Lumpy Investment, Business Cycles, and Stimulus Policy

Thomas Winberry
Chicago Booth and NBER
Thomas.Winberry@chicagobooth.edu

October 31, 2018

Abstract

I study the aggregate implications of micro-level lumpy investment in a model consistent with the empirical dynamics of the real interest rate. I find that the elasticity of aggregate investment with respect to shocks is procyclical because more firms are likely to make an extensive margin investment in expansions. Matching the dynamics of the real interest rate is key to generating this result; otherwise, counterfactual behavior of the model would eliminate most of the procyclical responsiveness in general equilibrium. Therefore, data on interest rates places important discipline on the role of general equilibrium in aggregating micro-level investment.

*I am extremely grateful to Richard Rogerson, Pablo Ottonello, Greg Kaplan, Mike Golosov, and Esteban Rossi-Hansberg for many conversations that shaped this paper. I also thank Mark Aguiar, Diane Alexander, Adrien Auclert, Gabe Chodorow-Reich, Jesus Fernandez-Villaverde, Oleg Itskhoki, Nobu Kiyotaki, Yann Koby, Ben Moll, Ezra Oberfield, Chris Sims, Christian Vom Lehn, Christian Wolf, and Eric Zwick. Financial support from Chicago Booth is gratefully acknowledged.
1 Introduction

Aggregate investment is one of the most volatile components of GDP over the business cycle, accounting for 38% of the decline in GDP during recessions.\footnote{Computed as the average contribution to the percentage change in GDP from BEA Table 1.1.2 during NBER recession dates, 1953-2018.} These large swings in aggregate investment are primarily driven by changes in the number of firms undertaking an investment project – the extensive margin – rather than changes in the size of investment projects – the intensive margin.\footnote{See, for example, the evidence in \textit{Doms and Dunne} (1998) or \textit{Gourio and Kashyap} (2007).} However, most DSGE macro models make the simplifying assumption that all changes in aggregate investment are the result of a representative firm operating along the intensive margin. Therefore, a key question for business cycle modeling is: does this abstraction matter for understanding aggregate investment? In a benchmark real business cycle (RBC) model, the answer to this question is a resounding no; general equilibrium changes in the real interest rate bring the aggregate investment series in line with the preferences of the representative household, regardless of whether that investment occurs along the extensive or intensive margin at the micro-level.\footnote{This irrelevance result was established in an important series of papers by \textit{Thomas} (2002), \textit{Khan and Thomas} (2003), \textit{Khan and Thomas} (2008), and further elaborated upon by \textit{House} (2014).}

In this paper, I argue that accounting for the importance of the extensive margin \textit{does} matter for our understanding of aggregate investment. My argument has two main components. First, I show that the dynamics of the real interest rate which drive the irrelevance results in the RBC environment are at odds with the data. Second, I build a heterogeneous firm model consistent with both the importance of the extensive margin of investment and the observed dynamics of the real interest rate, and find that the behavior of aggregate investment in my model is substantially different than in the representative firm model. In particular, the elasticity of aggregate investment with respect to shocks is procyclical; in expansions, more firms are close to making an extensive margin investment, so an additional shock generates more total investment than it would otherwise. These results illustrate how data on interest rates place important discipline on the role of general equilibrium in determining the aggregate implications of micro-level investment behavior.

In the benchmark RBC environment, changes in the real interest rate are able to elimi-
nate these time-varying impulse responses because, as House (2014) forcefully demonstrates, investment is extremely price sensitive. Therefore, small but procyclical movements in the real interest rate restrain large movements in the extensive margin which would generate procyclical responses to shocks. However, I show that these movements in the interest rate are inconsistent with two key features of the data. First, the interest rate is negatively correlated with aggregate output and productivity, suggesting that it does not dampen cyclical movements in investment demand. Second, the interest rate is an order of magnitude more volatile in the data than in the RBC model. Since the model matches the volatility of investment in the data, this result suggest that investment is not as interest-sensitive as in the model.

Motivated by this evidence, my model extends a simple heterogeneous firm model to capture the empirical dynamics of the real interest rate. There is a fixed mass of firms with both fixed and convex capital adjustment costs; the presence of fixed costs generates the extensive margin of investment. There is a representative household whose preferences feature habit formation. Aggregate dynamics are driven by shocks to aggregate TFP. The equilibrium dynamics of the real interest rate are determined by the strength of habit formation – which controls the sensitivity of investment supply to aggregate shocks – and the overall strength of the adjustment costs – which control the sensitivity of investment demand to shocks. I calibrate these feature of the model to match both the dynamics of the real interest rate and the importance of the extensive margin in micro-level investment.

Quantitatively, my calibrated model predicts that aggregate investment is nearly 40% more responsive to an aggregate shock starting from a brisk expansion than starting from a deep recession. As described above, this procyclical responsiveness reflects the fact that more firms are close to making an extensive margin investment in expansions, so an additional shock induces more firms to invest. However, this mechanism alone is not enough to generate procyclical responses in general equilibrium; as Khan and Thomas (2008) show, embedding this extensive margin mechanism into an otherwise standard RBC model generates little variation in the responsiveness to shocks. In my version of their calibration, aggregate investment is only 8% more responsive to shocks in the expansion compared to recession.

My model generates substantial variation in responsiveness for two related reasons. First,
matching the negative comovement between the real interest rate and aggregate productivity implies that the interest rate does not directly dampen the effect of a shock. Second, separately matching the dynamics of the interest rate and the behavior of investment indirectly disciplines the interest-sensitivity of investment, which Koby and Wolf (2018) show is key to determining aggregation in this class of models. In my version of Khan and Thomas (2008)’s model, the semi-elasticity of aggregate investment with respect to the real interest rate is over $-1000$, so small changes in the interest rate have a strong influence on aggregate investment. In my model, the corresponding semi-elasticity is only $-7.55$, giving general equilibrium a much smaller influence over aggregate dynamics.

I also illustrate two implications of my model for investment stimulus policy. First, the aggregate effect of investment stimulus policy is also state dependent and falls in recessions; therefore, predictions based on linear models would overstate the effectiveness of stimulus policies in recessions. Second, I develop a simple size-dependent stimulus policy that increases cost effectiveness by 77% compared to existing size-independent policies. The main insight of my alternative policy is to avoid subsidizing inframarginal investment that would have been done even without the policy; because investment primarily occurs along the extensive margin, most of this inframarginal waste is accounted for by subsidizing firms that would have made an extensive margin investment without the policy. In my model, small firms grow faster than average and are therefore more likely to be inframarginal to the policy.

A key challenge throughout the analysis is efficiently computing the equilibrium of the model, which involves approximating the entire cross-sectional distribution of firms. I use the method developed concurrently in Winberry (2018), which approximates the distribution with a flexible but finite dimensional parametric family. I find that this approach captures how changes in the shape of the distribution affect the dynamics of aggregate variables more accurately than simply using the mean of the capital distribution.

**Related Literature** This paper contributes to three main strands of literature. First, it addresses the long-standing question of how the extensive margin of investment impacts aggregate dynamics. Early papers, analyzing firms’ decision rules with fixed prices, find
that the extensive margin generates a procyclical responsiveness to shocks as in my model. However, Thomas (2002), Khan and Thomas (2003), and Khan and Thomas (2008) show that this time-varying elasticity disappears when prices are endogenized in an otherwise standard RBC framework, rendering the extensive margin irrelevant for aggregate dynamics. House (2014) suggests that these irrelevance results are driven by the extreme sensitivity of investment to the relative price of investment goods in a stylized partial equilibrium model. In recent work, Koby and Wolf (2018) extend House (2014)’s argument to a quantitative general equilibrium model in terms of the interest-sensitivity of investment. I show that matching the dynamics of the real interest rate requires breaking this extreme sensitivity and, therefore, a key source of the irrelevance results.

To match the dynamics of the real interest rate, I follow Beaudry and Guay (1996) in using habit formation and capital adjustment costs. Boldrin, Christiano and Fisher (2001) also use this approach to match interest rate dynamics and the level of the equity premium. These papers work in a representative agent environment; my results show that many of their lessons carry over to a heterogeneous firm environment in which adjustment costs are disciplined with micro-level investment data.

Finally, this paper contributes to a large literature which studies investment stimulus policy. Many papers estimate the effect of stimulus policy through linear regression models, most recently in House and Shapiro (2008) and Zwick and Mahon (2017). Edge and Rudd (2011) introduce the Bonus Depreciation Allowance into a linearized New Keynesian model, which rules out state dependence by construction. I focus on the effect of stimulus policy over the cycle and how micro-level targeting can increase its cost effectiveness.

---

4 See, for example, Caballero et al. (1995) or Caballero and Engel (1999).

5 Other papers challenge the irrelevance results on other grounds. Gourio and Kashyap (2007) argue that the results are sensitive to the distribution of fixed adjustment costs and the degree of returns to scale. Bachmann, Caballero and Engel (2013) argue that Khan and Thomas (2008)’s calibrated fixed costs are implausibly small and that increasing them to empirically reasonable levels breaks the irrelevance results. Bachmann and Ma (2016) and Bayer and Tjaden (2016) argue that the irrelevance results are sensitive to the precise form of general equilibrium; Bachmann and Ma (2016) allow for savings in durable goods and Bayer and Tjaden (2016) allow for multiple countries. Finally, and most closely related to this paper, Cooper and Willis (2014) parameterize an interest rate process from the data and solve firms’ decision problems given this process. My paper produces such an interest rate process endogenously in general equilibrium.

6 Berger and Vavra (2015) analyze a related class of consumer durable stimulus policies in a model of lumpy durable investment. They find that stimulus policies are less effective in recessions for similar reasons as here; however, they focus on detailed features of the micro data while I focus on the role of real interest rate dynamics in aggregation and on designing more cost effective policies.
Road Map  The rest of this paper is organized as follows. Section 2 describes the role of the real interest rate in driving the existing irrelevance results in the literature and argues that interest rate behavior is counterfactual. Section 3 develops my quantitative heterogeneous firm model, which Section 4 parameterizes to jointly match micro-level investment behavior and macro-level interest rate dynamics. Section 5 show that the existence of the extensive margin implies that aggregate investment is more responsive to shocks in expansions than in recessions, and argues that matching the dynamics of the interest rate is key to generate this result in general equilibrium. Section 6 introduces stimulus policy into the model, shows that the effectiveness of these policies falls in recessions, and develops an alternative size-dependent stimulus to increase cost effectiveness. Finally, Section 7 concludes.

2 Role of Real Interest Rate Dynamics

This section motivates the features of real interest rate dynamics on which I will focus for the rest of the paper. Section 2.1 uses a simplified RBC model to illustrate the role of the interest rate in rendering the fixed costs irrelevant for aggregate dynamics. Section 2.2 show that the key features of the interest rate which drives irrelevance are inconsistent with the data.

2.1 Irrelevance of Fixed Costs in Simple RBC Model

I derive an analytical irrelevance result in a simple model which highlights a general mechanism in quantitative models: since investment is extremely sensitive to changes in the real interest rate, small and procyclical movements in the interest rate bring aggregate investment in line with the representative household’s desired path of smooth consumption. This analysis builds heavily on House (2014), who derives a similar mechanism in a different model, and Koby and Wolf (2018), who provide a thorough analysis of it in a quantitative DSGE model. The value added of my analysis here is to summarize the mechanism in a transparent way and to study its empirical implications for interest rate dynamics.
**Simple RBC Model with Fixed Costs** Consider a discrete time environment with heterogeneous firms indexed by $j \in [0, 1]$. Firm $j$ produces output $y_{jt}$ using the production function

$$y_{jt} = z_t \varepsilon_{jt} k_{jt}^\alpha,$$

where $z_t$ is aggregate productivity, $\varepsilon_{jt}$ is idiosyncratic productivity, $k_{jt}$ is the firm’s capital stock, and the parameter $\alpha \leq 1$ controls the returns to scale. Idiosyncratic productivity $\varepsilon_{jt}$ follows a first-order Markov process with finite support $\varepsilon \in \{\varepsilon_1, ..., \varepsilon_n\}$. Firms have perfect foresight over the path of aggregate productivity $z_t$; since firms are owned by the representative household (see below), the absence of aggregate uncertainty implies that firms use the risk-free rate $r_t$ to discount profits.\(^7\) The capital stock $k_{jt}$ is predetermined at time $t$ and each period the firm chooses next period’s capital $k_{jt+1} = (1 - \delta)k_{jt} + i_{jt}$, where $i_{jt}$ is gross investment and $\delta$ is the depreciation rate. Nonzero gross investment incurs a fixed resource adjustment cost $\xi$, which is rebated lump-sum to the representative household. The initial distribution of idiosyncratic productivity and capital across firms is invariant in absent changes in aggregate TFP $z_t$.

There is a representative household with preferences over consumption $C_t$ represented by the utility function

$$\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\sigma} - 1}{1 - \sigma},$$

where $\beta$ is the discount factor and $1/\sigma$ is the elasticity of intertemporal substitution. The household owns all firms in the economy. Total output can be used for consumption or investment, which implies the aggregate resource constraint

$$Y_t = C_t + K_{t+1} - (1 - \delta)K_t,$$

where $Y_t = \int y_{jt}dj$ and $K_t = \int k_{jt}dj$.

**Proposition 1.** Let the returns to scale parameter $\alpha < 1$. Let $k^*_t(\alpha)$ be the optimal capital accumulation policy at date $t$ for a firm with idiosyncratic productivity $\tilde{\varepsilon} = \max_i \mathbb{E}[\varepsilon'|\varepsilon_i]$.

