Online Appendix for "Lumpy Investment, Business Cycles, and Stimulus Policy" by Thomas Winberry

I. 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 $\overline{\xi}$, 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

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

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

$$1 + r_t = (1-\delta) + z_{t+1}\tilde{\varepsilon}. \tag{2}$$

If $1+r_t < (1-\delta)+z_{t+1}\tilde{\varepsilon}$, then the firms with $\mathbb{E}[\varepsilon_{jt+1}|\varepsilon_{jt}] = \tilde{\varepsilon}$ would strictly prefer to let $k_{jt+1} \to \infty$, violating the finite resource constraint.[1] If $1+r_t > (1-\delta)+z_{t+1}\tilde{\varepsilon}$, then no firms would find it profitable to invest, which would imply $C_{t+1} = 0$ and violate the consumer's Inada condition.

Condition (2) implies that only firms for which $\mathbb{E}[\varepsilon_{jt+1}|\varepsilon_{jt}] = \tilde{\varepsilon}$ 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 μ is the mass of firms with $\mathbb{E}[\varepsilon_{jt+1}|\varepsilon_{jt}] = \tilde{\varepsilon}$ and K_{t+1} is the aggregate capital accumulation implied by the household's Euler equation

$$C_t^{-\sigma} = \beta\left(z_{t+1}\tilde{\varepsilon} + 1 - \delta\right)C_{t+1}^{-\sigma}.$$

Aggregate output in period $t+1$ is therefore $Y_{t+1} = z_{t+1}\tilde{\varepsilon}K_{t+1}$. Hence, aggregate outcomes are identical to a representative firm with production function $Y_{t+1} =$

[1] Note that there is a positive mass of such firms because ε_{jt} has finite support.

$z_{t+1}\tilde{\varepsilon}K_{t+1}$. Note that average productivity among active firms in period $t+1$ is $\tilde{\varepsilon}$ by the law of large numbers.

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

$$\pi^*_t(\alpha) = -k^*_t(\alpha) + \frac{1}{1+r_t}\left(z_{t+1}\widetilde{\varepsilon}k^*_t(\alpha)^\alpha + (1-\delta)k^*_t(\alpha)\right)$$

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

Now consider the limit as $\alpha \to 1$. The optimal policy $k^*_{jt}(\alpha)$ will converge to the optimal policy with $\alpha = 1$ if the fixed cost does not outweigh flow profits, i.e., $\overline{\xi} \leq \pi^*_t(\alpha)$. Since $\pi^*_t(\alpha) \to 0$ as $\alpha \to 1$, this requires $\overline{\xi} \to 0$. Hence, by the same logic as above, in the limit it must be that $r_t + \delta \to z_{t+1}\widetilde{\varepsilon}$ 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 $Y_t = \mu_t \times z_{t+1}\widetilde{\varepsilon}\frac{K_{t+1}}{\mu_t} = z_{t+1}\widetilde{\varepsilon}K_{t+1}$.

II. Data

A. Data Sources and Variable Definitions

I use the following data in the empirical analysis:

- Bureau of Economic Analysis: Fixed Assets Table 1.1 Annual (Bureau of Economic Analysis, 1947-2016*c*).
- Bureau of Economic Analysis: Fixed Assets Table 1.3 Annual (Bureau of Economic Analysis, 1947-2016*d*).
- Bureau of Economic Analysis: Domestic Product and Income Table 1.1.5 Annual (Bureau of Economic Analysis, 1947-2016*a*).
- Bureau of Economic Analysis: Domestic Product and Income Table 1.1.5 Quarterly (Bureau of Economic Analysis, 1947q1-2016q4).
- Bureau of Economic Analysis: Domestic Product and Income Table 1.1.9 Quarterly (Bureau of Economic Analysis, 1947-2016*b*).
- Federal Reserve Bank of San Francisco Total Factor Productivity Series (Federal Reserve Bank of San Francisco, 1947q1-2016q4).
- Board of Governors of the Federal Reserve System: 3-Month Treasury Bill Secondary Market Rate (Board of Governors of the Federal Reserve System, 1954q1-2016q4).

- U.S. Bureau of Labor Statistics: Consumer Price Index for All Urban Consumers All Items in U.S. City Average (U.S. Bureau of Labor Statistics, 1954q1-2016q4*a*).

- Riccardo Dicecio: Relative Price of Investment Goods (DiCecio, 1954q1-2016q4).

