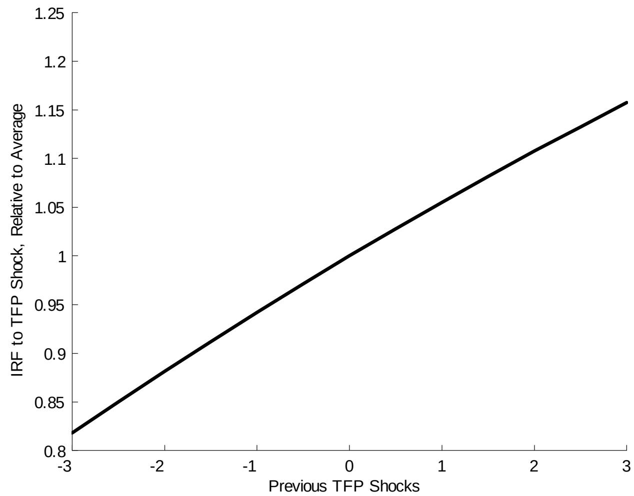
nearly linear.[20]

To show that state dependence is a general property of the model and not the product of the particular shocks chosen to generate the examples above, Figure 2 plots the impulse response upon impact as a function of the three previous period's shocks, ranging from a severe recession to an extreme expansion. This index varies from 20% lower than average in a particularly severe recession to 15% higher than average in a similarly extreme expansion. Thus, state dependence is a robust prediction of the model over all stages of the business cycle.

Mechanism The source of these state dependent impulse responses is variation in the mass of firms relative to their adjustment thresholds. When many firms are close to adjusting, an additional shock induces a lot of investment; when few firms are close to adjusting, an additional shock induces little investment. I confirm this in Figure 3, which reproduces Figure 2 for two cases: the full model, which contains this extensive margin mechanism, and the model without fixed costs, which does not contain this mechanism. Without fixed costs, the model generates virtually no

[20] In an exactly linear model, the impulse response function is independent of the aggregate state. The fact that the benchmark model is approximately linear is the reason linear approximation techniques work so well. The little state dependence I do find in the benchmark model is due to the dynamics of the capital stock. In the expansion, high previous investment increases the capital stock relative to average. This implies that the aggregate shock increases the marginal product of labor and capital more than on average. In the recession, the reverse is true.

Figure 2: State Dependent Impulse Responses Over the Cycle



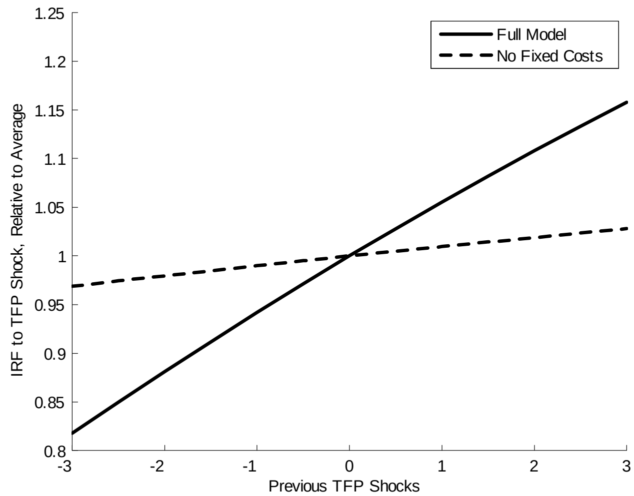
Notes: Impulse response upon impact to one standard deviation productivity shock, after different histories of shocks. Normalized so that response upon impact in steady state is 1.

state dependence in the impulse responses.

As discussed extensively in Bachmann, Caballero, & Engel (2013), the state dependence is procyclical because the mass of firms close to adjusting is itself procyclical. To understand why, recall that the investment policy of a firm is characterized by three objects: $k^a(\varepsilon, k, d; \mathbf{s})$, the target capital stock it would adjust to if it pays the fixed cost, $k^n(\varepsilon, k, d; \mathbf{s})$, the constrained capital stock it would adjust to if it does not pay the fixed cost, and $\widehat{\xi}(\varepsilon, k, d; \mathbf{s})$, the threshold fixed cost below which it pays the fixed cost. On average, $k^n(\varepsilon, k, d; \mathbf{s}) < k^a(\varepsilon, k, d; \mathbf{s})$, because once a firm adjusts capital depreciates over time. A history of good shocks, which pushes the economy into an expansion, further increases the distance between $k^a(\varepsilon, k, d; \mathbf{s})$ and $k^n(\varepsilon, k, d; \mathbf{s})$. This increases firms' probability of adjustment and makes them more sensitive to future shocks. Conversely, a history of bad shocks, which pushes the economy into a recession, decreases the target capital stock and therefore the distance between $k^a(\varepsilon, k, d; \mathbf{s})$ and $k^n(\varepsilon, k, d; \mathbf{s})$. This decreases firms' probability of adjustment and makes them less sensitive to future shocks.

Relation to the Literature There is already a large literature which studies whether micro-level lumpy investment generates state dependence in aggregate dynamics, as documented here. Early

Figure 3: State Dependent Impulse Responses, Full Model and No Fixed Costs



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.

papers in this literature, working with prices fixed in partial equilibrium, argue that this is indeed the case.[21] However, when Thomas (2002) and Khan & Thomas (2004, 2008) embedded this lumpy investment in an otherwise standard real business cycle model, they found that procyclical movements in the real interest rate eliminate most of this state dependence. In fact, they found that the resulting time series of aggregate investment is very close to the benchmark real business cycle model. Based on this result, it is tempting to conclude that lumpy investment is not important for understanding aggregate business cycle dynamics; for example, Thomas (2002)'s original irrelevance result inspired Prescott (2003) to argue that "partial equilibrium reasoning addressing an inherently general equilibrium question cannot be trusted."

I challenge this reading of the literature on the basis that the real interest rate dynamics crucial to Khan & Thomas' irrelevance result are inconsistent with key features of the data. In Table 5, I reproduce these features and compare them to the benchmark real business cycle model (which is roughly equivalent to Khan & Thomas (2008)'s framework, given their irrelevance result). The benchmark model predicts that the real interest rate is highly correlated with output, while in the data it is slightly negatively correlated. Additionally, the benchmark model underpredicts the

[21] See, for example, Bertola & Caballero (1994), Caballero, Engel, & Haltiwanger (1995), or Cooper, Haltiwanger, & Power (1999).

Table 4: Interest Rate Dynamics

Statistic	**Data**	**Model**	**Benchmark**
$\rho(\widehat{R}_t, \widehat{Y}_t)$	$-.21$	$-.20$	.65
$\sigma(\widehat{R}_t)$	.12%	.12%	.04%

Notes: Correlation of interest rate with output and standard deviation of interest rate, 1953 - 2012. Interest rate is return on 90 day treasury bills, adjusted for realized inflation. Output is real GDP. Both series are projected on lags of measured TFP, logged, and HP filtered. Model: full model in text. Benchmark: model with no adjustment costs or habit formation.

volatility of the interest rate by a factor of three.

The results documented in this section show that Khan & Thomas (2008)'s irrelevance result vanishes if one extends the model to generate realistic interest rate dynamics. The comovement of the interest rate with aggregate TFP is crucial to this result. Positive state dependence is the result of many firms wanting to make an extensive margin adjustment in expansions. In Khan & Thomas' model, the real interest rate increases in expansions and restrains this movement. However, in my model, as in the data, the interest rate instead decreases slightly. Thus, matching both micro-level lumpy investment and real interest rate dynamics is crucial for understanding the aggregate implications of lumpy investment. In Appendix E, I describe the differences between my framework and Khan & Thomas' in more detail.[22]

4.3 Model Matches Procyclical Volatility of Aggregate Investment Rate Time Series

Having shown that state dependent impulse responses is a distinctive prediction of my model relative to the benchmark model, I now assess whether this state dependence is a feature of the aggregate data, and if so, whether the model can quantitatively account for it. Intuitively, the procyclical responsiveness of aggregate investment to shocks in my model should show up as a procyclical volatility of aggregate investment in the data. To formalize this, I follow the approach

[22] In a general equilibrium environment, it is impossible to separate the effect of different adjustment frictions at firm level from different interest rate dynamics. In previous work I separated these two by exogenously feeding in a process for prices into Khan & Thomas (2008)'s calibrated model. In particular, I estimated a VAR for the dynamics of the interest rate and wage, and used this as the stochastic process for prices. I found that for this empirical price process, even with Khan & Thomas (2008)'s parameter values, the model generated state dependence.

of Bachmann, Caballero, & Engel (2013), who propose a simple statistical model to capture time variation in the impulse response of aggregate investment to shocks.

Specifically, Bachmann, Caballero, & Engel (2013) suggest the following statistical description of the aggregate investment rate:[23]

$$\begin{aligned} \frac{I_t}{K_t} &= \phi_0 + \phi_1 \frac{I_{t-1}}{K_{t-1}} + \sigma_t e_t,\ e_t \sim N(0,1) \\ \sigma_t^2 &= \beta_0 + \beta_1 \left(\frac{I_{t-1}}{K_{t-1}} \right) + u_t \end{aligned} \tag{9}$$

The first equation is a standard autoregression, except that the variance of the residuals is allowed to change over time.[24] The second equation specifies that this variance depends linearly on an average of past investment rates. Although I focus here on an AR(1)-type specification, I show in Appendix F that results are robust to more general specifications.

The possibility of heteroskedasticity in these residuals provides a simple test for state dependence. In particular, the procyclical state dependence in my model implies that $\beta_1 > 0$, i.e., that aggregate investment is more responsive to shocks in expansions, when past investment is high. Conversely, the benchmark real business cycle model predicts that $\beta_1 \approx 0$, since the response of aggregate investment to shocks is constant over time.[25]

Confirming the predictions of my model, Table 7 shows that the data exhibit substantial conditional heteroskedasticity. Because the units of the coefficient β_1 are hard to interpret, I instead report the spread of the estimated standard deviation $\widehat{\sigma}_t$ over the sample. For example, the $\log\left(\frac{\widehat{\sigma}_{90}}{\widehat{\sigma}_{10}}\right)$ shows that the residuals are approximately 16% more volatile at the 90^{th} percentile of

[23] To construct the quarterly investment rate time series, I follow the construction of Bachmann, Caballero, & Engel (2013).

[24] This assumes that the propagation of shocks is the same over time. However, as shown in Section 3.2, the model predicts that this changes over time as well through intertemporal substitution. I do not focus on this prediction here.