\(^7\)Section 4.3 argues that the risk-free rate captures nearly all of the implications of the empirical stochastic discount factor on firms’ investment decisions in a model with aggregate uncertainty.
conditional on paying the fixed cost. Let \( \pi^*_t(\alpha) \) denote the flow profits associated with that choice. If \( \xi(\alpha) \leq \pi^*_t(\alpha) \), then

\[
 r_t + \delta \to z_{t+1} \tilde{\varepsilon} \quad \text{as } \alpha \to 1. \tag{2}
\]

Furthermore, aggregate output \( Y_t \), investment \( I_t \), and consumption \( C_t \) approach the outcomes of a representative firm model with aggregate productivity \( \tilde{Z}_t = Z_t \tilde{\varepsilon} \) and initial capital stock \( K_0 \) equal to the aggregate capital stock in the invariant distribution.

**Proof.** See Appendix A. ■

In the limit of Proposition 1, fixed costs \( \xi \) are “irrelevant” in the sense that the aggregate outcomes can be derived from a representative firm without fixed costs. The key insight is that, as the returns to scale approaches one, the firm becomes infinitely elastic with respect to changes in the real interest rate (because their profit function becomes linear with respect to investment). Since general equilibrium requires that consumption be positive and finite, the real interest rate adjusts to ensure that firms with the highest value of expected future productivity earn zero profits. At this real interest rate, only firms with \( \mathbb{E}[\varepsilon' | \varepsilon] = \tilde{\varepsilon} \) accumulate capital and all other firms do not invest.

In this stark example, the irrelevance of fixed costs is ensured by the real interest rate adjusting to (2) in order to ensure an equilibrium which satisfies the household’s preference for positive consumption exists. Although quantitative models do not exactly satisfy the conditions in Proposition 1, they are often close to the limit, giving the representative household’s preferences for smooth consumption a powerful influence over aggregate investment. This occurs in quantitative models for two reasons. First, allowing for modest decreasing returns \( \alpha < 1 \) does not break the extreme sensitivity of investment with respect to the real interest rates; the semi-elasticity for adjusting firms is

\[
 \frac{\partial i_{jt}/i_{jt}}{\partial r_t} = -\frac{1}{\delta} \frac{1}{1 - \alpha} \left( \frac{1 + r_t}{r_t + \delta} \right). \tag{3}
\]

As \( \alpha \to 1 \), this semi-elasticity (3) becomes infinite. However, even under a more typical
calibration—for example, with $\alpha = 0.7$, $\delta = 0.025$, and $r_t = 0.01$—it is still over 3,847.\(^8\)

Second, the requirement of small fixed costs $\bar{\xi}$ is often guaranteed by assuming that fixed costs are a random draw from a $U[0, \bar{\xi}]$, which ensures that there is always a positive mass of firms with arbitrarily small fixed costs. Furthermore, House (2014) shows that even when fixed costs are non-random and positive, the elasticity of the timing of investment episodes with respect to the relative price of capital is infinite, playing a similar role to the infinite interest rate elasticity here.

### 2.2 Comparing Interest Rate Dynamics to the Data

The simple framework described in Section 2.1 implies that the real interest rate must move one-for-one with productivity to ensure that the zero variable profit condition (2) holds and generate aggregation. In more general quantitative models, the real interest rate will adjust to ensure that the zero profit approximately holds, generating approximate aggregation. In this subsection, I show that these real interest rate dynamics are counterfactual.

**Measurement** I study the joint dynamics of the real interest rate, aggregate TFP, and aggregate output in the U.S. data 1954q1 - 2016q4. I measure the real interest rate $r_t$ as the nominal return on 90-day Treasury bills adjusted for realized CPI inflation. I measure aggregate productivity $Z_t$ as the the aggregate Solow residual, adjusted for labor force composition. Finally, I measure output $Y_t$ as real GDP. Details of the data construction are contained in Appendix B.1. I compare the data to the benchmark RBC model, which is quantitatively close to models with fixed costs in the previous literature given the irrelevance results (e.g. Khan and Thomas (2008)); the detailed model specification and calibration, which follows standard practice in the business cycle literature, is contained in Appendix C.

**Descriptive Results** Table 1 shows that the RBC model is counterfactual along two key dimensions. First, the standard deviation of the interest rate is an order of magnitude higher in the data (1.73%) than in the model (0.16%). Second, the real interest rate is negatively correlated with aggregate TFP ($-0.20$), while in the model the two are extremely

\(^8\)Gourio and Kashyap (2007) calibrate a strong degree of decreasing returns and show that helps break this type of irrelevance result.
Table 1

**Cyclical Dynamics of Risk-Free Rate**

<table>
<thead>
<tr>
<th></th>
<th>$\sigma(r_t)$</th>
<th>$\rho(r_t, y_t)$</th>
<th>$\rho(r_t, z_t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole sample</td>
<td>1.73%</td>
<td>−0.11*</td>
<td>−0.06</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.09)</td>
<td>(0.33)</td>
<td></td>
</tr>
<tr>
<td>No Volcker</td>
<td>1.13%</td>
<td>0.07</td>
<td>−0.18***</td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Pre-1983</td>
<td>1.57%</td>
<td>−0.38***</td>
<td>−0.17*</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>Post-1983</td>
<td>1.86%</td>
<td>0.21**</td>
<td>−0.24***</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.00)</td>
<td></td>
</tr>
<tr>
<td>RBC</td>
<td>0.16%</td>
<td>0.95</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Notes: real interest rate measured as the return on 90-day Treasury bills adjusted for realized CPI inflation, expressed in annual percentage points. Output measured as real GDP. TFP measured as the aggregate Solow residual. All variables have been HP-filtered and expressed as percentage deviation from an HP trend. “Whole sample” refers to the 1954q1 - 2016q4 time series. “No Volcker” excludes 1979q1 - 1983q4. “Pre-1983” refers to the 1954q1-1982q4 sample. “Post-1983” refers to the 1983q1-2016q4 sample. “RBC” refers to the benchmark RBC model described in Appendix C.

positively correlated (0.97). Table 1 also shows that these two conclusions are robust in three different sub-samples of the data: pre-1983, post-1983, and excluding the years near the Volcker recession.

Figure 1 further illustrates the stability of these two statistics by plotting eight-year rolling windows over the postwar sample. Although the standard deviation of the real interest rate varies over time, it is consistently above the prediction of the RBC model. The correlation of the real interest rate with aggregate productivity is negative for nearly the entire sample. The correlation of the real interest rate with GDP is negative before 1983 but positive after 1983, consistent with the fact that productivity is less procyclical in the later period. Nevertheless, both correlations are consistently below the level implied by the RBC model.

**Impulse Response to TFP Shock** In order to make a consistent comparison between the data and the RBC model, which is solely driven by TFP shocks, I estimate the impulse response of the real interest rate to a TFP shock using a simple bivariate VAR:

$$X_t = \sum_{j=1}^{p} \Gamma_j X_{t-j} + e_t,$$

(4)
where \( X_t = (Z_t, r_t)^T \), \( p \) is the lag length, \( \Gamma_j \) are coefficient matrices, and \( e_t \) are residuals. I choose the lag length \( p = 3 \) following the AIC. I identify TFP shocks by assuming that shocks to the interest rate equation do not affect TFP upon impact.\(^9\) In order to that the responses are not driven by endogenous changes in utilization, I use the adjusted TFP measure from Fernald (2014).

Figure 2 shows that the RBC model fails along the same two dimensions highlighted in the descriptive analysis above.\(^{10}\) First, the empirical response is negative while the model’s is positive, consistent with the differences in correlations in Table 1. Second, the magnitude of the empirical response is larger than in the model, consistent with the differences in volatilities in Table 1.

---

\(^9\)Note that this identifying assumption is weaker than the assumption that TFP is exogenous with respect to the real interest rate.

\(^{10}\)Figure 2 also shows that the model’s theoretical impulse response is nearly identical the response estimated using the VAR (4) on model-simulated data.
Notes: impulse response of the real interest rate to a TFP shock identified from a bivariate VAR with TFP ordered first. TFP is adjusted for changes in utilization following Fernald (2014). Lag length of 3 chosen by the AIC criteria. “RBC theoretical” refers to the theoretical impulse response from the benchmark RBC model described in Appendix C. “RBC measured” refers to the impulse response identified using the VAR estimation on simulated data from the model. “Empirical (90% CI)” refers to the empirical impulse response and 90% error bands.

Robustness Appendix B.2 show that these results continue to hold if one uses a VAR to construct expected inflation, and are robust to different choices of business cycle filters. It also displays the impulse response of the ex-ante real interest rate to a TFP shock, which is targeted in the model calibration in Section 4.

3 Model

Motivated by the results in Section 2, I now develop a quantitative heterogeneous firm model to be consistent with both the behavior of investment at the micro-level and the dynamics of the real interest interest rate at the macro level.

3.1 Environment

The model is set in discrete time.
Firms

The firm side of the model builds heavily on Khan and Thomas (2008), extended to include convex adjustment costs and the corporate tax code.\(^\text{11}\) There is a fixed mass of firms \(j \in [0, 1]\) that produce output \(y_{jt}\) using the production function

\[ y_{jt} = z_t \varepsilon_{jt} k_{jt}^\theta n_{jt}^\nu, \]

where \(z_t\) is an aggregate productivity shock, \(\varepsilon_{jt}\) is an idiosyncratic productivity shock, \(k_{jt}\) is the firm’s capital stock, \(n_{jt}\) is its labor input, and \(\theta\) and \(\nu\) are parameters satisfying \(\theta + \nu < 1\). The aggregate shock \(z_t\) is common to all firms and follows the AR(1) process

\[ \log z_{t+1} = \rho \log z_t + \omega^z_{t+1}, \text{ where } \omega^z_{t+1} \sim N(0, \sigma^2_z). \]

The idiosyncratic shock \(\varepsilon_{jt}\) is independent across firms but within firm follows the AR(1) process

\[ \log \varepsilon_{jt+1} = \rho \log \varepsilon_{jt} + \omega^\varepsilon_{t+1}, \text{ where } \omega^\varepsilon_{t+1} \sim N(0, \sigma^2_\varepsilon). \]

Each period, a firm \(j\) observes these two shocks, uses its pre-existing capital stock, hires labor from a competitive labor market at wage \(w_t\), and produces output \(y_{jt}\).

After production, the firm decides how much capital in which to invest for the next period. Gross investment \(i_{jt}\) 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 \(i_{jt} \neq 0\), then the firm must pay a fixed cost \(\xi_{jt}\) in units of labor.\(^\text{12}\) The fixed cost \(\xi_{jt}\) is a uniform random variable with support \([0, \xi]\), distributed independently across firms and time. Second, any nonzero amount of investment incurs the quadratic adjustment cost \(-\frac{\phi}{2} (\frac{i_{jt}}{k_{jt}})^2 k_{jt}\) units of output.

After production and investment, the firm pays a linear tax rate \(\tau\) on its revenue \(y_{jt}\) net of two deductions. First, the firm deducts its labor costs \(w_t n_{jt}\). Second, it deducts capital depreciation costs according to the following geometric schedule. The firm enters the period with a stock of depreciation allowances \(d_{jt}\), of which it writes off \(\hat{\delta} d_{jt}\) from its tax bill. The firm also writes off the same fraction \(\hat{\delta}\) of new investment \(i_{jt}\) from its tax bill. The remaining

\(^{11}\)I include the tax code in order to study investment stimulus policy in Section 6.
\(^{12}\)Khan and Thomas (2008) allow for any investment in \([-ak_{jt}, ak_{jt}]\) to be free of the fixed costs. I set \(a = 0\) for simplicity, but results are robust to allowing for empirically reasonable values of \(a\).
portion is then carried into the next period, so that $d_{jt+1} = (1 - \hat{\delta})(d_{jt} + i_{jt})$.\(^{13}\) In total, the tax bill in period $t$ is

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

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

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

where $C_t$ is consumption, $N_t$ is labor supplied to the market, and $X_t$ is habit stock. The habit stock $X_t$ is simply

$$
X_t = \lambda \hat{C}_{t-1},
$$

where $\hat{C}_t = C_t - \chi \frac{N_t^{1+\eta}}{1+\eta}$ is consumption net of the disutility of work and $\lambda$ controls the sensitivity of habit with respect to $\hat{C}_t$.\(^{14}\) I assume that the household does not take into account the fact that their choices impact the habit stock $X_t$. The total time endowment per period is 1, so that $N_t \in [0, 1]$. The household owns all firms in the economy and markets are complete.

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

$$
T_t = \tau \left( Y_t - w_t n_t - \hat{\delta}(D_t + I_t) \right),
$$

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

\(^{13}\)In reality, the U.S. tax code follows an annual straight line depreciation schedule with a half-year purchase convention rather than this simple geometric schedule. However, Section 3.2 shows that only the present value of this schedule per unit of investment affects firm’s decisions, so this modeling assumption is without loss of generality. See Xu and Zwick (2018) for an analysis of the implications of the details of the tax code using a model of firm-level tax management.

\(^{14}\)I assume this particular form of preferences for two reasons. First, following Greenwood, Hercowitz and Huffman (1988), they eliminate the wealth effect on labor supply and allow the model to generate procyclical hours worked with a countercyclical real interest rate. With standard KPR preferences, the fact that the real interest rate falls in expansions potentially induces households to intertemporally substitute future leisure for current leisure in expansions, leading to a fall in hours worked. Second, assuming habit formation over the consumption bundle $\hat{C}_t$ simplifies the analysis of the stochastic discount factor below, but the results also hold if habit is defined over actual consumption $C_t$ only.
3.2 Firm Optimization

I characterize the firm’s optimization problem recursively. The firm’s individual state variables are $\varepsilon_{jt}$, 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 $\xi_{jt}$, its current draw of the fixed cost. I denote the aggregate state vector $s_t$ and postpone discussion of its elements until I define the recursive competitive equilibrium in Section 3.4.