- U.S. Bureau of Labor Statistics: Nonfarm Business Sector Hours of All Persons (U.S. Bureau of Labor Statistics, 1954q1-2016q4*c*).

- U.S. Bureau of Labor Statistics: Unemployment Rate (U.S. Bureau of Labor Statistics, 1954q1-2016q4*d*).

- U.S. Bureau of Labor Statistics: Consumer Price Index for All Urban Consumers All Items Less Food and Energy in U.S. City Average (U.S. Bureau of Labor Statistics, 1957q2-2016q4*b*).

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

- Real interest rate r_t: $400\left(\frac{1+r_t^{\text{nom}}}{1+\pi_{t+1}}-1\right)$, where r_t^{nom} is the average yield on 90-day Treasury bills (FRED series DTB3) and π_{t+1} is realized CPI inflation (FRED series CPIAUSCL).

- Relative price of investment goods q_t: constructed by Riccardo DeCicio (FRED series PIRIC).

- 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). Converted each to 2009q1 dollars so units are comparable.

- 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-tfp/`

Table 1—Cyclical Behavior of Risk-Free Rate, Core CPI Inflation

	$\sigma(r_t)$	$\rho(r_t, y_t)$	$\rho(r_t, z_t)$
Whole sample	1.21%	0.09	−0.05
(p-value)		(0.17)	(0.41)
No Volcker	1.01%	0.26***	0.01
		(0.00)	(0.86)
Pre-1983	1.50%	−0.09	0.05
		(0.34)	(0.63)
Post-1983	0.92%	0.46***	−0.23***
		(0.00)	(0.01)

Notes: real interest rate measured as the return on 90-day Treasury bills adjusted for core 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.

B. *Robustness of Empirical Results*

I perform four robustness checks on the empirical results in Table 1. First, I show in Table 1 that the main empirical results hold when I use core CPI rather than headline CPI to correct for inflation. Second, I show in Table 2 that the results hold when inflation expectations are computed from a VAR (rather than realized inflation as in the main text).[2] Third, I show in Figure 1 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 presented in the main text. Fourth, I show in Table 3 that the results hold when I detrend the data using a linear trend, a bandpass filter, or first differences.

III. 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 which 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 χ controls the disutility of labor supply and the $1/\eta$ is the Frisch elasticity.

[2]The VAR contains four lags of inflation, output growth, consumption growth, investment growth, and unemployment.

Table 2—Cyclical Behavior of Risk-Free Rate, VAR Inflation Expectations

	$\sigma(r_t)$	$\rho(r_t, y_t)$	$\rho(r_t, z_t)$
Whole sample	1.19%	0.01	-0.13^{**}
(p-value)		(0.91)	(0.04)
No Volcker	1.07%	0.12^{*}	-0.11^{*}
		(0.07)	(0.10)
Pre-1983	1.16%	-0.06	0.06
		(0.55)	(0.54)
Post-1983	1.21%	0.09	-0.34^{***}
		(0.30)	(0.00)

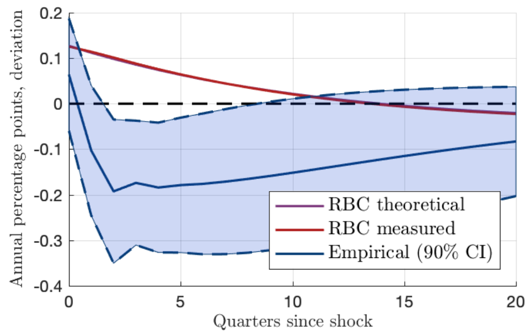
Notes: real interest rate measured as the return on 90-day Treasury bills adjusted for inflation expectations from a VAR, expressed in annual percentage points. The VAR contains four lags of inflation, output growth, consumption growth, investment growth, and unemployment. Output measured as real GDP. 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 3—Cyclical Behavior of Risk-Free Rate, Different Filters

	$\sigma(r_t)$	$\rho(r_t, y_t)$	$\rho(r_t, z_t)$
HP filter	1.73%	-0.11^{*}	-0.20^{***}
(p-value)		(0.09)	(0.00)
Linear trend	2.58%	0.15^{**}	-0.14^{**}
		(0.02)	(0.02)
Bandpass	1.22%	-0.18^{***}	-0.34^{***}
		(0.01)	(0.00)
First differences	2.58%	0.06	0.05
		(0.39)	(0.45)

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.