[25] I have interpreted changes in the residual variance as reflecting changes in the responsiveness of aggregate investment to underlying homoskedastic shocks. This was motivated by my model, which made exactly this prediction. However, an alternative explanation of this heteroskedasticity is that there is no state dependence, but rather heteroskedasticity in the shocks themselves. This alternative explanation has different policy implications than the ones I document in Section 5, and so is important to rule out.

In Appendix F, I argue that state dependence is the more plausible interpretation of the data. I do so in two steps. First, I show that there is no heteroskedasticity in measured TFP shocks. This is important given the prominent role of these shocks in my model as well as in generating fluctuations in the data. Second, I show that there is little heteroskedasticity in output; if there was substantial heteroskedasticity in the underlying shocks, one would expect that it shows up in these series as well. In fact, the little heteroskedasticity I do find is consistent with spillovers from the state dependence in investment.

Table 5: Conditional Heteroskedasticity

Statistic	**Data**	**Model**	**Benchmark**
$\log\left(\frac{\widehat{\sigma}_{90}}{\widehat{\sigma}_{10}}\right)$	.153** (.031)	.159	.017
$\log\left(\frac{\widehat{\sigma}_{75}}{\widehat{\sigma}_{25}}\right)$	.082** (.020)	.082	.008

Results from estimating equation (9) in the text. Standard errors computed using a bootstrapping procedure.

the sample than the 10^{th}. In addition, my model generates the same amount of conditional heteroskedasticity as exists in the data; when I estimate the same statistical model (9) on simulated data from my model, the resulting $\log\left(\frac{\widehat{\sigma}_{90}}{\widehat{\sigma}_{10}}\right)$ is .159, compared to .153 in the data.[26]

5 Policy Analysis

I now introduce investment stimulus policy into the model and obtain two main results. First, the effect of the policy is state dependent, becoming less powerful in recessions. Second, I develop a micro targeted policy which increases cost effectiveness more than five times compared to existing policies. But before deriving these aggregate predictions, I first show that the direct effect of policy on firm-level investment is in line with what is measured in the micro data.

5.1 Institutional Background

Since investment stimulus is mainly implemented by increasing the tax writeoff for new investment, I begin with a discussion of the US corporate tax code. According to the tax code, firms pay taxes on their revenues net of business expenses. Most of these expenses are on nondurable production inputs, like labor, energy, or materials. Since these inputs are completely used up in the fiscal year, they can be fully deducted. Capital investment, however, is durable, and so it is deducted over time. The current schedule for these deductions is given by the IRS' Modified Accelerated Cost Recovery System, or MACRS. MACRS assigns different tax depreciation schedules to different goods, in order to reflect the true economic depreciation of the capital used in production.

Historically, there have been two main implementations of investment stimulus policy. First,

[26] The little conditional heteroskedasticity that is present in the benchmark real business cycle model is again due to the dynamics of the capital stock, as discussed in footnote 20.

Table 6: 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.

the investment tax credit was commonly used before the 1986 tax reform. Second, the bonus depreciation allowance, or accelerated depreciation, was used as countercyclical stimulus in the last two recessions.

To understand how these stimulus policies work, consider the example of a firm purchasing a $1000 computer.[27] In Table 6, I reproduce the tax depreciation schedule for this $1000 purchase under three regimes: the standard MACRS schedule, which is the baseline case without stimulus policy; a 50% bonus depreciation allowance; and a 5% investment tax credit. First consider the standard tax depreciation schedule. The schedule specifies that the recovery period for a computer is five years, and the fraction of the $1000 purchase the firm writes off in each of those years. This fraction declines over time to reflect the true economic depreciation of the computer. At the end of the five years, the firm will have written off the full $1000 purchase price.[28]

Now consider how this standard schedule changes under the two investment stimulus policies. The 50% bonus depreciation allowance allows the firm to immediately deduct 50% of the $1000,

[27] This example draws heavily from Table 2 in Zwick & Mahon (2014).

[28] I abstract from potentially important features of the tax code for computational tractability. For example, firms making an operating loss or carrying previous losses forward do not have taxable income, and therefore do not face the depreciation schedule in the period of the investment.

or \$500. However, the firm can only apply the standard depreciation schedule to the remaining \$500. This does not change the total amount of the tax writeoffs, but does change the timing from the future to the present. This increases the present value of the tax deductions, again making investment more attractive.

The 5% investment tax credit reduces the firm's tax bill by 5% of the \$1000, or \$50; expressed in terms of tax writeoffs, this is $\frac{\$50}{35\%}$, where 35% is the example tax rate. The firm then applies the standard schedule to the remaining \$950. The investment tax credit thus increases both the total amount of tax deductions and the present value of these deductions, making investment more attractive.

5.2 Introducing Stimulus Policy into the Model

In the previous two examples, the present value of the tax deductions over time is a useful summary of how the various tax depreciation schedules affect the incentive to invest. In my model, the present value *completely* characterizes how the tax depreciation schedule affects the incentive to invest:

Proposition 2 *The only way the tax depreciation schedule affects a firm's investment policy rule is through the effective price* q_t*:*

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

Proof. In Appendix A. ■

Proposition 2 states that the effective price of a unit of investment is 1, the price of a unit of output, minus $\tau \times PV_t$, the present value of the depreciation allowances for that unit of investment. The reason that only the present value of deductions matters, and not the timing, is that financial markets are frictionless. In this case, tax deductions are equivalent to a riskless asset which depreciates at rate $\widehat{\delta}$.

With this result in hand, I model investment stimulus policy as an exogenous change to the present value of tax depreciation schedule. In particular, I modify the effective price to be

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

where sub_t is an implicit subsidy representing stimulus policy. The two real world examples described above map into different choices for the implicit subsidy sub_t. Specifically, the 50% bonus depreciation allowance implies that $sub_t = .5 \times (1 - PV_t)$; it is as if the firm receives the baseline depreciation schedule on all its investment, plus on one half of its investment gets an extra writeoff, equal to how much the firm values a unit of output in the current period, 1, relative to a unit of output through the depreciation schedule, PV_t. The 5% investment tax credit implies that $sub_t = .05 \times \left(\frac{1}{\tau} - PV_t\right)$; the subsidy equals how much the firm values the equivalent tax writeoff $\frac{1}{\tau}$ relative to the baseline schedule PV_t.

The subsidy sub_t follows the simple stochastic process:[29]

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

The policy shock ε_t is distributed $N(0, \sigma_{sub}^2)$. I choose $\overline{sub} = .01$, 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.[30]

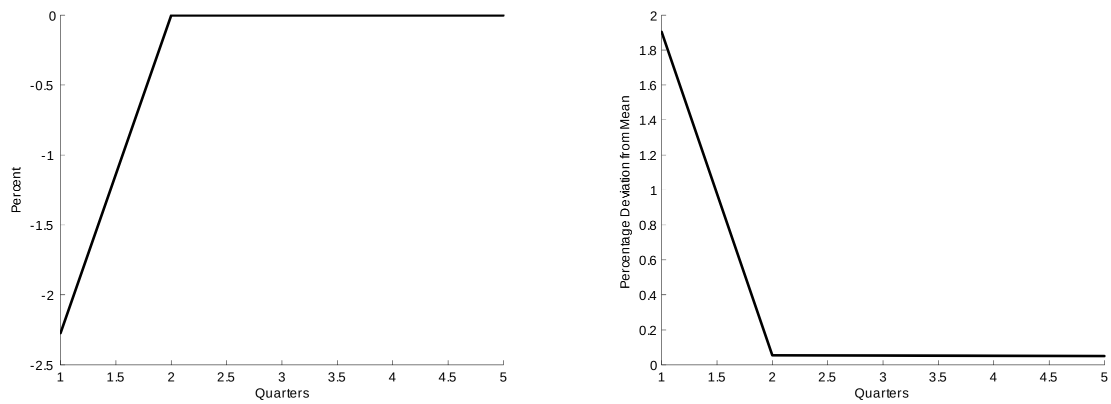
Average Impulse Response to Policy Shock In the model, stimulus policy increases aggregate investment by lowering the effective price q_t. I illustrate this process in Figure 4, which plots impulse responses to a one standard deviation policy shock, starting from steady state. The left panel shows that the shock decreases the effective price by roughly 2.5 percentage points; in the right panel, this translates into a nearly 2% increase in investment. In both cases the impulse response dies out after one period, reflecting the transitory nature of the policy shock.

[29] Given the small macro literature on investment stimulus policy, there is no consensus on the form of the policy rule. The choice here is motivated by two considerations. First, I specify the process in logs to ensure the subsidy is always positive. Second, I specify a log-linear rule to be able to use the perturbation method of my numerical algorithm. Given these two conditions, this was the simplest choice available.

[30] I also experimented with including a term correlated with aggregate productivity, to capture expectations about the countercyclical nature of the policy. However, for the results I show here, this was not quantitatively important.

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.

Taken together, the two panels of Figure 4 imply that the average semi-elasticity of aggregate investment to the subsidy is nearly 1. However, in what follows, I do not focus on this average elasticity. Instead, I focus on two new implications of my model: how this elasticity changes over the business cycle, and how it can be increased by designing a more cost-effective policy.

5.3 Model Matches Direct Effect of Policy on Firm-Level Decisions

The effect of stimulus on aggregate investment is comprised of two components: the direct effect of policy on firm-level investment decisions, and the general equilibrium effect of prices. Measuring the direct effect of policy on firm-level decisions in the data is the subject of a large literature, so as a validation exercise, I first compare this estimated direct effect in the data to the same measurement in my model.

Empirical Study: Zwick & Mahon (2014) To compare my model to the data, I reproduce a recent study by Zwick & Mahon (2014), which analyzes the effect of the bonus depreciation allowance on firm-level investment in the last two recessions. I choose this particular study for two reasons. First, it uses the same IRS microdata used to calibrate the model in Section 3. Zwick & Mahon (2014) show that the sample selection in alternative datasets leads to severe bias.

Second, Zwick & Mahon (2014)'s empirical strategy is particularly transparent, and therefore easy to interpret.

The core of Zwick & Mahon (2014)'s empirical strategy exploits variation in the implicit subsidy across firms. They formulate this implicit subsidy as

$$sub_{jt} = BDA_t \times (1 - PV_j),$$

where BDA_t is the fraction of new investment eligible for the bonus and PV_j the average present value of the baseline schedule for firm j. The present value PV_j varies because firms invest in different types of capital goods, which have different depreciation schedules. Long-lived goods, which have a low baseline present value PV_j, act as a treatment group because the bonus pulls more deductions from the future to the present. Short-lived goods, which have a high baseline present value PV_j, act as a control group because the bonus pulls fewer deductions from the future to the present.