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

$$
v(\varepsilon, k, d, \xi; s) = \tau \hat{\delta} d + \max_n \left\{ (1 - \tau) \left( e^{\varepsilon} e^{\delta} k^\theta n^\nu - w(s)n \right) \right\}
+ \max \left\{ v^a(\varepsilon, k, d; s) - \xi w(s), v^n(\varepsilon, k, d; s) \right\}. \tag{7}
$$

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

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

$$
v^a(\varepsilon, k, d; s) = \max_i \left( 1 - \tau \hat{\delta} \right) i - \frac{\varphi}{2} \left( \frac{i}{k} \right)^2 k + E[\Lambda(z'; s)v(\varepsilon', k', d', \xi'; s')|\varepsilon, k, d] \tag{8}
\text{s.t. } k' = (1 - \delta)k + i \text{ and } d' = \left( 1 - \hat{\delta} \right) (d + i),
$$

where $\Lambda(z'; s)$ is the stochastic discount factor. I denote the implied “target” capital stock $k^a(\varepsilon, k, d; s) = (1 - \delta) k + i^a(\varepsilon, k, d; s)$.

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

$$
v^n(\varepsilon, k, d; s) = E[\Lambda(z'; s)v(\varepsilon', k', d', \xi'; s')|\varepsilon, k, d] \tag{9}
\text{s.t. } k' = (1 - \delta)k \text{ and } d' = \left( 1 - \hat{\delta} \right) d.
$$

The only difference from the unconstrained Bellman equation (8) is that investment is constrained to be $i = 0$. I call the implied capital stock $k^a(\varepsilon, k, d; s) = (1 - \delta) k$ the constrained
capital stock.

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 \( v^a(\varepsilon, k, d; s) - \xi w(s) \geq v^n(\varepsilon, k, d; s) \). For each tuple \((\varepsilon, k, d; s)\), there is a unique threshold \( \hat{\xi}(\varepsilon, k, d; s) \) which makes the firm indifferent between these two options. The threshold satisfies

\[
\hat{\xi}(\varepsilon, k, d; s) = \frac{v^a(\varepsilon, k, d; s) - v^n(\varepsilon, k, d; s)}{w(s)}.
\]

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

It is possible to simplify this problem by eliminating the tax depreciation allowances \( d \) from the firm’s state vector. They key insight is that firms only care about the present value of the depreciation allowances generated by their investment because the allowances enter the firm’s flow profits separately from the other terms. Therefore, the tax depreciation schedule only affects firm’s decisions by changing the effective price of investment to include the present value of tax writeoffs. I formalize this logic in Proposition 2:

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

\[
v^1(\varepsilon, k, \xi; s) = \pi(\varepsilon, k; s) + \max_i \left\{ -q(s)i - \frac{(\xi - \hat{\delta})^2}{2i^2} k - \xi w(s) \mathbb{1}\{i \neq 0\} + \mathbb{E}[\Lambda(z'; s)v^1(\varepsilon', (1 - \delta) k + i, \xi'; s')] \right\},
\]

where \( q(s) = (1 - \tau PV(s)) \) is the tax-adjusted relative price of investment.

**Proof.** See Appendix D.

In Appendix D, I show that this result can be leveraged to simplify the model’s equilibrium following the strategy developed by Khan and Thomas (2003).
3.3 Household Optimization

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

$$\max_{C,N} \log \left( C - \chi \frac{N^{1+\eta}}{1+\eta} - X(s) \right) \text{ subject to } C \leq w(s)N + \Pi(s) + T(s), \quad (12)$$

where $\Pi(s)$ are profits from the firms and $T(s)$ is government transfers. Markets are complete with respect to aggregate risk, so the stochastic discount factor used by firms is equal to the household’s intertemporal marginal rate of substitution state by state:

$$\Lambda(z';s) = \beta \frac{\hat{C}(s) - X(s)}{\hat{C}(s') - X(s')}.$$  \quad (13)

3.4 Definition of Equilibrium

The aggregate state vector is $s = (z, X, \mu)$, where $z$ is the aggregate productivity shock, $X_{-1}$ is the household’s habit stock, and $\mu$ is the distribution of firms over their individual state vector $(\varepsilon, k, \xi, d)$.

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

(i) (Household Optimization) Taking $w(s)$, $\Pi(s)$, and $T(s)$ as given, $C(s)$ and $N(s)$ solve the utility maximization problem (12).

(ii) (Firm Optimization) Taking $w(s)$, $\Lambda(z'; s)$, $X'(s)$, and $\mu'(s)$ as given, $v(\varepsilon, k, d, \xi; s)$, $n(\varepsilon, k; s)$, $i^a(\varepsilon, k; s)$, and $\hat{\xi}(\varepsilon, k; s)$ solve the firm’s maximization problem (7) - (10).

(iii) (Government) For all $s$, $T(s)$ is given by (6).

(iv) (Consistency) For all $s$,

(a) $\Pi(s) = \int [ (1 - \tau) (e^{\varepsilon} e^{\xi} k^d n(\varepsilon, k; s) - w(s)n(\varepsilon, k; s)) + \tau \delta d - (1 - \tau \hat{\delta}) i(\varepsilon, k, \xi; s) - \frac{\varphi}{2} \left(\frac{i(\varepsilon, k, \xi; s)}{k}\right)^2 k - \xi w(s)1\{\frac{i(\varepsilon, k, \xi; s)}{k} \neq 0\} ] \mu(d\varepsilon, dk, dd, d\xi)$, where $i(\varepsilon, k, d, \xi; s)$
\( i^a(\varepsilon, k, d; \xi; s) \) if \( \xi \leq \bar{\xi}(\varepsilon, k; s) \) and \( i(\varepsilon, k, \xi; s) = 0 \) otherwise.

(b) \( \Lambda(z'; s) \) is given by (13).

(c) \( X'(s) \) follows (5).

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

(v) (Market Clearing) For all \( s, \) \( N(s) = \int n(\varepsilon, k, \xi; s) \mu(d\varepsilon, dk, dd, d\xi). \)

3.5 Solution Method

The key challenge to solving the model is that the aggregate state vector \( s \) contains the cross-sectional distribution of firms, which is an infinite-dimensional object. I overcome this challenge using the computational method concurrently developed in Winberry (2018). The method approximates the distribution at any point in time using a flexible but finite-dimensional parametric family; the parameters of that family are then endogenous aggregate state variables. I have found that a good approximation of the distribution requires 5-10 endogenous parameters, leaving globally accurate approximation methods infeasible due to the curse of dimensionality. Therefore, I solve for the aggregate dynamics of the model using a second-order perturbation. See Appendix E for details of the implementation.

I have found that Winberry (2018)’s method has two related over the usual approach of approximating the distribution with moments, as in Krusell and Smith (1998). First, forecasts of key variables based only on the aggregate capital stock are inaccurate (which I show in Appendix E). This fact indicates that higher-order features of the distribution are relevant in determining aggregate dynamics. Second, the method is computationally efficient due to the use of perturbation methods with respect to the aggregate state vector. This efficiency allows me to calibrate the model in order to explicitly match moments in the data.\(^{15}\)

\(^{15}\)The key advantage of Krusell and Smith (1998)’s method is that their solution is globally accurate with respect to aggregate shocks. In this model, a local approximation with respect to aggregate shocks is sufficient because those shocks are small.
Table 2
Fixed Parameter Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>Discount factor</td>
<td>.99</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Inverse Frisch</td>
<td>1</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Labor share</td>
<td>.64</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Capital share</td>
<td>.21</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Capital depreciation</td>
<td>.025</td>
</tr>
<tr>
<td>$\rho_z$</td>
<td>Aggregate TFP AR(1)</td>
<td>.95</td>
</tr>
<tr>
<td>$\sigma_z$</td>
<td>Aggregate TFP AR(1)</td>
<td>.007</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Tax rate</td>
<td>.35</td>
</tr>
<tr>
<td>$\hat{\delta}$</td>
<td>Tax depreciation</td>
<td>.119</td>
</tr>
</tbody>
</table>

Notes: parameters fixed in calibration.

4 Model Parameterization

I parameterize the model to jointly match the dynamics of the real interest rate and the behavior of investment at the micro level.

4.1 Paramaterization

My parameterization proceeds in two steps. First, I fix a set of parameters to match standard macroeconomic targets in steady state. Second, given the values of those parameters, I choose the remaining parameters to match targets in the data.

Fixed Parameters Table 2 lists the parameters that I fix. A model period is one quarter, so I set the discount factor $\beta = 0.99$. I set the Frisch elasticity of labor supply to 2, within the range of macro elasticities identified by Chetty et al. (2011). I set the labor share $\theta = 0.64$ and choose the capital share so that the total returns to scale is 85%. The returns to scale lies within the range considered in the literature, from 60% in Gourio and Kashyap (2007) to 92% in Khan and Thomas (2008). I set $\delta = 0.025$ so that the steady state aggregate investment rate is 10%, roughly in line with the average in the postwar data. I set the stochastic process for TFP to $\rho_z = 0.95$ and $\sigma_z = 0.007$ as in King and Rebelo (1999).

I set the tax rate $\tau = 0.35$ to match the top marginal income tax rate in the U.S. federal
income tax code over most of the sample period. Given the tax rate, I choose the slope of the tax depreciation schedule \( \delta \) to match the average present value of tax depreciation allowances per unit of investment from Zwick and Mahon (2017). Proposition 2 shows that the present value summarizes how the tax depreciation schedule affects firms’ decisions.

**Fitted Parameters** I choose the remaining parameters, listed in Table 4, in order to match the targets in Table 3.\(^{16}\) The micro-level investment targets are computed from annual IRS corporate income tax returns, reported in Zwick and Mahon (2017).\(^ {17}\) The IRS sample features significant micro-level lumpiness, in line with previous findings in Census data (see, for example, Cooper and Haltiwanger (2006)). About one fourth of firm-years in the sample feature essentially zero investment while simultaneously one sixth of firm-years have investment rate spikes greater than 20%.

I target two statistics related to the dynamics of the real interest rate. First, I target the one-year cumulative response of the expected real interest rate to a TFP shock identified using the VAR (4).\(^ {18}\) Second, I target the volatility of aggregate investment relative to the volatility of aggregate output. As I discuss in Section 4.2 below, increasing the strength of either habit formation or adjustment costs has similar implications for the dynamics of the real interest rate. However, they have opposite implications for the volatility of investment; adjustment costs make investment less volatile while habit formation makes it more volatile.

Targeting the volatility of investment therefore places some discipline on the overall strength

\(^{16}\)I exogenously fix the persistence of the productivity shocks \( \rho_e = 0.9 \) because Clementi and Palazzo (2015) show that the persistence is weakly identified separately from the volatility of shocks \( \sigma_e \) using investment data alone.

\(^{17}\)Much of the literature with firm heterogeneity and investment calibrates models to match investment behavior from Census of manufacturing firms, reported by Doms and Dunne (1998) or Cooper and Haltiwanger (2006). Zwick and Mahon (2017)'s data has three important advantages over Census data in the context of this paper. First, it covers all sectors of the economy rather than just manufacturing, and therefore allows for a more representative sample of the economy than previous studies. Second, it covers a more recent sample period (1998-2010) than the Census data (1972-1988). Third, the IRS data is at the firm level, which is the appropriate unit of analysis for studying tax policy in Section 6. However, Zwick and Mahon (2017)'s data also has two disadvantages relative to Cooper and Haltiwanger (2006). First, the tax data only record investment expenditures, while the Census data also records retirement and sales of capital. Second, measured investment in the Zwick and Mahon (2017) sample mainly includes equipment goods while measured capital includes both equipment and structures. On net, I prefer to work with the Zwick and Mahon (2017) data because of its greater sample coverage across sectors and time. I conjecture that my results are robust to calibrating the model to the Census data since both datasets indicate substantial lumpiness of investment.

\(^{18}\)The impulse response of the ex-ante real interest rate is plotted in Figure 8 in Appendix B.2.
Table 3
Empirical Targets

<table>
<thead>
<tr>
<th>Micro Investment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average investment rate (%)</td>
<td>10.4%</td>
<td>10.7%</td>
</tr>
<tr>
<td>Standard deviation of investment rates</td>
<td>0.16</td>
<td>0.15</td>
</tr>
<tr>
<td>Spike rate (%)</td>
<td>14.4%</td>
<td>19.0%</td>
</tr>
<tr>
<td>Positive investment rates (%)</td>
<td>85.6%</td>
<td>81.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interest Rate Dynamics</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative impulse response</td>
<td>−0.49</td>
<td>−0.31</td>
</tr>
<tr>
<td>$\sigma(I_t)/\sigma(Y_t)$</td>
<td>2.87</td>
<td>2.88</td>
</tr>
</tbody>
</table>

Notes: micro investment moments from annual firm-level IRS data, 1998 - 2010, as reported in Zwick and Mahon (2017) Appendix Table B.1. Statistics drawn from distribution of investment rates pooled over firms and time. Spike rate is fraction of observations with investment rate greater than 20%. Positive investment is fraction of observations less than 20%. “Cumulative impulse response” refers to the cumulated response of the ex-ante real interest rate to a TFP shock over the first year, identified from the empirical VAR (4). $\sigma(C_t)/\sigma(Y_t)$ is the standard deviation of HP-filtered aggregate consumption relative to the standard deviation of HP-filtered aggregate output.

Although the model is nonlinear and overidentified, with six moments pinning down four parameters, it nonetheless fits the targets in Table 3 fairly well. The model captures the frequency of spikes relative to the frequency of non-spike observations, which is informative about the strength of fixed costs. The model also captures the dispersion of investment rates across firms, which is informative about the size of idiosyncratic shocks and strength of the convex adjustment costs. While the model matches the negative response of the interest rate to a TFP shock, it only captures around two thirds of the overall decline. This failure

---

19 Habit formation has also been shown to improve the empirical performance of DSGE models along two other important dimensions as well. First, habit improves the internal propagation of shocks onto aggregate consumption; for example, Christiano, Eichenbaum and Evans (2005) show that habit formation allows their model to match the hump-shaped response of consumption to a monetary policy shock. Second, habit helps match various features of asset prices (see, for example, Boldrin, Christiano and Fisher (2001)).