FIGURE 1. IMPULSE RESPONSE OF THE EX-ANTE REAL INTEREST RATE TO TFP SHOCK



Notes: impulse response of the real interest rate to a TFP shock identified from a bivariate VAR with TFP ordered first. Lag length of 3 chosen by the AIC. Real interest rate measured as the nominal return on 90-day treasury bills adjusted for expected inflation. Expected inflation computed using a VAR with four lags of inflation, output growth, consumption growth, investment growth, and unemployment. "RBC theoretical" refers to the theoretical impulse response. "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.

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 χ 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 V, I also solve for the aggregate dynamics of the heterogeneous firm model using a second-order perturbation in `Dynare`.

IV. Characterizing Equilibrium

In this Appendix I characterize the recursive competitive equilibrium defined in Section II.D. I use this characterization to numerically compute the equilibrium in Appendix V. 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:

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

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

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

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

$$v^1(\varepsilon, k, \xi; \mathbf{s}) = \pi(\varepsilon, k; \mathbf{s}) + \max_i \left\{ \begin{array}{c} -(1 - \tau PV(\mathbf{s})) i - \frac{\varphi}{2} \left(\frac{i}{k}\right)^2 k - \xi w(\mathbf{s}) 1\{i \notin [-ak, ak]\} \\ + \mathbb{E}[\Lambda(z'; \mathbf{s}) v^1(\varepsilon', (1-\delta) k + i, \xi'; \mathbf{s}')] \end{array} \right\} \tag{3}$$

PROOF:

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

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

Collecting terms,

$$\begin{aligned} T(f)(\varepsilon, k, \xi; \mathbf{s}) &= \pi(\varepsilon, k; \mathbf{s}) + \tau \left(\widehat{\delta} + (1-\widehat{\delta}) \mathbb{E}[\Lambda(z'; \mathbf{s}) PV(\mathbf{s}')] \right) d \\ &+ \max_i \left\{ \begin{array}{c} -\left(1 - \tau\widehat{\delta} - \tau(1-\widehat{\delta}) \mathbb{E}[\Lambda(z'; \mathbf{s}) PV(\mathbf{s}')]\right) i - \frac{\varphi}{2} \left(\frac{i}{k}\right)^2 k \\ -\xi w(\mathbf{s}) 1\{i \notin [-ak, ak]\} + \mathbb{E}[\Lambda(z'; \mathbf{s}) f(\varepsilon', (1-\delta)k + i, \xi'; \mathbf{s}')] \end{array} \right\} \end{aligned} \tag{4}$$

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

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

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

Plugging this back into (4) and rearranging gives

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

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

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

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

Rearranging terms as before shows that

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

Subtracting $\tau PV(\mathbf{s})d$ from both sides establishes (3).

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

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

PROPOSITION 2: *The investment decision rule is independent of d and given by*

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

where

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

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

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

PROOF:

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

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

PROPOSITION 3: $v^0(\varepsilon, k; \mathbf{s})$ *solves the Bellman equation*

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

PROOF:

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

A Characterization of the Equilibrium

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

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

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

PROPOSITION 4: *The recursive competitive equilibrium from Definition 1 is characterized by a list of functions $v(\varepsilon, k; \mathbf{s})$, $w(\mathbf{s})$, $p(\mathbf{s})$, $X'(\mathbf{s})$, and $\mu'(\mathbf{s})$ such that*

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

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

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

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

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

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

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

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

2) (Labor market clearing)

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

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

3) (Consistency)

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

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

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

4) (Law of motion for habit stock)

$$X'(\mathbf{s}) = \lambda \left(C(\mathbf{s}) - \chi \frac{\left(\left(\frac{w(\mathbf{s})}{\chi}\right)^{\frac{1}{\eta}}\right)^{1+\eta}}{1+\eta} \right)$$

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

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

PROOF:

Condition (i) follows from Propositions 1 - 3, using the definition $v(\varepsilon, k; \mathbf{s}) = p(\mathbf{s}) v^0(\varepsilon, k; \mathbf{s})$ and noting that $\Lambda(z'; \mathbf{s}) = \frac{\beta p(\mathbf{s}')}{p(\mathbf{s})}$. Condition (ii) follows from the household's FOC, the firms' FOC, and labor market clearing. Condition (iii) follows from output market clearing and the definition of $p(\mathbf{s})$. Condition (iv) directly reproduces conditions iv(c) and iv(d) from Section 2.4 in the main text. Condition (v) follows from the original law of motion in condition iv(e) in the main text, eliminating d as an individual state variable and integrating out ξ.