Zwick & Mahon (2014) implement this idea through a difference in differences design. They specify the firm-level regression

$$\log i_{jt} = \alpha_i + \delta_t + \beta \times sub_{jt} + \varepsilon_{jt}, \tag{11}$$

where α_i is a firm fixed effect, δ_t is a time fixed effect, sub_{jt} is the implicit subsidy, and ε_{jt} is a residual. I use this measure of the direct effect of the policy to validate my model. As with any differences in differences exercise, the crucial identifying assumption is that "parallel trends" holds, that is, that investment growth in long-lived and short-lived goods would have been the same without the policy. Zwick & Mahon (2014) go to great lengths to argue that this is the case.

Reproducing Zwick & Mahon (2014) in the Model The main obstacle to reproducing Zwick & Mahon (2014)'s regression (11) is that my model does not have variation in the subsidy across firms. To introduce this variation, I alter my model to include two types of firms of equal mass. These firms differ only in their baseline depreciation schedule $\widehat{\delta}_j$ for $j \in \{\text{treatment}, \text{control}\}$, which controls the present value PV_j. I choose $\widehat{\delta}_{\text{treatment}}$ and $\widehat{\delta}_{\text{control}}$ to match two targets: their average matches the value in the baseline model, and their ratio reproduces the ratio of the 90^{th} percentile

Table 7: Model Matches Effect of Policy on Firm-Level Investment

	Data	**Model**	**No Fixed Cost**
Semi Elasticity	3.69** (0.53)	3.41	1.95

Notes: Estimates from the model $\log i_{jt} = \alpha_j + \delta_t + \beta \times sub_{jt} + \varepsilon_{jt}$. Data: Zwick & Mahon (2014). Model: Simulated data from the model. No fixed cost: model with $\overline{\xi} = 0$.

of PV_j to the 10^{th} percentile of PV_j in the data.[31] I feed in a history of two negative productivity shocks, pushing the economy into a recession. Starting from this point, I then feed two years of a 50% bonus, and compute the simple difference in differences coefficient $\widehat{\beta}$.

Results Table 7 shows that the estimated semi elasticity in my model is 3.41, within one standard error of Zwick & Mahon (2014)'s estimate. Therefore, my model passes this important validation test; it reproduces the direct effect of policy on firm-level decisions in the data. Lumpy investment is crucial to the model's fit; when I eliminate the fixed cost, the estimated semi elasticity falls to 1.95. Without the fixed cost, the estimate only measures the intensive margin response to the policy, which is small.[32] However, in the full model, the estimate also measures the extensive margin response to the policy, since firms with a higher subsidy are more likely to adjust. Given the importance of the extensive margin in accounting for firm-level response to other shocks, it is not surprising that it also accounts for a substantial portion of the response to the policy.

5.4 Policy Multiplier Declines in Recessions

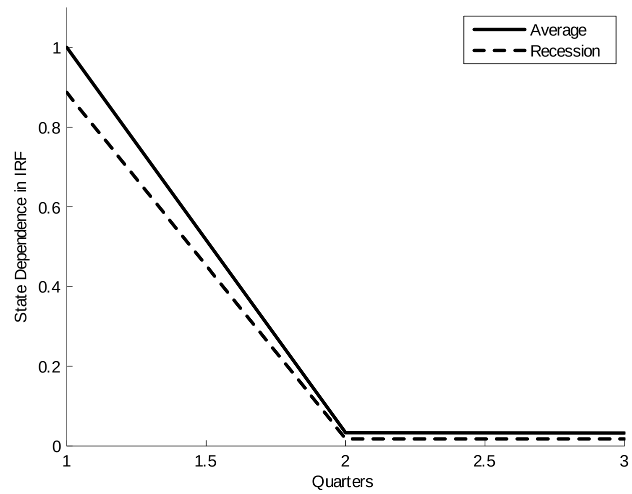
To show that the policy multiplier varies over the cycle, I feed in different histories of aggregate TFP shocks, one which leaves the economy in steady state and one which pushes the economy into a recession.[33] I choose these specific histories to compare the effect of policy on average to the effect in a recession accompanied by 6% drop in GDP, roughly comparable to the recent recession.

[31] Recall that I chose $\widehat{\delta}$ in the baseline model to reproduce the average present value in Zwick & Mahon (2014)'s data. The aggregate dynamics of this augmented model are very similar to the original model, reflecting the fact that $\widehat{\delta}_{\text{treatment}}$ and $\widehat{\delta}_{\text{control}}$ average to the $\widehat{\delta}$ used in the original.

[32] It is important to note that when I eliminate the fixed costs I do not recalibrate the remaining parameters of the model, which would presumably require a more responsive intensive margin to match the data.

[33] This the state dependence in the policy multiplier does not necessarily follow from the state dependence with respect to productivity shocks, given the different general equilibrium implications of the two shocks.

Figure 5: State Dependent Impulse Response to Policy Shock



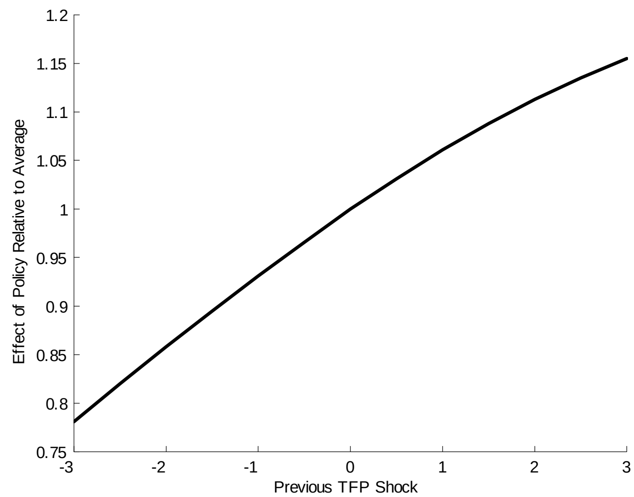
Notes: Impulse response to one standard deviation policy shock, after different history of TFP shocks. Normalized so that response upon impact in steady state is 1.

Starting from these different histories, I then compute the impulse response of aggregate investment to a one standard deviation policy shock.

In my model, the effect of the policy is substantially reduced in recessions, as shown in Figure 5. In the recession, the policy generates 15% less investment upon impact and 17% less investment in total than it would on average. The mechanism is the same as before; in the recession, fewer firms are close to their adjustment thresholds, so the policy induces fewer to adjust. More generally, this mechanism implies that the policy multiplier varies with the state of the business cycle. To show this, in Figure 6 I plot the total amount of investment generated by the policy, as a function of the previous year's shocks. The fact that the policy multiplier varies over the cycle implies that a constant multiplier cannot be used to reliably forecast the effect of the policy at a given point in the cycle. In fact, such predictions will be biased upward in recessions, with the bias increasing in the severity of the recession.[34]

[34] Standard user cost of capital or tax-adjusted q models, which are often used to forecast the effect of stimulus, are linear, which implies a constant multiplier by construction. Furthermore, these models are often estimated using policy variation from both expansion and recessions. Therefore, the estimates loosely correspond to the average effect.

Figure 6: State Dependence in Policy Multiplier

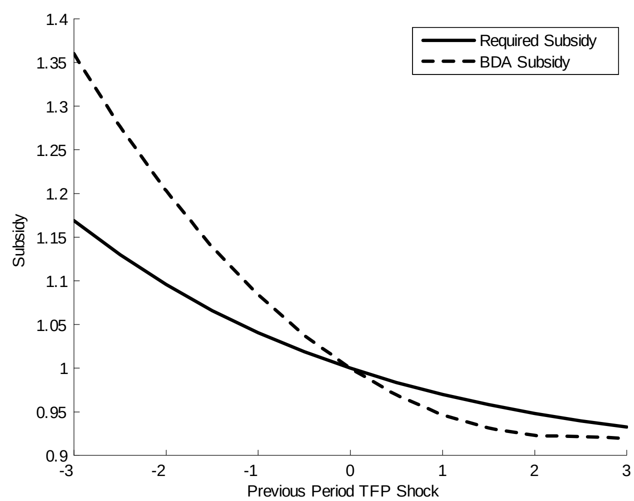


Notes: Total amount of investment generated by policy shock after different histories of TFP shocks. Normalized so that response in steady state is 1.

Bonus Depreciation Allowance Overcomes Negative State Dependence The results above show that a subsidy of a given size generates less investment in a recessions. A natural way to overcome this negative state dependence is to simply offer a larger subsidy in a recession. The solid line in Figure 7 plots the subsidy required to raise the same amount of investment as in steady state. This required subsidy is larger in recessions, when the policy multiplier is low, and smaller in expansions, when the policy multiplier is high.

In fact, the implicit subsidy offered by the bonus depreciation allowance follows just this countercyclical pattern: it is larger in recessions, when the policy multiplier is low, and smaller in expansions, when the policy multiplier is high. In particular, the implied subsidy is $sub_t = BDA \times (1 - PV_t)$, where BDA is the size of the bonus and PV_t is the present value of the baseline depreciation schedule. The dotted line in Figure 7 plots the implied subsidy from $BDA = .5$ at different points in the business cycle, and shows that the implied subsidy more than compensates for the state dependent policy multiplier. This happens because $1 - PV_t$ is countercyclical; in recessions, firms place a greater value on immediate writeoffs relative to future writeoffs, as reflected in higher interest rates used to discount future payoffs.

Figure 7: Subsidy Required to Match Average Effect and Subsidy Implied by Bonus Depreciation Allowance



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 the bonus depreciation allowance. Normalized so that subsidy in steady state is 1.

5.5 Increasing Cost Effectiveness with Micro Targeting

I now turn to investigating the cost effectiveness of stimulus policy. Most of the cost of the policies studied so far comes from subsidizing investment that would have been done even without the policy. A common prescription to reduce costs is to avoid subsidizing this inframarginal investment, while keeping the same incentives for investment that is due to the policy on the margin. In my model, the amount of inframarginal investment at the firm level is mainly determined by whether the policy induces the firm to adjust. Firms who do not adjust without the policy contribute very little to the total inframarginal costs, so a powerful way to reduce costs is to provide targeted subsidies to these firms. In principle, this micro-targeted policy could be made conditional on the entire state vector of the firm; however, I view such a policy as unrealistic, and do not pursue it.

Instead, I propose a simple micro-targeted policy which conditions on a single observable: employment. This observable is useful because in the model, as in the data, small firms grow faster than other firms in the economy. An important component of the growth of these firms is investment, so on average small firms are more likely to adjust without the policy. Conversely, medium and large firms are less likely to adjust without the policy, so I target these firms. I implement this target with the following 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 larger firms. I assume this policy is implemented for one period and is completely unexpected.