20 A natural target which I omit is the occurrence of investment inaction, often defined as the frequency of observations with investment rates less than 1% annually. I do not target inaction for two reasons. First, as discussed by Cooper and Haltiwanger (2006), the precise definition of inaction in the data is rather arbitrary given heterogeneity in investment goods or in the types of investment episodes (e.g., maintenance vs. large new projects). Second, since the IRS data only reports investment expenditure, observed inaction may erroneously reflect firms which do not purchase capital but which nonetheless sell or retire capital. It is straightforward to allow for some degree of inaction in the model by allowing for a nonzero amount of investment to not be subject to the fixed costs.
Table 4  
Fitted Parameter Values

<table>
<thead>
<tr>
<th>Micro Heterogeneity</th>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\xi$</td>
<td>Upper bound on fixed costs</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>$\varphi$</td>
<td>Convex adjustment cost</td>
<td>2.34</td>
</tr>
<tr>
<td></td>
<td>$\rho_\varepsilon$</td>
<td>Idiosyncratic productivity AR(1) (fixed)</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>$\sigma_\varepsilon$</td>
<td>Idiosyncratic productivity AR(1)</td>
<td>0.056</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Habit Formation</th>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda$</td>
<td>Sensitivity of habit w.r.t. consumption bundle</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Notes: Parameters chosen to match moments in Table 3. I exogenously fix the persistence of idiosyncratic productivity following the discussion in footnote 16.

is primarily due to the fact that the hump-shaped nature of empirical response, which the model cannot generate with the simple habit formation process in (5).

Table 4 shows that the calibrated parameter values are broadly comparable to previous findings in the literature. The upper bound on the fixed cost $\xi$ is within the admittedly wide range of 0.0083 in Khan and Thomas (2008) and 4.4 in Bachmann, Caballero and Engel (2013). The calibrated value implies that the average fixed cost paid conditional on adjusting is 8.1% of firms’ average output. The dispersion of idiosyncratic TFP shocks is in line with direct measures surveyed in Syverson (2011). The average size of the habit stock is 73% of the households consumption bundle $\hat{C}_t$, similar to the 65% in Christiano, Eichenbaum and Evans (2005).

4.2 Identification

The identification of these parameters can be understood in two broad steps. First, the dynamics of the real interest rate pin down the overall strength of habit formation and adjustment costs. This occurs because the model’s equilibrium real interest rate is determined by the interaction of habit formation, which controls the sensitivity of investment supply to productivity shocks, and adjustment costs, which control the sensitivity of investment demand to productivity shocks. Second, given the overall strength of adjustment costs, the micro-level investment data pins down the dispersion of idiosyncratic shocks and the split
between fixed and convex adjustment costs.

In order to more formally understand how interest rate dynamics pin down the strength of habit formation and adjustment costs, Figure 3 plots the impulse response of key model variables under four different parameter configurations: no habit formation or adjustment costs; only habit formation; only adjustment costs; and the calibrated model with both habit formation and adjustment costs.

First consider the cases without habit formation. The household’s Euler equation is then

\[ 1 + r_t = \frac{1}{\beta} \mathbb{E}_t \left[ \frac{\hat{C}_t}{\hat{C}_{t+1}} \right]^{-1}, \]

which relates the real interest rate to expected consumption growth.\(^\text{21}\) While a positive TFP shock unambiguously increases both consumption \(C_t\) and investment \(I_t\), its affect on consumption growth \(C_{t+1}/C_t\) – and therefore the real interest rate – depends on the re-

---

\(^\text{21}\)Due to Greenwood, Hercowitz and Huffman (1988) preferences, the Euler equation is technically in terms of the consumption bundle \(\hat{C}_t = C_t - \lambda^{1+\eta} \). Quantitatively the dynamics of this consumption bundle are similar to the dynamics of consumption \(C_t\) itself.
sponsiveness of investment (which determines the capital stock in period \( t + 1 \)). Without adjustment costs, investment responds enough that consumption in the next period \( C_{t+1} \) rises relative to \( C_t \), therefore causing the real interest rate to rise. Adjustment costs dampen the response of investment, which dampens the rise of consumption growth and therefore of the real interest rate. However, adjustment costs alone are not quantitatively strong enough to fully account for the response of the real interest rate in the data.

Habit formation further brings the model in line with the data by breaking the tight link between consumption growth and the real interest rate. With habit formation, the household’s Euler equation is

\[
1 + r_t = \frac{1}{\beta} \mathbb{E}_t \left[ \frac{\hat{C}_t - X_t}{\hat{C}_{t+1} - X_{t+1}} \right]^{-1},
\]

where \( X_t \) is the stock of habit at time \( t \). This equation implies that the real interest rate will fall following a positive TFP shock if the growth in the habit-adjusted consumption bundle \((\hat{C}_t - X_t)\) falls. And indeed, holding the path of the consumption bundle \( \hat{C}_t \) fixed, stronger habit formation decreases the response of habit-adjusted consumption growth; since \( X_t \) is predetermined during the period of the shock while \( X_{t+1} \) is not, \( \hat{C}_t - X_t \) increases by more than \( \hat{C}_{t+1} - X_{t+1} \). However, stronger habit formation also increases the incentive to smooth the consumption bundle \( \hat{C}_t \); in fact, Figure 3 shows that consumption smoothing undoes nearly all the effect of habit formation on the real interest rate. It is only when habit formation is combined with adjustment costs, which impede the household’s ability to smooth consumption over time, that the real interest rate substantially falls in response to a TFP shock. Hence, both habit formation and adjustment costs are necessary to match the dynamics of the real interest rate.

### 4.3 Relationship Between the Real Interest Rate and the SDF

The dynamics of the risk-free rate that I target in my calibration are tightly related to the stochastic discount factor \( \Lambda(z';s) \) which firms use to value the benefits of investment. Note
that the expected present value of the firm between any two periods can be decomposed as

$$
\mathbb{E}[\Lambda(z'; s)v(\varepsilon', k', \xi'; s')|\varepsilon, k, s] = \frac{1}{1 + r(s)} \mathbb{E}[v(\varepsilon', k', \xi'; s')|\varepsilon, k, s]
+ \text{Cov}(\Lambda(z'; s), v(\varepsilon', k', \xi'; s')|\varepsilon, k, s)
$$

(14)

where $r(s) = \frac{1}{\mathbb{E}[\Lambda(z'; s)|s]} - 1$ is the risk-free rate. The first term in the decomposition (14) captures the risk-free discounting of the value function, i.e. the implications of the stochastic discount factor (SDF) for intertemporal comparisons. The second term captures the covariance between the SDF and the value function, i.e. the implications of the SDF for risk. Hence, my calibration strategy directly targets the intertemporal component of the SDF and places no direct discipline on the risk component.\(^{22}\)

Capturing the dynamics of the risk component is outside the scope of my paper for two related reasons. First, in order to capture movements in the expected risk premium – which are informative about the risk component – the macro asset-pricing literature often appeals to stochastic volatility or a time-varying market price of risk, both of which are outside the class of models driven by trend-stationary TFP shocks. Second, results from Golosov and Winberry (2018) suggest that movements in the covariance term are quantitatively unimportant in this class of models.\(^{23}\) They show that the empirical covariance between the excess return on equity and aggregate TFP, which is tightly linked to the covariance between the SDF and aggregate TFP, is small. Furthermore, changes in the excess return and aggregate TFP are essentially uncorrelated, which suggests that the time-series comovement between the SDF and aggregate TFP is also small.

### 4.4 Model Validation

Before presenting the main results of the paper, I show that the model performs well along dimensions that were not targeted in the calibration.

\(^{22}\)Note that my empirical analysis ignores changes in the inflation risk premium, which are likely to be small on a quarter-to-quarter basis.

\(^{23}\)Golosov and Winberry (2018) estimate a process for the SDF using time-series data on the risk-free rate and equity premium, feed that empirical process into the investment decision problem of an aggregate firm subject to aggregate TFP shocks, and find that the variation in the covariance term accounts for less than 1% of the total variation in investment.
Figure 4: Distribution of Annualized Investment Rates in Steady State

Notes: histogram of investment rates in the model’s steady state. Investment rates are time-aggregated to the annual level in order to compare to the data.

Table 5

<table>
<thead>
<tr>
<th>Unconditional Business Cycle Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Volatility</strong></td>
</tr>
<tr>
<td>$\sigma(Y_t)$</td>
</tr>
<tr>
<td>$\sigma(C_t)/\sigma(Y_t)$</td>
</tr>
<tr>
<td>$\sigma(I_t)/\sigma(Y_t)$</td>
</tr>
<tr>
<td>$\sigma(N_t)/\sigma(Y_t)$</td>
</tr>
</tbody>
</table>

Notes: All series have been logged and HP-filtered with smoothing parameter 1600. “Data” refers to the data described in Appendix B.1. “Model” refers to calibrated model.

Micro Investment Behavior  The stationary distribution of realized investment rates across firms is broadly comparable to the empirical distribution reported in Zwick and Mahon (2017). Figure 4 plots the histogram of investment rates in the model’s stationary distribution. Due to the fixed cost, there is a large mass of observations with zero investment. At the same time, there is a large mass of observations with large positive investment spikes. Overall, the distribution is highly non-normal; it features both positive skewness (1.97, compared to 3.60 in the data) and excess kurtosis (10.16, compared to 17.6 in the data).
Unconditional Business Cycle Statistics  Table 5 shows that the model matches standard business cycle statistics roughly as well as the benchmark RBC model, despite the fact that the model has much stronger habit formation and adjustment costs. Investment is more volatile than output and consumption less volatile than output in both the model and the data. The volatility of hours is lower in the model than in the data, which is well-known defect of the benchmark RBC model as well. Finally, all macroeconomic aggregates are highly correlated with each other due to the fact that there is a single aggregate shock.

5  Procyclical Responsiveness to Shocks

I now use my calibrated model to analyze the quantitative implications of fixed costs for the dynamics of aggregate investment. Section 5.1 shows that fixed costs imply that aggregate investment is more responsive to shocks in expansions than in recessions because more firms are close to making an extensive margin investment in expansions. Section 5.2 shows that quantitatively matching the dynamics of the real interest rate is crucial to generate this procyclical responsiveness in general equilibrium, complementing the qualitative discussion in Section 2.

5.1  Role of Fixed Costs

I begin by describing how fixed adjustment costs generate the the procyclical responsiveness of aggregate investment to an TFP shock. In order to isolate the role of firms’ behavior in driving aggregate dynamics, I perform this analysis in “partial equilibrium,” i.e. by aggregating firms’ decision rules with prices held fixed at their steady state values. This analysis provides a natural benchmark against which to compare the general equilibrium results in Section 5.2.

Procyclical Impulse Responses  Figure 5 illustrates the procyclical responsiveness by comparing the response of aggregate investment to a TFP shock starting from two different points in the business cycle. The first is an “expansion” generated by a history of one standard deviation positive shocks over the previous year and the second is a “recession” generated
Figures 5: Procyclical Impulse Responses of Aggregate Investment

Notes: left panel plots the impulse responses to a one standard deviation positive shock to aggregate TFP computing following Koop, Pesaran and Potter (1996). “Expansion” refers to a history of four one standard deviation positive shocks and “recession” refers to history of four one standard deviation negative shocks. “Partial equilibrium” refers to aggregating firms’ decisions holding prices fixed at their steady state values and “general equilibrium” refers to the full general equilibrium model. Since the model is nonlinear, I compute the impulse responses by (i) drawing a random series of aggregate shocks, (ii) adding the history of shocks to generate an expansion and recession, (iii) computing the difference between the simulations in which there is the additional shock and the original simulation, and (iv) repeating this procedure 200 times and taking the average of all the differences produced in step (iii).

Right panel plots how the adjustment probability for firms (conditional on a high realization of idiosyncratic productivity ε) responds to a positive aggregate shock starting from steady state. The blue line (measured against the right panel) plots the steady state distribution of firms over capital k. the red lines (measured against the left panel) plot the probability paying the fixed cost and adjusting capital. The solid red line is in steady state and the dashed line is following a one standard deviation positive TFP shock (with prices held fixed at their steady state values).

by a symmetric history of negative shocks. Since the model is nonlinear, I replicate this procedure over many simulations and then take the average of these responses.

The left panel of Figure 5 shows that the response of aggregate investment to the shock is significantly larger starting from the expansion than the recession. In the expansion, the shock generates 35% more investment upon impact and 8% more in total. The fact that the cumulative difference is smaller than the impact difference reflects intertemporal substitution; starting from the expansion, some firms pull forward investment they would have done in the future into the period of the shock.
Role of Fixed Costs  The state dependence in these impulse responses is due to more firms making an extensive margin investment starting from the expansion. In order to understand this result, first note that in steady state the average firm holds less capital than its target stock $k^a(\varepsilon, k; s)$ because of capital depreciation and convex adjustment costs $\varphi$ (see footnote 25 below). Now consider a history of negative shocks which generates a recession. Since the negative shocks decrease the marginal product of capital, they decrease the target capital stock and therefore bring the average firm closer to its target. In this case, the probability of a firm paying its fixed cost – which is proportional to the adjustment threshold $\hat{\xi}(\varepsilon, k)$ – falls. Furthermore, additional shocks will have a relatively small effect on the adjustment probability $\hat{\xi}(\varepsilon, k)$ as well. On the other hand, a history of positive shocks will move the average firm even further below its target, i.e. $k^a(\varepsilon, k; s) << k^a(\varepsilon, k; s)$, and increase the adjustment probability $\hat{\xi}(\varepsilon, k; s)$. In this region of the state space, it turns out that further changes in $k^a(\varepsilon, k; s)$ have larger effects on the adjustment probability $\hat{\xi}(\varepsilon, k; s)$. Hence, the fact that the adjustment probability $\hat{\xi}(\varepsilon, k; s)$ is increasing in the distance from target $|k^a(\varepsilon, k; s) - k^a(\varepsilon, k; s)|$ is the key source of procyclical responses to shocks.