The main goal of this exercise is to illustrate the size of the potential cost savings associated with targeting the extensive margin, rather than to push for this particular size-dependent specification. In reality, other factors may affect investment by firm size that my model does not capture, and are therefore important to take into account when designing the policy. Furthermore, it is likely that firms will adjust their employment decisions to take advantage of this policy. However, these complications are only important to take into account if the potential cost savings are large.

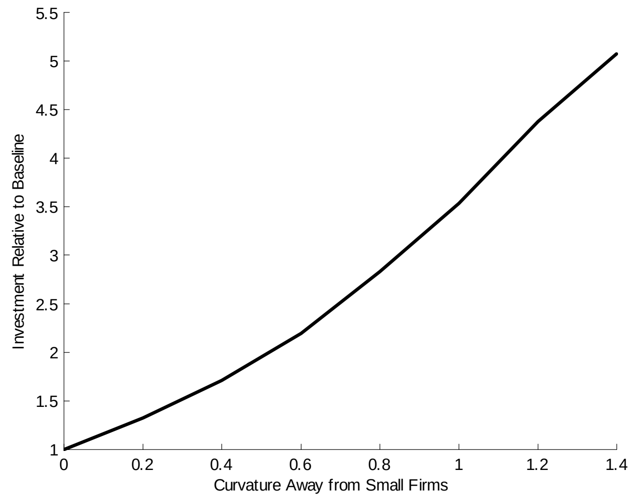
Figure 8 shows that these potential cost savings are indeed large, generating up to five time more investment than existing policies.[35] On the horizontal axis I plot a range of α_2, the weight on employment in the subsidy. Given a value of α_2, I then solve for the baseline slope α_1 which raises the same amount of revenue as the baseline size-independent subsidy. I then plot the total amount of investment generated by this combination of α_1 and α_2 on the vertical axis. Hence, increasing the weight on larger firms α_2 increases the total amount of investment generated by the micro-targeted policy.

Comparison to a Phase In To emphasize the power of this size-dependent policy, I compare its performance to a commonly suggested alternative to improve cost effectiveness: a phase in. A phase in avoids paying for inframarginal investment by only subsidizing investment above a certain threshold. In representative firm models, this phase in can completely avoid paying for inframarginal investment if the threshold is chosen correctly. I approximate this with the following subsidy per unit of investment:

$$sub_{jt} = \beta_1 \times i_{jt}^{\beta_2},$$

[35] I cannot consider higher values of α_2 because it implies such a large baseline subsidy that the tax-adjusted price of investment becomes negative. This is because for higher values of α_2, the policy is targeting a small mass of firms very heavily. To keep the policy budget equivalent, it therefore has to offer these firms a very high subsidy.

Figure 8: 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.

where β_1 controls the baseline slope and β_2 controls how quickly the subsidy is increasing in investment.[36] I then compare the amount of investment generated in the baseline case where $\beta_2 = 0$ to the case where $\beta_2 = 1$, choosing β_1 to ensure the policies are budget equivalent.[37] This specification captures the idea that higher levels of investment receive a higher subsidy.

The budget-equivalent phase in only increases cost effectiveness by 45%, less than ten times the improvement generated by my size-dependent proposal. Intuitively, the phase in targets investment along the intensive margin, while my size-dependent policy targets investment along the extensive margin. Given that the extensive margin accounts for a large share of firm-level investment, my proposal is more effective.

6 Conclusion

In this paper, I showed that accounting for lumpy investment and realistic interest rate dynamics is crucial for understanding the dynamics of aggregate investment, in response to both business cycle and policy shocks. To emphasize the role of these key features, I abstracted from others which are

[36] I model the phase in as a continuous schedule rather than a discrete cutoff for computational tractability.

[37] As in footnote 35, higher levels of β_2 again imply that the tax-adjusted price of investment becomes negative.

also potentially important. Incorporating such features is an active part of my research agenda, and I briefly discuss two here.

First, in adopting a real business cycle framework, I abstracted from the aggregate demand externality which at least partly motivated the use of stimulus policy in the first place. Incorporating this externality would likely raise the average size of the policy multiplier through the aggregate demand channel. Additionally, it would allow for a careful discussion of the welfare effects of investment stimulus. The results presented here are an important input into both of these analyses, as these results are determined by adjustment costs on the supply side of the economy.

Second, I abstracted from financial frictions, which many have argued is important in accounting for microeconomic investment behavior. Although such frictions are not required to match the data I used in this paper, it is possible 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 important, because in contrast to my results, financial frictions imply that investment stimulus is more powerful in recessions; stimulus decreases the amount of external funds required to finance investment, and external funds are more costly in recessions. Recognizing this possibility, my paper serves as an important benchmark which analyzes the effect of adjustment costs in isolation.

References

[1] Algan, Yann, Olivier Allais, and Wouter Den Haan (2008), "Solving Heterogeneous-Agent Models with Parameterized Cross-Sectional Distributions," *Journal of Economic Dynamics and Control*, 32(3), 875-908.

[2] 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.

[3] Bachmann, Rudiger and Lin Ma (2012), "Lumpy Investment, Lumpy Inventories," *NBER Working Paper No 17924*.

[4] Basu, Susanto, Miles Kimball, and John Fernald (2006), *American Economic Review*, 96(5), 1418-1448.

[5] 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.

[6] Berger, David and Joseph Vavra (2014), "Consumption Dynamics During Recessions," *Econometrica (forthcoming).*

[7] Bertola, Giuseppe, and Ricardo Caballero (1990), "Kinked Adjustment Costs and Aggregate Dynamics," *NBER Macroeconomics Annual 1990, Volume 5*, 237-296.

[8] Bertola, Giuseppe, and Ricardo Caballero (1994), "Irreversibility and Aggregate Investment," *The Review of Economic Studies*, 61(2), 223-246.

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

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

[11] Caballero, Ricardo and Eduardo Engel (1993), "Microeconomic Adjustment Hazards and Aggregate Dynamics," *The Quarterly Journal of Economics,* 108(2), 359-383.

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

[13] 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.

[14] Chen, Andrew (2014), "Precautionary Volatility and Asset Prices," *Unpublished Manuscript.*

[15] Cooley, Timothy and Edward Prescott, editors (1995), *Frontiers of Business Cycle Research.*

[16] Cooper, Russell, John Haltiwanger, and Laura Power (1999), "Machine Replacement and the Business Cycle: Lumps and Bumps," *American Economic Review*, 921-946.

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

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

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

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

[21] Goolsbee, Austan (1998), "Investment Tax Incentives, Prices, and the Supply of Capital Goods," *The Quarterly Journal of Economics*, 113(1), 121-148.

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

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

[24] Hall, Robert and Dale Jorgenson (1967), "Tax Policy and Investment Behavior," *The American Economic Review*, 391-414.

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

[26] Jaimovich, Nir and Sergio Rebelo (2009), "Can News About the Future Drive the Business Cycle?," *The American Economic Review*, 99(4), 1097-1118.

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

[28] Judd, Kenneth (1998), *Numerical Methods in Economics*, MIT Press.

[29] Khan, Aubhik and Julia Thomas (2003), "Nonconvex Factor Adjustments in Equilibrium Business Cycle Models: Do Nonlinearities Matter?," *Journal of Monetary Economics,* 50(2), 331-360.

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

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

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

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

[34] Krusell, Per and Anthony Smith (2006), "Quantitative Macroeconomic Models with Heterogeneous Agents," *Econometric Society Monographs*, 41, 298.

[35] Lettau, Martin and Harald Uhlig (2001), "Can Habit Formation be Reconciled with Business Cycle Facts?", *Review of Economic Dynamics*, 3(1), 79-99.

[36] Prescott, Edward (1986), "Theory Ahead of Business Cycle Measurement," in *Carnegie-Rochester Conference Series on Public Policy* (Vol. 25, pp. 11-44), North-Holland.

[37] Prescott, Edward (2003), "Non-Convexities in Quantitative General Equilibrium Studies of Business Cycles," No. 312 Federal Reserve bank of Minneapolis.

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

[39] Romer, Christina and Jared Bernstein (2009), "The Job Impact of the American Recovery and Reinvestment Plan."

[40] Sims, Christopher (2002), "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1), 1-20.

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

[42] Veracierto, Marcelo (2002), "Plant-level Irreversible Investment and Equilibrium Business Cycles," *The American Economic Review*, 92(1), 181-197.

[43] Young, Eric (2010), "Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell-Smith Algorithm and Non-Stochastic Simulations," *Journal of Economic Dynamics and Control*, 34(1), 36-41.

[44] Zwick, Eric and James Mahon (2014), "Do Financial Frictions Amplify Fiscal Policy? Evidence from Business Investment Stimulus," *Unpublished manuscript.*

A Characterizing Equilibrium

In this Appendix, I characterize the recursive competitive equilibrium defined in Section 2.4. I use this characterization to numerically solve the model in Appendix B, and to model investment stimulus policy in Proposition 2 of the main text. To keep the description manageable, I proceed in two steps. First, in Appendix A.1, I simplify the firm's decision problem, taking all prices as given. Second, in Appendix A.2, I embed the firm's decision problem into a characterization of the equilibrium.

A.1 Firm's Decision Problem

I simplify the firm's decision problem through a series of three propositions. These propositions eliminate two individual state variables from the problem, which is crucial for the feasibility of the numerical algorithm.

To ease notation, I define

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

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 3 *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*

$$v^1(\varepsilon, k, \xi; \mathbf{s}) = \pi(\varepsilon, k; \mathbf{s}) + \max_i \left\{ \begin{array}{c} -(1 - \tau PV(\mathbf{s}))\, i - \frac{\phi}{2}\left(\frac{i}{k}\right)^2 k - \xi w(\mathbf{s}) 1\{i \notin [-ak, ak]\} \\ +E[\Lambda(z'; \mathbf{s}) v^1\left(\varepsilon', (1-\delta)\, k + i, \xi'; \mathbf{s}'\right)] \end{array} \right\} \quad (12)$$

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 being by showing that the operator T defined by the right hand side of (3) 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:

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

Collecting terms,

$$\begin{aligned} T(f)(\varepsilon, k, \xi; \mathbf{s}) &= \pi(\varepsilon, k; \mathbf{s}) + \tau\left(\widehat{\delta} + (1-\widehat{\delta})E[\Lambda(z'; \mathbf{s})PV(\mathbf{s}')]\right) d \quad (13) \\ &+ \max_i \left\{ \begin{array}{c} -\left(1 - \tau\widehat{\delta} - \tau(1-\widehat{\delta})E[\Lambda(z'; \mathbf{s})PV(\mathbf{s}')]\right) i - \frac{\phi}{2}\left(\frac{i}{k}\right)^2 k \\ -\xi w(\mathbf{s}) 1\{i \notin [-ak, ak]\} + E[\Lambda(z'; \mathbf{s}) f(\varepsilon', (1-\delta)k + i, \xi'; \mathbf{s}')] \end{array} \right\} \end{aligned}$$

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

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

Plugging this back into (13) and rearranging gives

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

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 (3) to get

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

Rearranging terms as before shows that

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

Subtracting $\tau PV(\mathbf{s})d$ from both sides establishes (12). ■

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:

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

Proposition 4 *The investment decision rule is independent of d and given by*

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

where

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

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

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

Proof. The form of $i^a(\varepsilon, k; \mathbf{s})$ follows directly from the Bellman equation (12), 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 (12), 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 5 $v^0(\varepsilon, k; \mathbf{s})$ *solves the Bellman equation*

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

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 1 and the form of the policy function from Proposition 2. ■

A.2 A Characterization of the Equilibrium

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

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

This 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 6 *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*

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

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

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*

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

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

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

and $PV(\mathbf{s})$ *is defined by the recursion* $p(\mathbf{s})PV(\mathbf{s}) = p(\mathbf{s})\widehat{\delta} + \left(1 - \widehat{\delta}\right)\beta E\left[p(\mathbf{s}')PV(\mathbf{s}')|\mathbf{s}\right]$.

2. *(Labor market clearing)*

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

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

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

4. *(Laws of motion)*

$$\begin{aligned} S'_{-1}(\mathbf{s}) &= S(\mathbf{s}) \\ C'_{-1}(\mathbf{s}) &= C(\mathbf{s}) \end{aligned}$$

5. *(Law of motion for measure) For all measurable sets* $\Delta_\varepsilon \times \Delta_k$,

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

Proof. Condition 1 follows from Propositions 1-3, 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 2 follows from the household's FOC, the firms' FOC, and labor market clearing. Condition 3 follows from output market clearing and the definition of $p(\mathbf{s})$. Condition 4 directly reproduces conditions 4(c) and 4(d) from the main text. Condition 5 follows from the original law of motion in condition 4(e) in the main text, eliminating d as an individual state variable and integrating out ξ. ■

B Solution Algorithm

In this Appendix, I detail the solution algorithm outlined in Section 2.5 of the main text. This description draws heavily from the characterization of equilibrium presented in Appendix A. As in the text, I proceed in three main steps. First, I represent all equilibrium objects with finite dimensional approximations. Second, I solve for the steady state. Third, I solve for the dynamics around steady state.

Before describing these steps in detail, I switch from the aggregate state space representation of the equilibrium in Proposition 6 to a sequence representation of the equilibrium. This does not change the equilibrium dynamics, but makes the presentation of the algorithm much simpler. I keep the state space representation with respect to individual states. This mixture of state space and sequence representations of equilibrium objects is at the core of the mixed global and local approximation I present below.

For each t, the equilibrium objects $v_t(\varepsilon,k)$, w_t, p_t, S_t, C_t, and $\mu_t(\varepsilon,k)$ must satisfy:

1. (Firm optimization)

$$\begin{aligned} v_t(\varepsilon,k) &= p_t\pi_t(\varepsilon,k) \\ &+\frac{\widehat{\xi}_t(\varepsilon,k)}{\bar{\xi}}\left\{\begin{array}{c} -p_t\left(1-\tau PV_t\right)i_t^a(\varepsilon,k)-p_t\frac{\phi}{2}\left(\frac{i_t^a(\varepsilon,k)}{k}\right)^2 k \\ -p_t\frac{\widehat{\xi}_t(\varepsilon,k)}{2}w_t+\beta E_t[E_{\varepsilon'|\varepsilon}[v_{t+1}(\varepsilon',(1-\delta)k+i_t^a(\varepsilon,k))]] \end{array}\right\} \\ &+\left(1-\frac{\widehat{\xi}_t(\varepsilon,k)}{\bar{\xi}}\right)\left\{\begin{array}{c} -p_t\left(1-\tau PV_t\right)i_t^n(\varepsilon,k)-p_t\frac{\phi}{2}\left(\frac{i_t^n(\varepsilon,k)}{k}\right)^2 k \\ +\beta E_t[E_{\varepsilon'|\varepsilon}[v_{t+1}(\varepsilon',(1-\delta)k+i_t^n(\varepsilon,k))]] \end{array}\right\} \end{aligned} \tag{14}$$

where $i_t^a(\varepsilon,k)$, $i_t^n(\varepsilon,k)$, and $\widehat{\xi}_t(\varepsilon,k)$ are derived from v_{t+1} analogously to before. Note that in this representation, I have separated the expectation with respect to the idiosyncratic shock ε' from the expectation with respect to the aggregate shock z_{t+1}.

2. (Labor market clearing)

$$\left(\frac{w_t}{\chi}\right)^{\frac{1}{\eta}} = \int \left(n_t(\varepsilon, k) + \frac{\widehat{\xi}_t(\varepsilon, k)^2}{2\bar{\xi}}\right) \mu_t(d\varepsilon, dk)$$

where $n_t(\varepsilon, k) = \left(\frac{e^z e^\varepsilon k^\theta \nu}{w_t}\right)^{\frac{1}{1-\nu}}$.

3. (Consistency)

$$p_t = \left(C_t \times S_t - \chi \frac{\left(\left(\frac{w_t}{\chi}\right)^{\frac{1}{\eta}}\right)^{1+\eta}}{1+\eta}\right)^{-\sigma}$$

where C_t and S_t are consistent with firms' decisions analogously to before.

4. (Law of motion for measure) For all measurable sets $\Delta_\varepsilon \times \Delta_k$,

$$\begin{aligned}\mu_{t+1}(\Delta_\varepsilon \times \Delta_k) &= \int p(\varepsilon' \in \Delta_\varepsilon|\varepsilon)(\frac{\widehat{\xi}_t(\varepsilon, k)}{\bar{\xi}} 1\{(1-\delta)k + i_t^a(\varepsilon, k) \in \Delta_k\} + \\ &\left(1 - \frac{\widehat{\xi}_t(\varepsilon, k)}{\bar{\xi}}\right) 1\{(1-\delta)k + i_t^n(\varepsilon, k) \in \Delta_k\} d\varepsilon' \mu_t(d\varepsilon, dk)\end{aligned}$$

B.1 Step 1: Approximate Equilibrium with Finite Dimensional Objects

The first step of my algorithm is to represent the equilibrium objects using finite-dimensional approximations. Some objects – w_t, p_t, C_t, and S_t – are scalars, and so do not need to be approximated at all. However, both $v_t(\varepsilon, k)$ and $\mu_t(\varepsilon, k)$ are infinite-dimensional. Below I discuss how I approximate these objects, but the general approach does not rely on these particular choices.

Value Function I approximate the value function $v_t(\varepsilon, k)$ with a weighted sum of orthogonal polynomials:

$$v_t(\varepsilon, k) = \sum_{i,j=1,1}^{n_\varepsilon, n_k} \theta_t^{ij} C^i(\varepsilon) C^j(k)$$

where θ_t^{ij} is a scalar coefficient, $C^i(\cdot)$ is the Chebyshev polynomial of order i, and n_ε, n_k are the orders of approximation. Denote the vector of the coefficients by $\boldsymbol{\theta}_t$. See Judd (1998) for a detailed discussion of the properties of Chebyshev polynomials.

To compute the Bellman equation, I need to compute the expectation over idiosyncratic shocks

ε' (the expectation over aggregate shocks will be computed implicitly in the third step). To do so, I first transform this to an expectation over ω_ε, the innovation in the AR(1) process $\varepsilon' = \rho_\varepsilon \varepsilon + \omega_\varepsilon$. Since ω_ε is normally distributed, I approximate the integral using Gaussian quadrature. This quadrature scheme specifies nodes ω_ε^m and weights w_ε^m such that, for any function $g(\omega_\varepsilon)$, $\int g(\omega_e)p(\omega_\varepsilon)d\omega_\varepsilon \approx \sum_{m=1}^{n_m} w_\varepsilon^m g(\omega_\varepsilon^m)$, where p is the p.d.f. of ω_ε and n_m is the order of the quadrature. Again, see Judd (1998) for details of Gaussian quadrature.

To pin down the coefficients $\boldsymbol{\theta}_t$, I use a collocation scheme. This scheme imposes that the Bellman equation (14) holds exactly at a set of grid points $\{\varepsilon_o, k_p\}_{o,p=1,1}^{n_\varepsilon, n_k}$, i.e.,

$$\sum_{i,j=1,1}^{n_\varepsilon, n_k} \theta_t^{ij} C^i(\varepsilon_o) C^j(k_p) = p_t \pi_t(\varepsilon_o, k_p)$$

$$+\frac{\widehat{\xi}_t(\varepsilon_o, k_p)}{\bar{\xi}} \left\{ \begin{array}{c} -p_t(1-\tau PV_t)\, i_t^a(\varepsilon_o, k_p) - p_t \frac{\phi}{2}\left(\frac{i_t^a(\varepsilon_o,k_p)}{k_p}\right)^2 k_p \\ -p_t \frac{\widehat{\xi}_t(\varepsilon_o,k_p)}{2} w_t + \beta E_t[\sum_{m=1}^{n_m} w_\varepsilon^m \sum_{i,j=1,1}^{n_\varepsilon,n_m} \theta_{t+1}^{ij} C^i(\rho_\varepsilon \varepsilon_o + \omega_\varepsilon^m) C^j((1-\delta)k_p + i_t^a(\varepsilon_o, k_p))] \end{array} \right\}$$

$$+\left(1 - \frac{\widehat{\xi}_t(\varepsilon_o, k_p)}{\bar{\xi}}\right) \left\{ \begin{array}{c} -p_t(1-\tau PV_t)\, i_t^n(\varepsilon_o, k_p) - p_t \frac{\phi}{2}\left(\frac{i_t^n(\varepsilon_o,k_p)}{k_p}\right)^2 k_p \\ +\beta E_t[\sum_{m=1}^{n_m} w_\varepsilon^m \sum_{i,j=1,1}^{n_\varepsilon,n_m} \theta_{t+1}^{ij} C^i(\rho_\varepsilon \varepsilon_o + \omega_\varepsilon^m) C^j((1-\delta)k_p + i_t^n(\varepsilon_o, k_p))] \end{array} \right\}$$

Distribution I approximate the distribution of firms $\mu_t(\varepsilon, k)$ using a mixture of normals over $(\varepsilon, \log k)$ pairs. This mixture distribution is characterized by a p.d.f. of the form:

$$p_t(\varepsilon, \log k) = \sum_{i=1}^{n_\mu} \pi_{it} p_i\left(\varepsilon, \log k | \mathbf{m}_t^i, \boldsymbol{\Sigma}_t^i\right)$$

where $p_i(\varepsilon, \log k | \mathbf{m}, \boldsymbol{\Sigma})$ is the p.d.f. of a normal with mean $\mathbf{m}$ and covariance matrix $\boldsymbol{\Sigma}$, π_{it} is the weight on component i, and $\mathbf{m}_t^i$ and $\boldsymbol{\Sigma}_t^i$ are the mean and covariance matrices of component i. Hence, the distribution is completely characterized by the vector $\boldsymbol{\mu}_t = \left(\{\pi_{it}\}_{i=1}^{n_\mu}, \{\mathbf{m}_t^i\}_{i=1}^{n_\mu}, \{\boldsymbol{\Sigma}_t^i\}_{i=1}^{n_\mu}\right)$.