The right panel of Figure 5 plots how the adjustment probabilities of firms responds to a positive productivity shock starting from steady state (conditional on a high realization of idiosyncratic productivity $\varepsilon$). The adjustment probability is a convex function of capital $k$ and increasing in $|k^a(\varepsilon, k; s) - k^a(\varepsilon, k; s)|$. The positive shock increases the target capital stock for all firms and therefore shifts the adjustment probability function up and to the right.

5.2 Role of Prices in General Equilibrium

The left panel of Figure 5 shows that the procyclical responses described above survive in general equilibrium. The degree of state dependence is lower in general equilibrium for two reasons. First, the real wage $w_t$ is procyclical, which decreases the marginal revenue product.

---

24 Since firms’ draws of the fixed cost $\xi$ are i.i.d., for each value of productivity and capital $(\varepsilon, k)$ a fraction $\hat{\xi}(\varepsilon, k; s)$ of firms will adjust while the remaining fraction will not.

25 The adjustment probability function is positive throughout the distribution of firms due to the existence of convex adjustment costs $\varphi$. Convex costs imply that the target capital stock is a decreasing function of current capital since capital decreases the marginal adjustment cost.
of capital in response to the shock. Second, as Figure 6 shows, the real interest rate $r_t$ falls by more in the recession than in the expansion, reflecting the fact that marginal utility growth falls by more starting from the recession.\footnote{Table 7 shows that this nonlinearity in the interest rate process is partly driven by habit formation.} Despite these dampening forces, the degree of procyclical responsiveness is still quantitatively large in general equilibrium.

While this stylized example has been instructive, in order to quantify the amount of time-variation in the impulse response function over a long simulation of the model, I follow Bachmann, Caballero and Engel (2013) and compute the “responsiveness index” $RI_t$

$$RI_t = 100 \times \log \left( \frac{I(z_t + \sigma_z, X_t, \mu_t) - I(z_t, X_t, \mu_t)}{I(\sigma_z, X^\ast, \mu^\ast) - I(0, X^\ast, \mu^\ast)} \right),$$

where $I(z, X, \mu)$ is aggregate investment given the aggregate state $s = (z, X, \mu).$\footnote{Bachmann, Caballero and Engel (2013) use a more general measure that accounts for asymmetries in the response to a positive and negative shock. These asymmetries are small in my model, so I ignore them for the sake of simplicity.} The responsiveness index measures the impact effect of a TFP shock at a given point in time, relative to the effect starting from steady steady state.
Table 6
Fluctuations In Responsiveness Index Over Time

<table>
<thead>
<tr>
<th></th>
<th>95-5 ratio</th>
<th>90-10 ratio</th>
<th>75-25 ratio</th>
<th>( \rho(\text{RI}_t, \log Y_t) )</th>
<th>( \rho(\text{RI}_t, \text{adj}_t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benchmark Calibration</strong> (PE interest elasticity ( d \log I_t / dr_t = -7.55 ))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial Equilibrium</td>
<td>64%</td>
<td>50%</td>
<td>25%</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>General Equilibrium</td>
<td>31%</td>
<td>23%</td>
<td>15%</td>
<td>0.99</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Khan and Thomas (2008) Calibration</strong> (PE interest elasticity ( d \log I_t / dr_t = -1055.41 ))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial Equilibrium</td>
<td>49%</td>
<td>38%</td>
<td>18%</td>
<td>0.92</td>
<td>0.94</td>
</tr>
<tr>
<td>General Equilibrium</td>
<td>7%</td>
<td>5%</td>
<td>3%</td>
<td>0.98</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Notes: responsiveness index \( \text{RI}_t \) defined in (15) in the main text. \( \text{adj}_t \) computes the fraction of firms who pay their fixed cost. “Partial equilibrium” refers to aggregating firms’ decision rules, holding prices fixed at their steady state values. “General equilibrium” refers to the full general equilibrium model. “Benchmark calibration” refers to the calibrated model. “Khan and Thomas (2008) Calibration” refers to eliminating the convex adjustment costs (\( \varphi = 0 \)), reducing the fixed costs (\( \xi = 0.0083 / 4 \)), changing the idiosyncratic shock process (\( \rho_\varepsilon = 0.859^{1/4} \) and \( \sigma_\varepsilon = 0.022 / 4 \)), increasing the returns to scale (\( \theta = 0.256 \)), eliminating the tax code (\( \tau = 0 \)), eliminating habit formation (\( \lambda = 0 \)), and using separable preferences between consumption and labor supply (\( \log C_t - \chi \frac{N^{1+\gamma}}{1+\gamma} \)). “PE interest elasticity” is the effect of a one-time unexpected change in the real interest rate in steady state.

Table 6 shows that the model generates a significant amount of procyclical responsiveness over a long simulation of the model; the 95th percentile of the responsiveness index \( \text{RI}_t \) is 31% higher than the 5th percentile. Furthermore, the responsiveness index is positively correlated with aggregate output and the fraction of firms paying their fixed cost, consistent with the mechanism described in Section 5.1.

**Relationship to Proposition 1** These quantitative results suggest that the model is far from the limiting case of Proposition 1, in which fixed costs are irrelevant for aggregate dynamics. In that case, movements in the real interest rate are able to bring the dynamics of investment in line with the desires of the representative household because the elasticity of aggregate investment with respect to the interest rate is nearly infinite. In my calibrated model, that elasticity is only \(-7.55\), so general equilibrium has less influence of aggregate investment dynamics.\(^{28}\)

In contrast to my model, Khan and Thomas (2008) provide an important quantitative model in which fixed costs are irrelevant for aggregate dynamics. Table 6 roughly replicates

\(^{28}\)I compute the interest elasticity as the effect of a one-time unexpected increase in the real interest rate starting at steady state. I aggregate firm-level elasticities according to the stationary distribution.
their calibration at the quarterly frequency and compares it to my model. The partial equilibrium version of their model generates a sizable amount of variation in the responsiveness index; the 95th percentile is nearly 50% higher than the 5th percentile. However, general equilibrium eliminates nearly 85% of the variation in the responsiveness index. Consistent with the logic of Proposition 1, general equilibrium is so powerful in Khan and Thomas (2008)’s model because investment is extremely interest sensitive; the elasticity of aggregate investment with respect to the real interest rate is over −1000. Furthermore, the remaining degree of variation in the responsiveness index $RI_t$ is also present in the version of the model without fixed costs, which aggregates to a representative firm. Hence, in the Khan and Thomas (2008) model fixed costs are essentially irrelevant for aggregate dynamics relative to the representative firm model.

Building on insights from House (2014) and an earlier draft of this paper, Koby and Wolf (2018) provide a thorough analysis of the role of price-sensitivity in driving the irrelevance results of the previous literature and argue that quantitative models should target the interest-sensitivity of investment. My calibration implicitly targets the interest-sensitivity of investment by separately targeting the volatility of investment and the dynamics of the interest rate in the data. The presence of convex adjustment costs $\varphi$ is key to matching the interest-sensitivity because it leads to an upward-sloping marginal cost curve for investment.

**Role of Model Ingredients** Table 7 decomposes the role of three key model ingredients in driving the procyclical responsiveness to shocks in my model. First, decreasing the size of the fixed costs decreases the variation in the responsiveness index, which is natural given that the fixed costs are the source of the state dependence (as described in Section 5.1).

---

29 I roughly replicate Khan and Thomas (2008) calibration by eliminating the convex adjustment costs ($\varphi = 0$), reducing the fixed costs ($\bar{F} = 0.0083/4$), changing the idiosyncratic shock process ($\rho_\varepsilon = 0.859^{1/4}$ and $\sigma_\varepsilon = 0.022/4$), increasing the returns to scale ($\theta = 0.256$), eliminating the tax code ($\tau = 0$), eliminating habit formation ($\lambda = 0$), and using separable preferences between consumption and labor supply ($\log C_t - \frac{\chi N^{1+\eta}}{1+\eta}$).

30 In fact, Khan and Thomas (2008) show that the entire distribution of aggregate investment rates is essentially the same between the two models.

31 Koby and Wolf (2018)’s analysis applies to a broader class of models in which the relative price of investment is not tied as directly to the real interest rate as it is in my model. They discipline the price-sensitivity of investment by targeting the response of investment to the Bonus Depreciation Allowance estimated in Zwick and Mahon (2017). While my strategy is less direct than Koby and Wolf (2018), it does not require the additional structure to accommodate Zwick and Mahon (2017)’s difference-in-differences empirical specification.
Table 7
Role of Key Model Ingredients

<table>
<thead>
<tr>
<th></th>
<th>$\sigma(RI_t)$</th>
<th>$RI_t$ 95-5 ratio</th>
<th>90-10 ratio</th>
<th>PE $d\log I_t/dr_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full model</td>
<td>8.49</td>
<td>31%</td>
<td>24%</td>
<td>−7.55</td>
</tr>
<tr>
<td>Smaller fixed costs $(\xi/2)$</td>
<td>4.76</td>
<td>17%</td>
<td>14%</td>
<td>−9.03</td>
</tr>
<tr>
<td>Smaller convex costs $(\varphi/2)$</td>
<td>4.56</td>
<td>16%</td>
<td>13%</td>
<td>−10.14</td>
</tr>
<tr>
<td>Smaller habit $(\lambda/2)$</td>
<td>8.83</td>
<td>33%</td>
<td>25%</td>
<td>−7.55</td>
</tr>
</tbody>
</table>

Notes: Responsiveness index $RI_t$ defined in (15) in the main text. “Full model” refers to the calibrated model. “Smaller fixed costs” keeps all parameters the same as the full model except decreases the upper bound on the distribution of fixed costs $\xi$ by 50%. “Smaller quadratic costs” keep all parameters the same as the full model except decreases the convex adjustment cost $\varphi$ by 50%. “Smaller habit” keeps all parameters the same as in the full model except decreases the habit formation parameter $\lambda$ by 50%. “PE $d\log I_t/dr_t$” is the effect of a one-time unexpected change in the real interest rate in steady state.

Second, decreasing the size of the convex adjustment costs also decreases the variation in the responsiveness index because, as described above, it increases the interest-sensitivity of investment. Third, and perhaps surprisingly, decreasing the strength of habit formation slightly increases the variation in the responsiveness index $RI_t$. This result occurs because, as discussed in Figure 6, habit formation creates nonlinearities in the dynamics of the real interest rate; therefore, decreasing the strength habit decreases these nonlinearities and therefore increases the nonlinearities in aggregate investment.

It is important to emphasize that habit formation nevertheless plays a crucial role in generating the quantitative variation in responsiveness index in my model because it allows the model to match interest rate dynamics for empirically reasonable values of adjustment costs. As discussed in Section 4.2, matching interest rate dynamics with adjustment costs alone would require significantly larger adjustment costs and therefore imply an even lower interest-sensitivity of investment. The comparative static in Table 7 indicates that this particular function form for habit formation generates slight nonlinearities in the interest rate. While it is presumably possible to reverse engineer a process for habit formation that would eliminate these nonlinearities (see, for example, Campbell and Cochrane (1999)), I do not pursue that approach here given that the ultimate impact on aggregate investment dynamics is small.

Appendix F contains three additional robustness checks on the results in this section.
First, it shows that the results also hold with separable preferences over consumption and labor supply. Second, it shows that the results hold when the tax code is eliminated. Third, it shows that the results are robust to changes the returns to scale.

6 Implications for Stimulus Policy

In this section, I briefly study two implications of my model for investment stimulus policies. First, as with productivity shocks, the effectiveness of investment stimulus is state dependent and falls in recessions. Second, the importance of the extensive margin at the micro level implies that a micro-targeted policy can increase cost effectiveness up to 77% compared to existing size-independent policies.

I model investment stimulus as an exogenous shock to the tax-adjusted price of capital, $q(s)$, derived in Proposition 2. In particular, for this section only, I assume that the relative price is $q(s) = 1 - \tau(PV(s) + \omega)$, where $\omega$ is the investment stimulus shock. Appendix G shows that the two most common investment stimulus policies in the U.S., the investment tax credit and the bonus depreciation allowance, map into different values for the shock $\omega$. I assume that the shock $\omega$ follows an AR(1) process:

$$\omega' = \rho_\omega \omega + \varepsilon_\omega',$$

where $\varepsilon_\omega \sim N(0, \sigma^2_\omega)$. For illustrative purposes, I choose the standard deviation of the shock $\sigma_\omega = 0.035$ to roughly match the size of a 50% Bonus Depreciation Allowance and the quarterly autocorrelation $\rho_\omega = 0.91$ to match a half-life of two years.

Figure 7 plots the impulse response of aggregate investment and consumption to a one standard deviation positive stimulus shock starting from steady state. The shock immediately decreases the relative price of investment $q(s)$, which then increases investment by 2%. Since output is fixed upon impact, this higher investment must be met with lower consumption. Note that the stimulus shock is isomorphic to an investment-specific technological shock, which also induces a negative comovement between consumption and investment. Over time,

---

32State dependence in the effect of policy does not necessary follow from the results in Section 5 because policy shocks have different general equilibrium implications than productivity shocks.
Figure 7: Average Impulse Response to Investment Stimulus Shock

Notes: impulse response of aggregate investment and aggregate consumption to a one-standard deviation positive investment stimulus shock $\omega$.