With this mixture approximation, the integrals in the integrals in Conditions 2 and 3 can be computed using Gaussian quadrature. As before, this quadrature scheme specifies nodes $\left\{\left(\varepsilon_m^i, \log k_m^i\right)\right\}_{m=1}^{n_m}$ and weights $\left\{w_\mu^m\right\}_{m=1}^{n_m}$ for each component i. This implies that, for any function

$g(\varepsilon, \log k)$,

$$\int g(\varepsilon, \log k) p_t(\varepsilon, \log k)\, d\varepsilon d\log k \approx \sum_{i=1}^{n_\mu} \pi_{it} \sum_{m=1}^{n_m} w_\mu^m g(\varepsilon_m^i, \log k_m^i)$$

I compute the law of motion for the measure in Condition 4 using a simple moment matching scheme. Specifically, using the current period's measure $\boldsymbol{\mu}_t$ and decision rules derived from $\boldsymbol{\theta}_t$, I compute moments of next period's distribution $\mu_{t+1}(\varepsilon, k)$ using Gaussian quadrature and Proposition 7 below. I then choose the coefficients $\boldsymbol{\mu}_{t+1}$ to match these implied moments.

Proposition 7 *Let the $m(\varepsilon', \log k')$ be an integrable function of state variables $(\varepsilon', \log k')$. Then*

$$\int m(\varepsilon', \log k')\, p_{t+1}(\varepsilon', \log k')\, d\varepsilon' d\log k' \tag{15}$$

$$= \int \left\{ \begin{array}{c} \left(\frac{\widehat{\xi}_t(\varepsilon, \log k)}{\bar{\xi}}\right) g(\rho_\varepsilon \varepsilon + \omega'_\varepsilon, \log((1-\delta)k + i_t^a(\varepsilon, k))) \\ + \left(1 - \frac{\widehat{\xi}_t(\varepsilon, \log k)}{\bar{\xi}}\right) g(\rho_\varepsilon \varepsilon + \omega'_\varepsilon, \log((1-\delta)k + i_t^n(\varepsilon, k))) \end{array} \right\} p(\omega'_\varepsilon) d\omega'_\varepsilon p_t(\varepsilon, \log k)\, d\varepsilon d\log k$$

where, as before, ω'_ε is the innovation to the AR(1) process of idiosyncratic productivity $\varepsilon' = \rho_\varepsilon \varepsilon + \omega'_\varepsilon$, $p(\omega'_\varepsilon)$ is the p.d.f. of ω'_ε, $p_{t+1}(\varepsilon, \log k)$ is the p.d.f. of firms over state variables in period $t+1$, and $p_t(\varepsilon, \log k)$ is the p.d.f. of firms over state variables in period t.

Proof. By definition,

$$p_{t+1}(\varepsilon', \log k') = \int 1\{\rho_\varepsilon \varepsilon + \omega'_\varepsilon = \omega'\} \times \left[\begin{array}{c} \left(\frac{\widehat{\xi}_t(\varepsilon, \log k)}{\bar{\xi}}\right) 1\{\log k^a(\varepsilon, \log k) = \log k'\} \\ + \left(1 - \frac{\widehat{\xi}_t(\varepsilon, \log k)}{\bar{\xi}}\right) 1\{\log k^n(\varepsilon, \log k) = \log k'\} \end{array} \right]$$
$$\times p(\omega'_\varepsilon)\, d\omega'_\varepsilon p_t(\varepsilon, \log k)\, d\varepsilon d\log k$$

where $k^a(\varepsilon, \log k) = (1-\delta) k + i^a(\varepsilon, k)$ and $k^n(\varepsilon, \log k) = (1-\delta)k + i^n(\varepsilon, k)$. We want an expression for $\int m(\varepsilon', \log k')\, p_{t+1}(\varepsilon', \log k')\, d\varepsilon' d\log k'$. Plugging in the expression for $p_{t+1}(\varepsilon', \log k')$ above gives

$$
\begin{aligned}
&\int m\left(\varepsilon', \log k'\right) p_{t+1}\left(\varepsilon', \log k'\right) d\varepsilon' d\log k' \\
= &\int m\left(\varepsilon', \log k'\right) \times 1\left\{\rho_{\varepsilon}\varepsilon + \omega'_{\varepsilon} = \omega'\right\} \times \left[\begin{array}{c} \left(\frac{\widehat{\xi}_t(\varepsilon, \log k)}{\bar{\xi}}\right) 1\left\{\log k^a(\varepsilon, \log k) = \log k'\right\} \\ + \left(1 - \frac{\widehat{\xi}_t(\varepsilon, \log k)}{\bar{\xi}}\right) 1\left\{\log k^n(\varepsilon, \log k) = \log k'\right\} \end{array}\right] \\
&\times d\varepsilon' d\log k' \times p\left(\omega'_{\varepsilon}\right) d\omega'_{\varepsilon} p_t\left(\varepsilon, \log k\right) d\varepsilon d\log k
\end{aligned}
$$

Then integrate out ε' and $\log k'$ to arrive at (15). ■

Canonical Form With these approximations, the equilibrium can be written in the canonical form of Schmitt-Grohe & Uribe (2004):

$$E_t F(\mathbf{y}_{t+1}, \mathbf{y}_t, \mathbf{x}_{t+1}, \mathbf{x}_t) = 0 \tag{16}$$

where $\mathbf{y}_t$ is a vector of controls (i.e., non-predetermined variables), $\mathbf{x}_t$ is a vector of states (i.e., predetermined variables), and F collects the Conditions 1 - 4. In my model, the predermined variables are the aggregate shock z_t, the previous period's habit variables S_{t-1} and C_{t-1}, and the distribution $\boldsymbol{\mu}_t$, so $\mathbf{x}_t = (z_t, S_{t-1}, C_{t-1}, \boldsymbol{\mu}_t)$. The non-predetermined variables are the value function $\boldsymbol{\theta}_t$, prices w_t and p_t, and aggregates C_t and S_t, so that $\mathbf{y}_t = (\boldsymbol{\theta}_t, w_t, p_t, C_t, S_t)$.

B.2 Step 2: Solve for Steady State

The second step of my algorithm is to solve for the steady state of the model where $z_t = 0$ for all t. In terms of the canonical form (16), this amounts to solving for constant vectors $\mathbf{x}^*$ and $\mathbf{y}^*$ such that

$$F(\mathbf{y}^*, \mathbf{y}^*, \mathbf{x}^*, \mathbf{x}^*) = 0. \tag{17}$$

Note that in this steady state, there is still a full set of idiosyncratic shocks ε.

I solve for the $\mathbf{x}^*$ and $\mathbf{y}^*$ using the following iterative scheme.[38] First, I guess a wage w^*. I then solve the firm's problem to compute the coefficients $\boldsymbol{\theta}^*$.[39] Using this, compute the firms'

[38] I could directly solve the nonlinear equation (17), but doing so directly is unstable.

[39] Note that I do not need to know p^* to solve this problem, since it only appears through the stochastic discount

decision rules along the quadrature grid, and use these values to compute the invariant distribution $\boldsymbol{\mu}^*$ by iterating on the mapping defined by (15). I then compute aggregate labor demand from Condition 2, compute aggregate labor supply from $N^* = \left(\frac{w^*}{\chi}\right)^{\frac{1}{\eta}}$, and update the wage accordingly. Once this iteration converges on the market clearing wage, I can back out the other state state objects.[40]

B.3 Step 3: Solve for Dynamics Around Steady State

The third and final step of my algorithm is to solve for the dynamics of the model around steady state using perturbation techniques. Following Schmitt-Grohe & Uribe (2004), I denote the solution

$$\begin{aligned}\mathbf{y}_t &= g(\mathbf{x}_t;\sigma) \\ \mathbf{x}_{t+1} &= h(\mathbf{x}_t;\sigma) + \sigma\Upsilon\omega_{z,t+1}\end{aligned}$$

where σ is the perturbation parameter, Υ is a known matrix, and g and h are unknown functions to be approximated. I approximate these functions using Taylor expansions, and show how to solve for the coefficients of these Taylor expansions from the partial derivatives of F. This is a standard perturbation problem, and I refer the reader to Schmitt-Grohe & Uribe (2004) for more details.

To simulate the model, I do not directly use the dynamics implied by g and h. Instead, following Krusell & Smith (1997), I use g to forecast future values of $\mathbf{x}_{t+1}$, but compute prices to clear markets exactly. See Krusell & Smith (1997) for a discussion of this approach to simulation. Using this approach, the locally accurate dynamics implied by g and h have the same interpretation as a Krusell & Smith (1997) forecasting rule.

B.4 Details of the Implementation

I now provide some details of my exact implementation. I set the order of approximation n_ε to 6 and n_k to 8, which ensures that individual problems are solved accurately; increasing the order does not change the results. I set the order of the quadrature to $n_m = 5$, which exactly integrates

factor, which is just $\beta = \frac{\beta p^*}{p^*}$ in steady state.

[40] Although I have decribed this as an iteration on w^*, in practice it is more efficient to view this as a root finding problem in w^*.

Table 8: 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.

any polynomial of order less than 10; again, increasing the order does not change the results. I set the number of components $n_\mu = 1$, i.e., a single component normal. This provides an accurate approximation of the measure by two counts. First, higher order moments of the distribution are very close to those implied by the normal. Second, in steady state, the single component normal provides an accurate approximation to a fully nonparametric histogram. In the third step, I implement a first order approximation using Chris Sims' gensys solver. Most derivatives are taken by hand, though some are taken numerically using forward differences.

In principle, agents forecast the entire distribution $\boldsymbol{\mu}_{t+1}$. However, it is difficult to assess the accuracy of forecasts over individual elements of $\boldsymbol{\mu}_{t+1}$ in terms of economically meaningful units. Therefore, in Table 8 I report the accuracy of the two prices w_t and p_t, which are of course derived from the distribution.

C Calibration to Compustat Micro Data

As I discussed in Section 3, the IRS microdata used to calibrate the model has a significant measurement issue: its measure of investment only contains equipment goods, while its measure of capital contains both equipment and structures. In this Appendix, I show that the main results of the paper are robust to calibrating instead to Compustat microdata, which does not suffer from this measurement issue. First, I describe the Compustat data in more data. I then recalibrate the model to this data, continuing to target the same features of real interest rate dynamics described in the main text. Finally, I show that the recalibrated model has similar implications for the three main results in the paper: aggregate investment is more responsive to productivity shocks in expansions, the policy multiplier declines in recessions, and micro-targeting can significantly increase

Table 9: Alternative Data Targets

Micro Investment		
Target	**Data**	**Model**
Inaction rate (%)	5.3%	5.9%
Spike rate (%)	14.1%	14.1%
Positive investment rates (%)	74.9%	78.6%
Negative investment rates (%)	3.8%	1.3%
Negative spike rate (%)	1.1%	0%
Average investment rate (%)	2.6%	2.6%
Standard deviation of investment rates	.033	.053
Interest Rate Dynamics		
Target	**Data**	**Model**
Standard deviation of interest rates (%)	.12%	.12%
Correlation of interest rate and output	−.21	−.21

Notes: Moments targeted in calibration. Micro investment targets from firm-level Compustat data, 1990 - 2010. Inaction rate is amount investment rate less between -.25% and .25%. Spike rate is amount of investment rate greater than 5%. Positive investment is investment rate between .25% and 5%. Negative investment is investment rate between -.25% and -5%. Negative spike rate is investment rate less than -5%. Interest rate is return on 90-day treasury bill adjusted for realized inflation. Output is real GDP. Both series are 1953-2012, projected on lags of TFP, logged, and HP filtered.

cost effectiveness relative to existing policies.

C.1 Compustat Micro Data

The alternative firm-level investment moments are drawn from quarterly Compustat data, 1990 - 2010. As in the main text, I focus on the distribution of investment rates, which is investment divided by capital stock. I measure investment as capital expenditures minus sales, and measure capital as the book value of the capital stock. Because there is no entry or exit in the model, I use a balanced panel.

I target the seven moments from the Compustat data listed in the top panel of Table 9. The first five correspond to a coarse histogram of investment rates, similarly to the main text.[41] Since

[41] Besides sample coverage, there are two main differences between the Compustat data and the IRS data used in

Compustat contains mainly large firms, this histogram is considerably less lumpy than the IRS data used in the main text; there are less firms undergoing an investment spike (14% of observations, compared to 15% in the IRS data) and less firms inactive (5% of observations, compared to 25% in the IRS data). Finally, I also match the mean and standard deviation to ensure the model fits basic features of the distribution.

C.2 Calibration Results

Table 9 shows a rough recalibration of the model, which fits the Compustat investment moments and real interest rate moments reasonably well. It captures the amount of inactive observations and positive spikes, which are indicative of the lumpiness of investment in Compustat. However, it overpredicts the amount of observations with positive investment and underpredicts the amount with negative investment. Furthermore, it overpredicts the dispersion in investment rates. Finally, the fit to interest rate dynamics is almost exact.

Table 10 shows the parameter values that result from this rough recalibration, and compares them to the original calibration. Consistent with the fact that Compustat is less lumpy than the IRS data, the fixed cost are half the size and the size of the no fixed cost region is twice as large. The variance of the idiosyncratic productivity shocks is higher to match the amount of negative investment, but also overpredicts the dispersion of investment rates.

C.3 Robustness of Main Results

Figure 9 shows that the main results are robust to this rough recalibration. Analogously to Figure 2, panel (a) plots the impulse response upon impact of aggregate investment to a productivity shock, as a function of the three previous TFP shocks. The recalibrated model exhibits roughly the same amount of state dependence as in the main text. Analogously to Figure 6, panel (b) plots the impulse response upon impact to a policy shock, as a function of the previous year's TFP shocks. The recalibrated model also exhibits a procyclical policy multiplier, although the variation is smaller than in the main text. Finally, and analogously to Figure 8, panel (c) plots

the main text. First, since Compustat records sales of capital, the histogram includes two bins of disinvestment. The IRS data only records capital expenditure and thus does not record disinvestment. However, as shown in Table 9, disinvestment is rare, accounting for less than 5% of the Compustat observations. Second, Compustat is recorded at a quarterly frequency, while the IRS data is annual. I scale the size of the Compustat bins to take this into account.

Table 10: Fitted Parameter Values

Micro Heterogeneity			
Parameter	**Description**	**Alternative**	**Original**
$\overline{\xi}$	Upper bound on fixed costs	.25	.44
a	Size of no fixed cost region	.006	.003
ϕ	Quadratic adjustment cost	2.07	2.69
ρ_ε	Idiosyncratic productivity AR(1)	.97	.94
σ_ε	Idiosyncratic productivity AR(1)	.034	.026
Interest Rate Dynamics			
Parameter	**Description**	**Alternative**	**Original**
$\overline{S}$	Average surplus consumption	.65	.65
$\rho_{\overline{S}}$	Autocorrelation surplus consumption	.96	.95

Notes: Alternative: parameters chosen to match moments in Table 9. Original: parameters chosen to match moments in Table 2, as in main text.

the total amount of investment generated by the micro targeting policy, as a function of the weight on medium and large scale firms. The micro-targeting increases cost effectiveness up to three and a half times, which is smaller than the five times predicted in the main text.

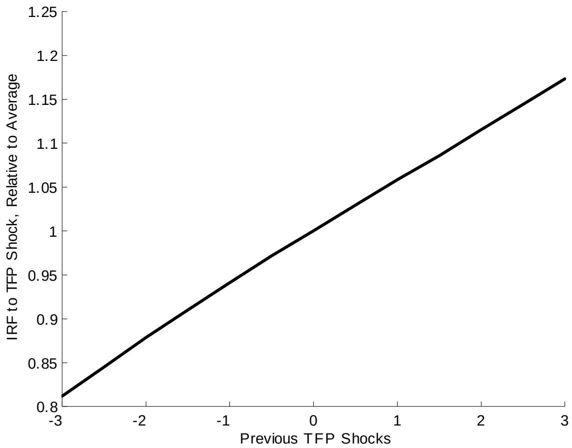
D Unconditional Second Moments

In this Appendix, I show that my model performs equally well as a benchmark real business cycle model with respect to unconditional second moments of aggregate series, as claimed in Section 4.1 of the main text. Specifically, I compare prediction about output, consumption, and hours. In the data, I measure output as real gross domestic product from BEA Table 1.1.6. I measure consumption as the sum of nondurable goods plus flow from durable services, again from BEA Table 1.1.6. I measure 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 1600.

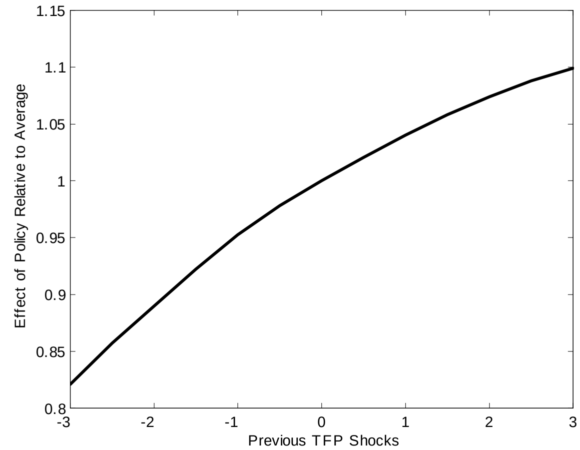
As can be seen in Table 4, my model matches these second moments as well as the benchmark

Figure 9: Robustness of Main Results to Alternative Calibration

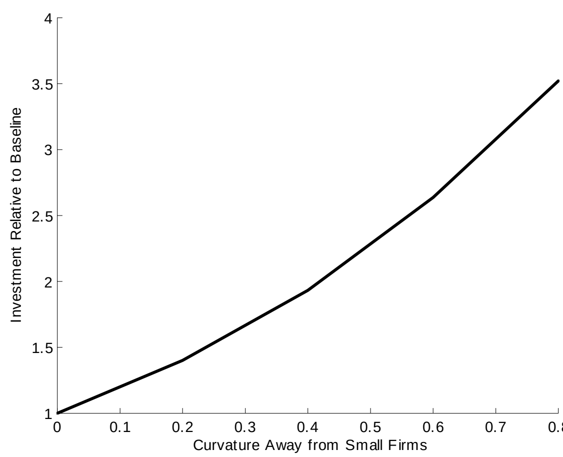
(a) State dependence w.r.t. productivity



(b) State dependence w.r.t. policy



(c) Micro-targeting policy proposal



Notes: Main results of paper, using alternative calibration to Compustat micro data in Table 9. (a) State dependence with respect to productivty shocks, analogous to Figure 2 in main text. (b) State dependence with respect to policy shocks, analogous to Figure 6 in main text. (c) Micro-targeting policy proposal, analogous to Figure 8 in main text.

Table 11: 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
(c) Correlation with Output							
Statistic	**Data**	**Model**	**RBC**				
$\rho(C, Y)$	.84	.99	.99				
$\rho(I, Y)$	.80	.99	.99				
$\rho(H, Y)$	.87	.99	.99				

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

real business cycle model. Both models reproduce the overall volatility of output, as well as the relative ranking of the volatility of consumption and investment. Both models underpredict the volatility of hours, which is a well known problem with the real business cycle framework (see Cooley & Prescott 1995).[42] The aggregate series are too highly correlated with output in both models, reflecting the fact that aggregate TFP is the only business cycle shock. Finally, both models underpredict the autocorrelation of the aggregate series.

E Role of Interest Rate Dynamics

In this Appendix, I discuss the role of real interest rate dynamics in explaining why my model generates state dependence while Khan & Thomas (2008)'s does not. In both models, the real interest rate decentralizes two key forces: the preferences of the household to smooth consumption over time, and the capital adjustment technology which allows them to do so. To build intuition, consider a simple supply and demand framework for the market for new investment, cleared by the

[42] My model has a slightly higher volatility of hours than the benchmark because labor is also used to pay the fixed cost.

real interest rate. This market is depicted in Figure 10. The supply of new investment comes from the household saving for future consumption; a higher interest rate increases the return on saving, so the household saves more. The demand for new investment comes from firms investing in capital for production next period; a higher interest rate increases the discount factor on next period's profits, so firms demand less. Hence, the supply curve reflects the preferences of the household to smooth consumption, and the demand curve reflects the ability to do so through capital adjustment.