Table 8
Responsiveness Index for Investment Stimulus Shock

<table>
<thead>
<tr>
<th>95-5 ratio</th>
<th>90-10 ratio</th>
<th>75-25 ratio</th>
<th>$\rho(RI_t, \log Y_t)$</th>
<th>$\rho(RI_t, \text{adj}_t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact</td>
<td>22%</td>
<td>15%</td>
<td>6%</td>
<td>0.86</td>
</tr>
<tr>
<td>Cumulative</td>
<td>17%</td>
<td>11%</td>
<td>5%</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Notes: time-variation in the responsiveness index for the stimulus policy shock. $\text{adj}_t$ computes the fraction of firms who pay their fixed cost. “Partial equilibrium” refers to aggregating firms’ decision rules, holding prices fixed at their steady state values. “General equilibrium” refers to the full general equilibrium model.

Persistently higher investment increases the capital stock, output, and therefore consumption.

State Dependent Effect of Investment Stimulus In order to quantify the degree of state dependence in response to the stimulus shock, I follow Section 5 and construct two “policy responsiveness indices.” The first measures the impact effect of the shock:

$$RI_t^{\omega,i} = 100 \times \log \left( \frac{I(z_t, \omega_t + \sigma, X_t, \mu_t) - I(z_t, \omega_t, X_t, \mu_t)}{I(0, \sigma, X^*, \mu^*)} - I(0, \sigma, X^*, \mu^*) \right)$$,
### Table 9

**Effectiveness of Size-Dependent Policy**

<table>
<thead>
<tr>
<th>Weight on firm size $\alpha$</th>
<th>$\alpha = 0$</th>
<th>$\alpha = 2$</th>
<th>$\alpha = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total investment generated per unit of cost</td>
<td>0.58</td>
<td>0.65</td>
<td>1.03</td>
</tr>
</tbody>
</table>

Notes: aggregate investment divided by aggregate cost of a size-dependent investment stimulus shock $\omega \times n(\varepsilon, k; s)^\alpha$.

where $I(z_t, \omega_t, X_t, \mu_t)$ is aggregate investment given the expanded aggregate state vector $s_t = (z_t, \omega_t, X_t, \mu_t)$. The second index measures the cumulative effect of the shock:

$$RI_t^{\omega,c} = 100 \times \log \left( \frac{\widehat{I}(z_t, \omega_t + \sigma_\omega, X_t, \mu_t) - \widehat{I}(z_t, \omega_t, X_t, \mu_t)}{\widehat{I}(0, \sigma_\omega, X^*, \mu^*) - \widehat{I}(0, 0, X^*, \mu^*)} \right),$$

where $\widehat{I}(z_t, \omega_t, X_t, \mu_t)$ measures the cumulative amount of investment starting from aggregate state $s_t$ and reverting back to study state. The cumulative effect is closely related to the total change in the capital stock and, therefore, output and consumption.

Table 8 shows that both the impact and cumulative effect vary significantly over time; the 95th percentile of the impact effect is 22% higher than the 5th percentile, and the 95th percentile of the cumulative effect is 17% higher than the 5th percentile. Both indices are positively correlated with output, implying that the effectiveness of policy falls in recessions.

A linear forecasting model – such as a VAR, user cost, or tax-adjusted $q$ model – would abstract from this state dependence and therefore be biased up in recessions.

**Increasing Cost Effectiveness with Micro-Targeting**

A general issue with investment stimulus policies is that much of their cost is due to subsidizing investment that would have been done even without the policy. To increase cost effectiveness, one would like to avoid paying for this inframarginal investment and instead only subsidize investment that is done in response to the policy. Therefore, an important challenge to policymakers is identifying and disregarding this inframarginal investment. A key insight of my model is that, because investment occurs primarily along the extensive margin, most of the subsidy to inframarginal investment is accounted for by subsidizing firms that would have made an extensive margin investment even without the policy. This insight simplifies the problem to identifying these
inframarginal firms.

Table 9 shows that a simple size-dependent implementation of this idea is quantitatively powerful. I assume that the change in a given firm’s relative price of investment \( q(s) \) is now

\[
\omega \times n(\varepsilon, k; s)^\alpha,
\]

where \( n(\varepsilon, k; s) \) is the firm’s employment and \( \alpha \) captures the weight of the policy on large firms. Table 9 shows that increasing the weight on large firms \( \alpha \) increases the amount of investment generated by the policy, per unit of its cost, by up to 77%. This occurs because small firms grow faster than the average firm due to mean reversion in idiosyncratic productivity \( \varepsilon \). In order to grow, these firms are more likely to invest, making them more likely to be inframarginal to the policy.\(^{33}\)

7 Conclusion

In this paper, I have argued that accounting for the importance of the extensive margin in micro-level investment decisions matters for our understanding of aggregate investment dynamics because it implies that aggregate investment is more responsive to shocks in expansions than in recessions. Matching the dynamics of the real interest rate is key to generating this result; in an otherwise standard RBC model, counterfactual movements in the real interest rate eliminate most of this procyclical responsiveness. More generally, these results show that data on interest rates place sharp discipline on the role of general equilibrium in determining the aggregate implications of firm-level investment behavior.

\(^{33}\)Of course, the quantitative effect of this size-dependent policy relies on this particular model of firm growth. Clementi and Palazzo (2016) have used a similar model to study firms’ lifecycle and argue that it provides a good fit to the data. However, other models may have different implications for the correlation between size and responsiveness to investment stimulus. For example, a model with financial frictions may imply that small firms are more likely to be financially constrained and therefore more responsive to the policy. The goal of my exercise here is simply to illustrate the magnitude of the cost savings associated with micro-targeting firms along the extensive margin, rather than strongly advocate for this particular size-dependent implementation.
References


Appendix (For Online Publication Only)

A Proof of Proposition 1

Consider the optimization problem of a firm $j$ choosing capital accumulation $k_{jt+1}$ in period $t$. Conditional on paying the fixed cost $\xi$, the choice $k_{jt+1}$ affects the discounted value of the firm’s profits through the terms

$$-k_{jt+1} + \frac{1}{1+r_t} \left( z_{t+1} \mathbb{E}[\varepsilon_{jt+1}|\varepsilon_{jt}] k_{jt+1}^\alpha + (1-\delta)k_{jt+1} \right).$$

First consider the limiting case $\alpha = 1$ and $\overline{\xi} = 0$; I will discuss convergence to this limit below. In this limiting case, the expression becomes

$$\left[ \frac{1}{1+r_t} (z_{t+1} \mathbb{E}[\varepsilon_{jt+1}|\varepsilon_{jt}] + (1-\delta)) - 1 \right] k_{jt+1}. \tag{16}$$

Since the expression (16) is linear in capital accumulation $k_{jt+1}$, the optimal policy of the firm is to set $k_{jt+1} = 0$ if $\frac{1}{1+r_t} (z_{t+1} \mathbb{E}[\varepsilon_{jt+1}|\varepsilon_{jt}] + (1-\delta)) - 1 < 0$, set $k_{jt+1} \rightarrow \infty$ if $\frac{1}{1+r_t} (z_{t+1} \mathbb{E}[\varepsilon_{jt+1}|\varepsilon_{jt}] + (1-\delta)) - 1 > 0$, and can be any $k_{jt+1} \in [0, \infty)$ otherwise.

General equilibrium requires that the firm with the highest expected future productivity earns zero variable profits, i.e.

$$1 + r_t = (1-\delta) + z_{t+1} \bar{\varepsilon}. \tag{17}$$

If $1 + r_t < (1-\delta) + z_{t+1} \bar{\varepsilon}$, then the firms with $\mathbb{E}[\varepsilon_{jt+1}|\varepsilon_{jt}] = \bar{\varepsilon}$ would strictly prefer to let $k_{jt+1} \rightarrow \infty$, violating the finite resource constraint.\footnote{Note that there is a positive mass of such firms because $\varepsilon_{jt}$ has finite support.} If $1 + r_t > (1-\delta) + z_{t+1} \bar{\varepsilon}$, then no firms would find it profitable to invest, which would imply $C_{t+1} = 0$ and violate the consumer’s Inada condition.

Condition (17) implies that only firms for which $\mathbb{E}[\varepsilon_{jt+1}|\varepsilon_{jt}] = \bar{\varepsilon}$ will accumulate capital for the next period; all other firms have strictly lower expected productivity and therefore set $k_{jt+1} = 0$. The choice $k_{jt+1} = \frac{K_{t+1}}{\mu}$ is optimal for the active firms, where $K_{t+1}$ is the...
aggregate capital accumulation implied by the household’s Euler equation

\[ C_t^{-\sigma} = \beta (z_t+1 \bar{\varepsilon} + 1 - \delta) C_{t+1}^{-\sigma}. \]

Aggregate output in period \( t+1 \) is therefore \( Y_{t+1} = z_{t+1} \bar{\varepsilon} K_{t+1} \). Hence, aggregate outcomes are identical to a representative firm with production function \( Y_{t+1} = z_{t+1} \bar{\varepsilon} K_{t+1} \). Note that average productivity among active firms in period \( t+1 \) is \( \bar{\varepsilon} \) by the law of large numbers.

Now consider the general firm’s problem with \( \alpha < 1 \) and \( \bar{\xi} > 0 \). Following the statement of the proposition, let \( k^*_j(\alpha) \) be the optimal policy of a firm with productivity \( \bar{\varepsilon} \) conditional on paying the fixed cost. Further denote the mass of these firms by \( \mu_t \), which may be time-varying depending on how many pay the fixed costs. Finally, let

\[
\pi^*_t(\alpha) = -k^*_t(\alpha) + \frac{1}{1 + r_t} (z_{t+1} \bar{\varepsilon} k^*_t(\alpha)^\alpha + (1 - \delta) k^*_t(\alpha))
\]

be the contribution of the capital choice to the value of the firm’s discounted profits, net of the fixed cost \( \bar{\xi}(\alpha) \).

Now consider the limit as \( \alpha \to 1 \). The optimal policy \( k^*_j(\alpha) \) will converge to the optimal policy with \( \alpha = 1 \) if the fixed cost does not outweigh flow profits, i.e., \( \bar{\xi} \leq \pi^*_t(\alpha) \). Since \( \pi^*_t(\alpha) \to 0 \) as \( \alpha \to 1 \), this requires \( \bar{\xi} \to 0 \). Hence, by the same logic as above, in the limit it must be that \( r_t + \delta \to z_{t+1} \bar{\varepsilon} \) to ensure that the finite resource constraint of the economy is respected. Since active firms will be indifferent, the choice \( \frac{K_{t+1}}{\mu_t} \) will be optimal, and aggregate output will be given by \( Y_t = \mu_t \times z_{t+1} \bar{\varepsilon} \frac{K_{t+1}}{\mu_t} = z_{t+1} \bar{\varepsilon} K_{t+1} \).

\section*{B Data}

\subsection*{B.1 Data Sources and Variable Definitions}

I construct the variables used in the empirical analysis as follows.

- Real interest rate \( r_t \): \( 400 \left( 1 + \frac{r_{t+1}^{\text{nom}}}{1 + \pi_{t+1}} - 1 \right) \), where \( r_{t+1}^{\text{nom}} \) is the average yield on 90-day Treasury bills (FRED series DTB3) and \( \pi_{t+1} \) is realized CPI inflation (FRED series CPI-AUSCL).
Real GDP $Y_t$: nominal GDP, quarterly (NIPA Table 1.1.5) divided by GDP deflator (NIPA Table 1.1.9).

Real consumption $C_t$: nominal expenditures on consumption goods (NIPA Table 1.1.5) divided by implicit price deflator (NIPA Table 1.1.9) plus nominal expenditures on services (NIPA Table 1.1.5) divided by implicit price deflator (NIPA Table 1.1.9).

Real investment $I_t$: nominal expenditures on nonresidential fixed investment (NIPA Table 1.1.5) divided by implicit price deflator (NIPA Table 1.1.9).

Hours worked $N_t$: hours of all persons in nonfarm business sector (FRED series HOANBS).

Total factor productivity $z_t$: downloaded from FRBSF database

https://www.frbsf.org/economic-research/indicators-data/total-factor-productivity-tpf/

**B.2 Robustness of Empirical Results**

I perform three robustness checks on the empirical results in Table 1. First, I show in Table 10 that the results hold when inflation expectations are computed from a VAR (rather than realized inflation as in the main text). Second, I show in Figure 8 that the impulse response of this ex-ante real interest rate to a TFP shock is similar to the response of the ex-post real interest rate in the main text. Third, I show in Table 11 that the results hold when I detrend the data using a linear trend, a bandpass filter, or first differences.

**C Benchmark Real Business Cycle Model**

There is a representative firm with production function $Y_t = Z_t K_t^\alpha N_t^{1-\alpha}$, where $Z_t$ is aggregate productivity, $K_t$ is the aggregate capital stock, and $N_t$ is labor supply. Aggregate productivity $Z_t$ follows the log-AR(1) process $\log Z_t = \rho \log Z_{t-1} + \omega_t$, where $\omega_t \sim N(0, \sigma_z^2)$.

There is a representative household has separable preferences over consumption $C_t$ and labor supply $N_t$ represented by the expected utility function $\mathbb{E} \sum_{t=0}^{\infty} \beta^t \left( \frac{C_t^{1-\sigma}}{1-\sigma} - \chi \frac{N_t^{1+\eta}}{1+\eta} \right)$, where $\chi$ controls the disutility of labor supply and the $1/\eta$ is the Frisch elasticity.