The dynamics of the real interest rate are determined by the relative importance of these two key forces. Again to build intuition, consider the effect of a productivity shock in the simple supply and demand framework. The supply curve shifts out because the shock increases the household's wealth, and the household saves out of this wealth. But the demand curve also shifts out because the shock increases the future marginal product of capital, and firms invest to take advantage of that. Whether the interest rate increases or decreases, and by how much, depends on the relative magnitude of these two shifts.

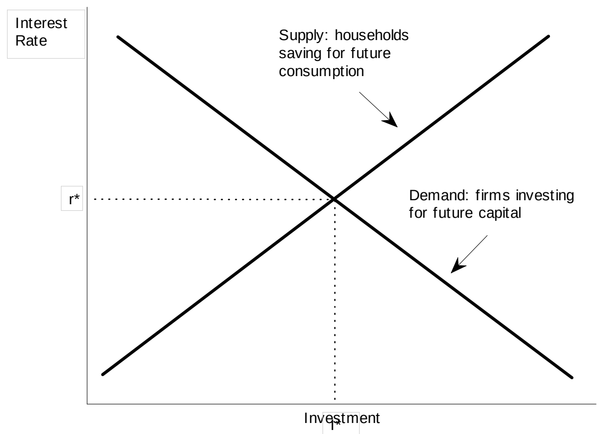
Ultimately, the main difference between my model and Khan & Thomas (2008) is in the relative weight we place on these two key forces. Khan & Thomas (2008) implicitly choose a configuration which generates procyclical real interest rate dynamics, i.e., the demand shift must be greater than the supply shift. This implies that the calibrated adjustment costs are low. Since these adjustment costs are the source of the state dependence, Khan & Thomas (2008)'s model generates little state dependence.

I choose a configuration of these two key forces which generates countercyclical interest rate dynamics, i.e., the supply shift is greater than the demand shift. On the one hand, habit formation increases the shift in supply, because households save more out of the shock. On the other hand, the convex adjustment costs decrease the shift in supply, because firms find it more costly to take advantage of the shock. The increased role of adjustment costs then leads to state dependence.

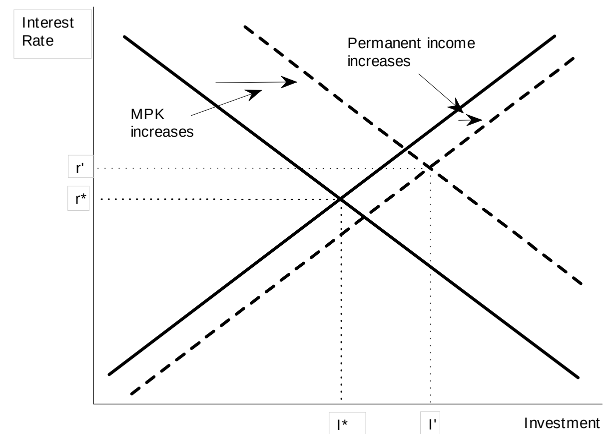
This discussion is closely related to Bachmann, Caballero, & Engel (2013)'s decomposition of what they call adjustment cost smoothing (AC-smoothing) and price response smoothing (PR-smoothing). Specifically, Bachmann, Caballero, & Engel (2013) note that the effect of a shock on aggregate investment is, in principle, attenuated by two forces: micro-level adjustment frictions, and general equilibrium price responses. They refer to this attenuation as smoothing, and la-

Figure 10: Supply and Demand Intuition for Interest Rate Dynamics

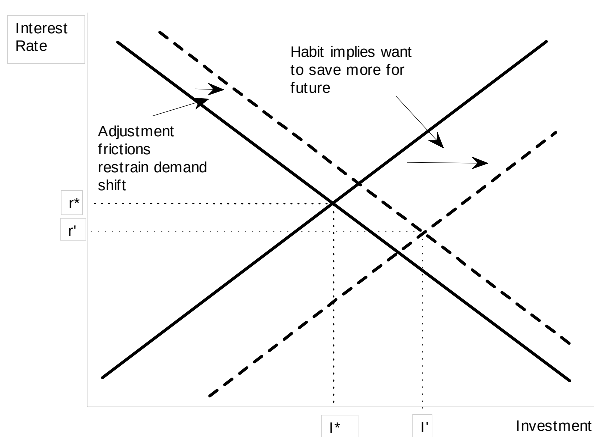
(a) Original equilibrium



(b) Positive shock without habit or adjustment costs



(c) Positive shock with habit and adjustment costs



Notes: Simple supply and demand intuition for the market for new investment, cleared by the real interest rate. (a) Original equilibrium. (b) Positive productivity shock without habit or adjustment costs. (c) Positive productivity shock with habit formation and adjustment costs.

bel the respective components AC-smoothing and PR-smoothing. In Khan & Thomas (2008)'s model, almost all the smoothing is due to PR-smoothing, because the interest rate is procyclical. However, when the interest rate is mildly countercyclical, as in the data, the contribution of PR-smoothing diminishes, and accordingly the contribution of AC-smoothing increases. As Bachmann, Caballero, & Engel (2013) discuss, it is exactly the importance of AC-smoothing which leads to state dependence.

F Heteroskedasticity

In this Appendix, I collect results about conditional heteroskedasticity which were referenced in the main text. Specifically, in Appendix F.1, I investigate the robustness of the results to an alternative specification of the statistical model. In Appendix F.2, I argue against the hypothesis that the heteroskedasticity in the aggregate investment rate is driven by heteroskedasticity in aggregate shocks.

F.1 Robustness to Alternative Specification

The statistical model in the main text is a special case of a more general class of models suggested by Bachmann, Caballero, & Engel (2013):

$$\begin{aligned} \frac{I_t}{K_t} &= \phi_0 + \sum_{s=1}^{p} \phi_j \frac{I_{t-j}}{K_{t-j}} + \sigma_t e_t,\ e_t \sim N(0,1) \\ \sigma_t^2 &= \beta_0 + \beta_1 \left(\frac{1}{p} \sum_{j=1}^{p} \frac{I_{t-j}}{K_{t-j}} \right) + u_t \end{aligned} \tag{18}$$

The specification in the main text imposes the lag length $p = 1$, which is motivated by the fact that the underlying shock process is an AR(1). However, a natural question is how robust the results are to alternative choices of lag length p. To address this issue, I focus on the case $p = 6$, which is the data's preferred lag length according to the Bayesian Information Criterion.

Table 12 shows that the qualitative results from the main text carry through for this alternative case: the data support my model's prediction of positive heteroskedasticity, and reject the benchmark model's prediction of no heteroskedasticity. However, quantitatively my model now

Table 12: Conditional Heteroskedasticity, Alternate Specification

Statistic	**Data**	**Model**	**Benchmark**
$\log\left(\frac{\widehat{\sigma}_{90}}{\widehat{\sigma}_{10}}\right)$	.263** (.041)	.112	.006
$\log\left(\frac{\widehat{\sigma}_{75}}{\widehat{\sigma}_{25}}\right)$	.134** (.025)	.057	.003

Notes: Results from estimating the model (18) with $p = 6$. Standard errors computed with bootstrapping procedure.

underpredicts the amount of heteroskedasticity in the data. This reflects the fact that my model is driven by AR(1) shocks, and therefore struggles with persistence at longer lags.

F.2 Ruling Out Heteroskedasticity in the Shocks

In this section, I argue against the hypothesis that the conditional heteroskedasticity in investment is driven by heteroskedastic shocks. First, I show that measured TFP shocks do not exhibit much heteroskedasticity. Berger & Vavra (2014) do a similar exercise for Basu, Fernald, & Kimball (2006)'s utilization-adjusted TFP and the fed funds rate and find that neither exhibit conditional heteroskedasticity. However, this analysis does not rule out the possibility that other, unmeasured shocks are heteroskedastic. To address that possibility, I then show that GDP does not contain significant heteroskedasticity either. To the extent that unmeasured shocks affect output and investment in similar ways, this provides further evidence against unmeasured heteroskedastic shocks.

Measured TFP To check for heteroskedasticity in measured TFP, I estimate the same type of model as equation (9) in the main text:

$$\begin{aligned} z_t &= \phi_0 + \phi_1 z_{t-1} + \sigma_t e_t,\ e_t \sim N(0,1) \qquad (19) \\ \sigma_t^2 &= \beta_0 + \beta_1 z_{t-1} + u_t \end{aligned}$$

Table 11 shows that there is virtually no conditional heteroskedasticity in measured TFP. This result is also robust to other choices of lag length, as in Appendix F.1.

Table 13: Conditional Heteroskedasticity in Measured TFP

Statistic	**Data**
$\log\left(\frac{\widehat{\sigma}_{90}}{\widehat{\sigma}_{10}}\right)$	$.032^{**}$ (.009)
$\log\left(\frac{\widehat{\sigma}_{75}}{\widehat{\sigma}_{25}}\right)$	$.015^{**}$ (.007)

Notes: Results from estimating the model (19). Standard errors computed using bootstrapping procedure.

Table 14: Conditional Heteroskedasticity in Real GDP

Statistic	**Data**
$\log\left(\frac{\widehat{\sigma}_{90}}{\widehat{\sigma}_{10}}\right)$	$-.124^{**}$ (.045)
$\log\left(\frac{\widehat{\sigma}_{75}}{\widehat{\sigma}_{25}}\right)$	$-.065^{**}$ (.032)

Notes: Results from estimating the model (20). Standard errors computed using bootstrapping procedure.

GDP To check for heteroskedasticity in GDP, I again estimate an equation of the form:

$$\begin{aligned} Y_t &= \phi_0 + \phi_1 Y_{t-1} + \sigma_t e_t,\, e_t \sim N(0,1) \\ \sigma_t^2 &= \beta_0 + \beta_1 Y_{t-1} + u_t \end{aligned} \tag{20}$$

where Y_t is GDP, detrended using a log-linear time trend. Table 13 shows that there is mild negative heteroskedasticity in GDP. This is consistent with Bloom (2009), who shows that GDP growth is more volatile in recessions. I interpret this as evidence that, if anything, underlying unmeasured shocks are more volatile in recessions. The fact that aggregate investment is more volatile in expansions is then evidence of the strength of the mechanism described in this paper.