---

35 The VAR contains four lags of inflation, output growth, consumption growth, investment growth, and unemployment.
Table 10
Cyclical Behavior of Risk-Free Rate, VAR Inflation Expectations

<table>
<thead>
<tr>
<th></th>
<th>( \sigma(r_t) )</th>
<th>( \rho(r_t, y_t) )</th>
<th>( \rho(r_t, z_t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole sample</td>
<td>1.27%</td>
<td>-0.02</td>
<td>-0.22***</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.77)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>No Volcker</td>
<td>1.19%</td>
<td>0.08</td>
<td>-0.19***</td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>Pre-1983</td>
<td>1.11%</td>
<td>-0.16*</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.39)</td>
<td></td>
</tr>
<tr>
<td>Post-1983</td>
<td>1.39%</td>
<td>0.13</td>
<td>-0.36***</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.00)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: real interest rate measured as the return on 90-day Treasury bills adjusted for inflation expectations from a VAR, expressed in annual percentage points. The VAR contains four lags of inflation, output growth, consumption growth, investment growth, and unemployment. Output measured as real GDP. TFP measured as the aggregate Solow residual. All variables have been HP-filtered and expressed as percentage deviation from an HP trend. “Whole sample” refers to the 1954q1 - 2016q4 time series. “No Volcker” excludes 1979q1 - 1983q4. “Pre-1983” refers to the 1954q1-1982q4 sample. “Post-1983” refers to the 1983q1-2016q4 sample.

Table 11
Cyclical Behavior of Risk-Free Rate, Different Filters

<table>
<thead>
<tr>
<th></th>
<th>( \sigma(r_t) )</th>
<th>( \rho(r_t, y_t) )</th>
<th>( \rho(r_t, z_t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP filter</td>
<td>1.73%</td>
<td>-0.11*</td>
<td>-0.20***</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.09)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Linear trend</td>
<td>2.58%</td>
<td>-0.13**</td>
<td>-0.35***</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.00)</td>
<td></td>
</tr>
<tr>
<td>Bandpass</td>
<td>1.22%</td>
<td>-0.17***</td>
<td>-0.33*</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.00)</td>
<td></td>
</tr>
<tr>
<td>First differences</td>
<td>2.58%</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.41)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: real interest rate measured as the return on 90-day Treasury bills adjusted for realized CPI inflation, expressed in annual percentage points. Output measured as real GDP. TFP measured as the aggregate Solow residual. All variables have been HP-filtered and expressed as percentage deviation from an HP trend. “HP filter” refers to detrending all variables using an HP filter. “Linear trend” refers to removing a linear trend from output and TFP. “Bandpass” refers to removing a bandpass filter from all variables with minimum periodicity of 6 quarters and maximum periodicity of 32 quarters. “First differences” refers to expressing output and TFP in log-differences. All statistics are computed over the 1954q1-2016q4 sample.
A model period is one quarter, so I set the discount factor $\beta = 0.99$. I set the elasticity of intertemporal substitution $1/\sigma = 1$ and the Frisch elasticity of labor supply $1/\eta = 2$. I choose the disutility of labor supply $\chi$ to ensure that steady state hours worked is $1/3$ of available time. I set the labor share $1 - \alpha = 0.64$ and the depreciation rate of capital $\delta = 0.025$. I set the process for aggregate TFP to the standard values $\rho = 0.95$ and $\sigma_z = 0.007$.

I solve the RBC model using a second-order perturbation implemented in Dynare. As I describe in Appendix E, I also solve for the aggregate dynamics of the heterogeneous firm model using a second-order perturbation in Dynare.

D Characterizing Equilibrium

In this Appendix I characterize the recursive competitive equilibrium defined in Section 3.4. I use this characterization to numerically compute the equilibrium in Appendix E. For the
sake of generality, I allow firms that do not pay the fixed cost to choose any investment $i \in [-ak, ak]$. The main text sets $a = 0$.

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

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

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

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

I begin by proving Proposition 2 in the main text. This proposition shows that the firm’s value function $v(\varepsilon, k, d, \xi; s)$ is linear in the pre-existing stock of depreciation allowances $d$. I exploit this property in the other propositions to simplify the decision rules. For ease of reading, I restate the proposition below:

**Proposition 3.** The firm’s value function is of the form $v(\varepsilon, k, d, \xi; s) = v_1(\varepsilon, k, \xi; s) + \tau PV(s)d$ where $PV(s)$ is defined by the recursion $PV(s) = \hat{\delta} + \left(1 - \hat{\delta}\right) \mathbb{E}[\Lambda(z'; s)PV(s')]$. Furthermore, $v_1(\varepsilon, k, \xi; s)$ is defined by the Bellman equation

$$
v_1(\varepsilon, k, \xi; s) = \pi(\varepsilon, k; s) + \max_i \left\{ -\left(1 - \tau PV(s)\right)i - \hat{\delta} \left(\frac{\theta}{\theta^2 + \delta}\right)^2 k - \xi w(s)1 \{ i \notin [-ak, ak]\} + \mathbb{E}[\Lambda(z'; s)v_1(\varepsilon', (1 - \delta)k + i, \xi'; s')] \right\}
$$

**Proof.** First, I show that the value function is of the form $v(\varepsilon, k, d, \xi; s) = v_1(\varepsilon, k, \xi; s) + \tau PV(s)d$ for some function $v_1(\varepsilon, k, \xi; s)$. I begin by showing that the operator $T$ defined by the right hand side of the Bellman equation maps functions of the form $f(\varepsilon, k, \xi; s) + \tau PV(s)d$ into functions of the form $g(\varepsilon, k, \xi; s) + \tau PV(s)d$. Applying $T$ to $f$, we get:

$$
T(f)(\varepsilon, k, \xi; s) = \pi(\varepsilon, k; s) + \tau \hat{\delta}d
$$

$$
+ \max_i \left\{ -\left(1 - \tau \hat{\delta}\right)i - \hat{\delta} \left(\frac{\theta}{\theta^2 + \delta}\right)^2 k - \xi w(s)1 \{ i \notin [-ak, ak]\} + \mathbb{E}[\Lambda(z'; s)(f(\varepsilon', (1 - \delta)k + i, \xi'; s') + \tau PV(s)(1 - \hat{\delta})(d + i))] \right\}
$$
Collecting terms,

\[
T(f)(\varepsilon, k, \xi; s) = \pi(\varepsilon, k; s) + \tau \left( \hat{\delta} + (1 - \hat{\delta}) \mathbb{E}[\Lambda(z'; s)PV(s')] \right) d + \max_i \left\{ \begin{array}{l}
- \left( 1 - \tau \hat{\delta} - \tau(1 - \hat{\delta}) \mathbb{E}[\Lambda(z'; s)PV(s')] \right) i - \frac{\varepsilon}{2} \left( \frac{i}{k} \right)^2 k \\

- \xi w(s) \{ i \notin [-ak, ak] \} + \mathbb{E}[\Lambda(z'; s)f(z', (1 - \delta)k + i, \xi'; s')] \end{array} \right\}
\]

By the definition of \( PV(s) \), we have that

\[
\tau \left( \hat{\delta} + (1 - \hat{\delta}) \mathbb{E}[\Lambda(z'; s)PV(s')] \right) d = \tau PV(s)
\]

\[
- \left( 1 - \tau \hat{\delta} - \tau(1 - \hat{\delta}) \mathbb{E}[\Lambda(z'; s)PV(s')] \right) i = - (1 - \tau PV(s)) i
\]

Plugging this back into (19) and rearranging gives

\[
T(f)(\varepsilon, k, \xi; s) = \tau PV(s)d + \pi(\varepsilon, k; s) + \max_i \left\{ \begin{array}{l}
- (1 - \tau PV(s)) i - \frac{\varepsilon}{2} \left( \frac{i}{k} \right)^2 k - \xi w(s) \{ i \notin [-ak, ak] \} \\

+ \mathbb{E}[\Lambda(z'; s)(v^1(z', (1 - \delta)k + i, \xi'; s') + \tau PV(s)(1 - \hat{\delta})(d + i))]
\end{array} \right\}
\]

which is of the form \( \tau PV(s)d + g(\varepsilon, k, \xi; s) \). Hence, \( T \) maps functions of the form \( \tau PV(s)d + f(\varepsilon, k, \xi; s) \) into functions of the form \( \tau PV(s)d + g(\varepsilon, k, \xi; s) \). 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 \( v(\varepsilon, k, d, \xi; s) = v^1(\varepsilon, k, \xi; s) + \tau PV(s)d \).

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

\[
v^1(\varepsilon, k, \xi; s) + \tau PV(s)d = \pi(\varepsilon, k; s) + \tau \hat{\delta}d + \max_i \left\{ \begin{array}{l}
- \left( 1 - \tau \hat{\delta} \right) i - \frac{\varepsilon}{2} \left( \frac{i}{k} \right)^2 k - \xi w(s) \{ i \notin [-ak, ak] \} \\

+ \mathbb{E}[\Lambda(z'; s)(v^1(z', (1 - \delta)k + i, \xi'; s') + \tau PV(s)(1 - \hat{\delta})(d + i))] \end{array} \right\}
\]

Rearranging terms as before shows that

\[
v^1(\varepsilon, k, \xi; s) + \tau PV(s)d = \pi(\varepsilon, k; s) + \tau PV(s)d + \max_i \left\{ \begin{array}{l}
- (1 - \tau PV(s)) i - \frac{\varepsilon}{2} \left( \frac{i}{k} \right)^2 k - \xi w(s) \{ i \notin [-ak, ak] \} \\

+ \mathbb{E}[\Lambda(z'; s)v^1(z', (1 - \delta)k + i, \xi'; s')] \end{array} \right\}
\]
Subtracting $\tau PV(s)d$ from both sides establishes (18).

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:

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

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

$$i(\varepsilon, k, \xi; s) = \begin{cases} i^a(\varepsilon, k; s) & \text{if } \xi \leq \hat{\xi}(\varepsilon, k; s) \\ i^n(\varepsilon, k; s) & \text{if } \xi > \hat{\xi}(\varepsilon, k; s) \end{cases}$$

where

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

$$i^n(\varepsilon, k; s) = \begin{cases} ak & \text{if } i^a(\varepsilon, k; s) > ak \\ i^a(\varepsilon, k; s) & \text{if } i^a(\varepsilon, k; s) \in [-ak, ak] \\ -ak & \text{if } i^a(\varepsilon, k; s) < -ak \end{cases}$$

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

**Proof.** The form of $i^a(\varepsilon, k; s)$ follows directly from the Bellman equation, using the law of iterated expectations and the fact that $\xi'$ is i.i.d. The form of $i^n(\varepsilon, k; s)$ also follows from the Bellman equation, which shows that the objective function in the no-adjust problem is the same as the adjust problem and the choice set is restricted. The form of $i(\varepsilon, k, \xi; s)$ 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 $\xi$. Therefore, there must be a
cutoff rule. Setting the adjust and no adjust payoffs equal gives the form of the threshold \( \hat{\xi} (\varepsilon, k; s) \).

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

**Proposition 5.** \( v^0(\varepsilon, k; s) \) solves the Bellman equation

\[
v(\varepsilon, k; s) = \pi(\varepsilon, k; s) + \frac{\hat{\xi}(\varepsilon, k; s)}{\xi} \left\{ \begin{array}{l} -(1 - \tau PV(s)) i^a(\varepsilon, k; s) - \frac{\varphi}{2} \left( \frac{i^a(\varepsilon, k; s)}{k} \right)^2 k \\
- \frac{\hat{\xi}(\varepsilon, k; s)}{\xi} w(s) + \mathbb{E}[\Lambda(\varepsilon'; s) v^0(\varepsilon'; (1 - \delta)k + i^a(\varepsilon, k; s); s')] \end{array} \right\} + \left(1 - \frac{\hat{\xi}(\varepsilon, k; s)}{\xi} \right) \left\{ \begin{array}{l} -(1 - \tau PV(s)) i^a(\varepsilon, k; s) - \frac{\varphi}{2} \left( \frac{i^a(\varepsilon, k; s)}{k} \right)^2 k \\
+ \mathbb{E}[\Lambda(\varepsilon'; s) v^0(\varepsilon', (1 - \delta)k + i^a(\varepsilon, k; s); s')] \end{array} \right\}
\]

**Proof.** This follows from integrating \( v^0(\varepsilon, k; s) = \int v^1(\varepsilon, k; \xi; s) \frac{1}{\xi} d\xi \), using the expression for \( v^1(\varepsilon, k; \xi; s) \) from Proposition 3 and the form of the policy function from Proposition 4.

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

\[
v(\varepsilon, k; s) = p(s) v^0(\varepsilon, k; s)
\]

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

**Proposition 6.** The recursive competitive equilibrium from Definition 1 is characterized by a list of functions \( v(\varepsilon, k; s), w(s), p(s), X'(s), \) and \( \mu'(s) \) such that
(i) (Firm optimization) $v(\varepsilon, k; s)$ solves the Bellman equation

$$v(\varepsilon, k; s) = p(s) \pi(\varepsilon, k; s) + \frac{\xi(\varepsilon, k; s)}{\xi} \left\{ -p(s) \left(1 - \tau PV(s)\right) i^n(\varepsilon, k; s) - p(s) \frac{\varphi}{2} \left(\frac{i^n(\varepsilon, k; s)}{k}\right)^2 k \right\}$$

$$+ \left(1 - \frac{\xi(\varepsilon, k; s)}{\xi}\right) \left\{ -p(s) \left(1 - \tau PV(s)\right) i^n(\varepsilon, k; s) - p(s) \frac{\varphi}{2} \left(\frac{i^n(\varepsilon, k; s)}{k}\right)^2 k \right\}$$

$$+ \beta E[v(\varepsilon', (1 - \delta)k + i^n(\varepsilon, k; s); s')]$$

where $i^n(\varepsilon, k; s)$, $i^n(\varepsilon, k; s)$, and $\xi(\varepsilon, k; s)$ are derived from $v(\varepsilon, k; s)$ using

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

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

$$\xi(\varepsilon, k; s) = \frac{1}{p(s)w(s)} \times \left\{ -p(s) \left(1 - \tau PV(s)\right)(i^n(\varepsilon, k; s) - i^n(\varepsilon, k; s)) \right\}$$

and $PV(s)$ is defined by the recursion

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

(ii) (Labor market clearing)

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

where $n(\varepsilon, k; s) = \left(\frac{\varepsilon e^{\varepsilon k^2}\mu}{w(s)}\right)^{\frac{1}{1-q}}$. 
(iii) (Consistency)

\[ p(s) = \left( C(s) - X(s) - \chi \left( \frac{\int w(s) \, \pi}{\chi} \right)^{\frac{1}{1+\eta}} \right)^{-\sigma} \]

where \( C(s) \) is derived from the decision rules by

\[ C(s) = \int (e^{z} e^{k} n(\varepsilon, k; s))^\nu - i(\varepsilon, k; s) - AC(\varepsilon, k; s) \, \mu(d\varepsilon, dk) \]

using \( i(\varepsilon, k; s) = \tilde{\xi}(\varepsilon, k; s) i^a(\varepsilon, k; s) \) and

\[ AC(\varepsilon, k; s) = \tilde{\xi}(\varepsilon, k; s) \left( \frac{\varphi}{2} \left( \frac{i^a(\varepsilon, k; s)}{k} \right)^2 \right) + \left( 1 - \tilde{\xi}(\varepsilon, k; s) \right) \left( \frac{\varphi}{2} \left( \frac{i^a(\varepsilon, k; s)}{k} \right)^2 \right). \]

(iv) (Law of motion for habit stock)

\[ X(s)' = \lambda C(s) \]

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

\[ \mu'(s)(\Delta_\varepsilon \times \Delta_k) = \int p(\varepsilon' \in \Delta_\varepsilon | \varepsilon) \left( \frac{\tilde{\xi}(\varepsilon, k; s)}{\xi} \right) 1 \{ (1 - \delta) k + i^a(\varepsilon, k; s) \in \Delta_k \} + \left( 1 - \tilde{\xi}(\varepsilon, k; s) \right) 1 \{ (1 - \delta) k + i^a(\varepsilon, k; s) \in \Delta_k \} d\varepsilon' \mu(d\varepsilon, dk) \]

Proof. Condition (i) follows from Propositions 3 - 5, using the definition \( v(\varepsilon, k; s) = p(s) v^0(\varepsilon, k; s) \) and noting that \( \Lambda(z'; s) = \frac{\partial p(s')}{p(s)} \). Condition (ii) follows from the household’s FOC, the firms’ FOC, and labor market clearing. Condition (iii) follows from output market clearing and the definition of \( p(s) \). Condition (iv) directly reproduces conditions iv(c) and iv(d) from Section 2.4 in the main text. Condition (v) follows from the original law of motion in condition iv(e) in the main text, eliminating \( d \) as an individual state variable and integrating out \( \xi \).

E Solution Method

I solve the model using the method concurrently developed in Winberry (2018). I provide a brief overview of the method in this appendix and refer to the interested reader to Winberry (2018) for details. Broadly, the method involves three key steps. First, for each period \( t \) I
approximate the equilibrium objects – including the cross-sectional distribution of firms – using a finite-dimensional parametric approximation. Second, I solve for the steady state of this discretized model in which there are no aggregate shocks (but there are still idiosyncratic shocks). Third, I solve for the dynamics of the discretized model by perturbing it around this steady state.

The main challenge in applying the method is approximating the value function \( v_t(\varepsilon, k) \) and distribution \( \mu_t(\varepsilon, k) \) in the first step. I approximate the value function using a weighted sum of Chebyshev polynomials indexed by the vector of weights \( \theta_t \).\(^{36}\) I approximate the density function of the distribution, denoted \( g(\varepsilon, \log(k)) \), using the parametric family

\[
g (\varepsilon, \log(k)) \cong g_0 \exp\{g_1^1 (\varepsilon - m_1^1) + g_1^2 (\log(k) - m_2^2) + \sum_{i=2}^{n_g} \sum_{j=0}^{i} g_i^j \left[ (\varepsilon - m_1^1)^{i-j} (\log(k) - m_2^2)^j - m_i^j \right] \},
\]

where \( n_g \) indexes the degree of approximation, \( \{g_i^j\}_{i,j=(1,0)}^{(n_g,i)} \) are parameters, and \( \{m_i^j\}_{i,j=(1,0)}^{(n_g,i)} \) are centralized moments of the distribution. The fact that the parameters and moments must be consistent provides a convenient method for approximating the law of motion of the distribution. With all of these approximations, the discretized equilibrium of the model is characterized by a sequence of state vectors \( x_t = (m_t, X_t, z_t) \) and control vectors \( y_t = (\theta_t, g_t, p_t, w_t) \) which satisfy

\[
\mathbb{E}_t[f(x_t, x_{t+1}, y_t, y_{t+1})] = 0,
\]

where \( f \) is a function returning equilibrium condition residuals. This is a standard canonical form in the perturbation literature and Winberry (2018) shows how it can be solved using Dynare.

Table 12 shows that “approximate aggregation” does not hold in this model. The table

\(^{36}\)The notation in this discussion follows the exposition of Winberry (2018), which provides further details.
Table 12
Forecast Accuracy Based on Aggregate Capital

<table>
<thead>
<tr>
<th></th>
<th>Maximum DH</th>
<th>Mean DH</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital accumulation $K_{t+1}$</td>
<td>2.4%</td>
<td>0.3%</td>
<td>0.999</td>
</tr>
<tr>
<td>Marginal utility $(\hat{C}_t - X_t)^{-1}$</td>
<td>1.3%</td>
<td>0.2%</td>
<td>0.996</td>
</tr>
<tr>
<td>Real wage $w_t$</td>
<td>0.2%</td>
<td>0.0%</td>
<td>0.999</td>
</tr>
</tbody>
</table>

Notes: results from forecasting using the system of equations (22) - (24). “Maximum DH” refers to the maximum absolute difference between realized series and series forecasting by iterating on (24) for 1,000 periods (as suggested by Den Haan (2010)). “Mean DH” refers to mean absolute difference between these two series. “$R^2$” refers to simple $R^2$ of the regressions.

reports results from the forecasting equations

$$\log(\hat{C}_t - X_t)^{-1} = \alpha_0^C + \alpha_1^C \log z_t + \alpha_2^C K_t$$ (22)
$$\log w_t = \alpha_0^C + \alpha_1^C \log z_t + \alpha_2^C K_t$$ (23)
$$\log K_{t+1} = \alpha_0^C + \alpha_1^C \log z_t + \alpha_2^C K_t.$$ (24)

If approximate aggregation holds, then forecasts of the path of prices (marginal utility and the real wage) based on equations (22) - (24) would be extremely accurate. Following Den Haan (2010), I assess the forecasting power of this system by iterating (24) forward $T = 10,000$ periods to compute a path of capital and then using equations (22) and (23) to compute an implied path of prices. Table 12 shows that the implied forecasts are at times substantially different than the actual prices which occur in equilibrium, which suggests that applying Krusell and Smith (1998)’s methodology would require adding additional moments to accurately summarize the distribution. This approach would be computationally costly and render the simulation-based calibration in Section 4.1 infeasible.

F Business Cycle Analysis Appendix

Table 13 shows that the key results from Table 6 in the main text hold true in three alternative specifications of the model. First, the results hold if households have separable preferences
Table 13
Alternative Model Specifications

<table>
<thead>
<tr>
<th>Model</th>
<th>$\sigma(\text{RI}_t)$</th>
<th>RI 95-5 ratio</th>
<th>90-10 ratio</th>
<th>$d\log I_t/dr_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full model</td>
<td>8.49</td>
<td>31%</td>
<td>24%</td>
<td>-7.55</td>
</tr>
<tr>
<td>Separable preferences</td>
<td>5.06</td>
<td>18%</td>
<td>14%</td>
<td>-7.55</td>
</tr>
<tr>
<td>No taxes</td>
<td>6.90</td>
<td>26%</td>
<td>19%</td>
<td>-9.92</td>
</tr>
<tr>
<td>Lower returns to scale</td>
<td>6.95</td>
<td>26%</td>
<td>20%</td>
<td>-7.29</td>
</tr>
</tbody>
</table>

Notes: responsiveness index $\text{RI}_t$ defined in (15) in the main text. “Full model” refers to the calibrated model. “Separable preferences” refers to the preference specification $E_0 \sum_{t=0}^{\infty} \beta^t \left( \log(C_t - X_t) - \chi \frac{N_t^{1+\eta}}{1+\eta} \right)$. “No taxes” refers to setting $\tau = 0$. “Lower returns to scale” refers to setting $\theta = 0.16$. All variables have been HP-filtered with smoothing parameter $\lambda = 1600$.

over consumption and labor supply:\footnote{37}{I found that using the value of the habit formation parameter $\lambda = 0.73$ with these preferences implies rather unstable aggregate dynamics. Therefore, I set $\lambda = 0.5$ for this exercise, which implies stable dynamics and a roughly similar response of the real interest rate to a TFP shock as in the main parameterization.}

$$E_0 \sum_{t=0}^{\infty} \beta^t \left( \log(C_t - X_t) - \chi \frac{N_t^{1+\eta}}{1+\eta} \right).$$

Second, the results hold if there are no taxes, which is more directly comparable to Khan and Thomas (2008). Third, the results are also stable when using lower returns to scale $\theta = 0.16$ than in the paper $\theta = 0.21$. My reading of this result is that, conditional on generating a similar interest-sensitivity of investment the exact degree of returns to scale is relative unimportant.

G Policy Analysis Appendix

In this appendix, I show how common investment stimulus policies can be mapped into the investment stimulus shock $\omega$ defined in the main text.

Institutional Details I begin with a brief description of the U.S. corporate tax code. Firms pay taxes on their revenues net of business expenses. Most of those expenses are for nondurable inputs such as labor, energy, or materials. These nondurable inputs are fully deducted from the firm’s tax bill because they are completely used in the fiscal year. However, since capital is a durable good, investment expenses are deducted over time. The schedule
for these deductions is given by the IRS' Modified Accelerated Cost Recovery System, or MACRS.

Historically, there have been two main implementations of investment stimulus policies in the U.S: the investment tax credit, which was commonly used before the 1986 tax reform, and the bonus depreciation allowance, which has been used as countercyclical stimulus in the last two recessions. In order to understand how these policies work, consider the example of purchases $1000 in computer equipment.\textsuperscript{38} Table 14 reproduces the depreciation schedule for this $1000 purchase under three regimes: the standard MACRS schedule, a 50% bonus depreciation allowance, and a 5% investment tax credit.

First consider the standard MACRS schedule. The schedule specifies that the recovery period for a computer is five years and also specifies the fraction of the purchase that can be written off each each of those years. This fraction declines over time to reflect the economic depreciation of the computer. At the end of five years, the firm will have written off the full $1000 purchase.\textsuperscript{39}

Now consider how the schedule changes under the two investment stimulus policies. The

\begin{table}[h]
\centering
\caption{TAX DEPRECIATION SCHEDULE}
\begin{tabular}{lcccccc}
\hline
 & \multicolumn{6}{c}{Standard MACRS Schedule (No policy)} \\
 & Year & 0 & 1 & 2 & 3 & 4 & 5 & Total & PV, 7\% \\
\hline
Deductions & 200 & 320 & 192 & 115 & 115 & 58 & 1000 & 890 \\
\hline
\multicolumn{7}{l}{50\% Bonus Depreciation} \\
 & Year & 0 & 1 & 2 & 3 & 4 & 5 & Total & PV, 7\% \\
\hline
Deductions & 500+100 & 160 & 96 & 57.5 & 57.5 & 29 & 1000 & 945 \\
\hline
\multicolumn{7}{l}{5\% Investment Tax Credit} \\
 & Year & 0 & 1 & 2 & 3 & 4 & 5 & Total & PV, 7\% \\
\hline
Deductions & \frac{50}{30\%}+190 & 304 & 182.4 & 109.3 & 109.3 & 55 & 1093 & 1093 \\
\hline
\end{tabular}
\end{table}

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. Example drawn heavily from Table 1 in Zwick and Mahon (2017).

\textsuperscript{38} This example draws heavily from Table 1 in Zwick and Mahon (2017).

\textsuperscript{39} This discussion abstracts from the fact that firms do not pay taxes if they make a loss in that fiscal year; I abstract from loss carryforwards/carrybacks for computational simplicity.
50% bonus depreciation allowance allows the firm to immediately deduct 50% of the $1000, or $500. The firm then applies the standard MACRS schedule to the remaining $500. Hence, the bonus does not change the total amount of tax writeoffs, just their timing. Since more writeoffs are taken in the present, the bonus increases the present value of tax deductions, making investment more attractive to the firm. 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 $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.

**Introducing Stimulus Policy into the Model**  In these two examples, the present value of tax deductions is a useful summary of how various schedules affect the incentive to invest. Proposition 2 shows that, in my model, the present value completely characterizes how the tax code affects the incentive to invest through the tax-adjusted price \( q(s) = 1 - PV(s) \). Any changes in tax depreciation allowances can therefore be mapped into changes in this tax-adjusted price, which I defined as \( \omega \) in the main text. Specifically, the 50% bonus depreciation allowance is captured by \( \omega = 0.5 \times (1 - PV(s)) \); it is as if the firm receives the baseline depreciation schedule on all investment, plus on 50% of its investment gets an extra subsidy. The extra subsidy \( \omega \) is equal to how much the firm values output in the current period, 1, relative to a stream of output through the depreciation schedule, \( PV(s) \). Similarly, the 5% investment tax credit is captured by \( \omega = 0.05 \times \left( \frac{1}{\tau} - PV(s) \right) \); the implicit subsidy \( \omega \) equal how much the firm values the tax writeoff \( \frac{1}{\tau} \) relative to the baseline schedule \( PV(s) \).