V. 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 θ_t.[3] I approximate the density function of the distribution, denoted $g(\varepsilon, \log(k))$, using the parametric family

$$g(\varepsilon, \log(k)) \cong g_0 \exp\{g_1^1 \left(\varepsilon - m_1^1\right) + g_1^2 \left(\log(k) - m_1^2\right) + \tag{5}$$

$$\sum_{i=2}^{n_g} \sum_{j=0}^{i} g_i^j \left[\left(\varepsilon - m_1^1\right)^{i-j} \left(\log(k) - m_1^2\right)^j - m_i^j\right]\}, \tag{6}$$

where n_g indexes the degree of approximation, $\left\{g_i^j\right\}_{i,j=(1,0)}^{(n_g,i)}$ are parameters, and $\left\{m_i^j\right\}_{i,j=(1,0)}^{(n_g,i)}$ are centralized moments of the distribution. The fact that the parameters and moments must be consistent with each other implies that the parameters $\mathbf{g}_t$ are pinned down by the moments $\mathbf{m}_t$. I then approximate the law of motion of the distribution using the law of motion for these moments. With all of these approximations, the discretized equilibrium of the model is characterized by a sequence of state vectors $\mathbf{x}_t = (\mathbf{m}_t, X_t, z_t)$ and control vectors $\mathbf{y}_t = (\theta_\mathbf{t}, \mathbf{g}_t, p_t, w_t)$ which satisfy

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

where f is a function returning equilibrium condition residuals. This is a standard canonical form in the perturbation literature and Winberry (2018) shows how it can be solved using `Dynare`. For the analysis of state-dependence in Sections 5 and 6, I solve the model using a second-order perturbation in order to capture the nonlinear aggregate dynamics. However, for the remaining analysis, I solve the model using a first-order perturbation. A first-order perturbation features the same average behavior and is computationally feasible enough to perform the calibration. I have verified that the features of the model targeted in the calibration are nearly indistinguishable in a first- vs. second-order calibration.

Table 4 shows that "approximate aggregation" does not hold in this model. The table reports results from the forecasting equations

$$\log(\hat{C}_t - X_t)^{-1} = \alpha_0^C + \alpha_1^C \log z_t + \alpha_2^C K_t \tag{7}$$

$$\log w_t = \alpha_0^C + \alpha_1^C \log z_t + \alpha_2^C K_t \tag{8}$$

$$\log K_{t+1} = \alpha_0^C + \alpha_1^C \log z_t + \alpha_2^C K_t. \tag{9}$$

If approximate aggregation holds, then forecasts of the path of prices (marginal

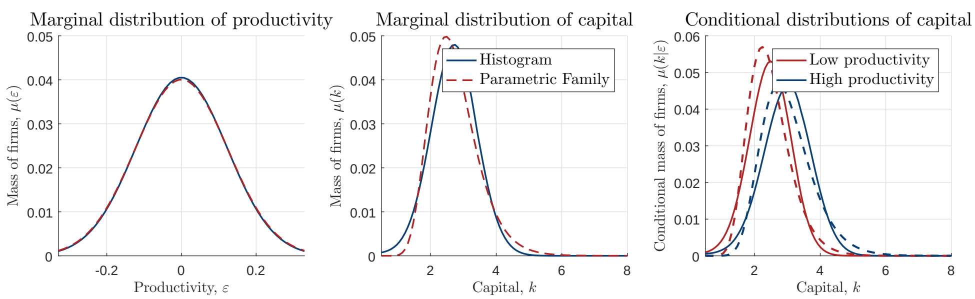
[3]The notation in this discussion follows the exposition of Winberry (2018), which provides further details.

Table 4—Forecast Accuracy Based on Aggregate Capital

	Maximum DH	Mean DH	R^2
Capital accumulation K_{t+1}	1.2%	0.3%	0.999
Marginal utility $(\hat{C}_t - X_t)^{-1}$	3.5%	0.3%	0.996
Real wage w_t	13.0%	1.0%	0.989

Notes: results from forecasting using the system of equations (7) - (9). "Maximum DH" refers to the maximum absolute difference between realized series and series forecasting by iterating on (9) for $10,000$ 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.

Figure 2. Steady State Distribution, Histogram vs. Parametric Family

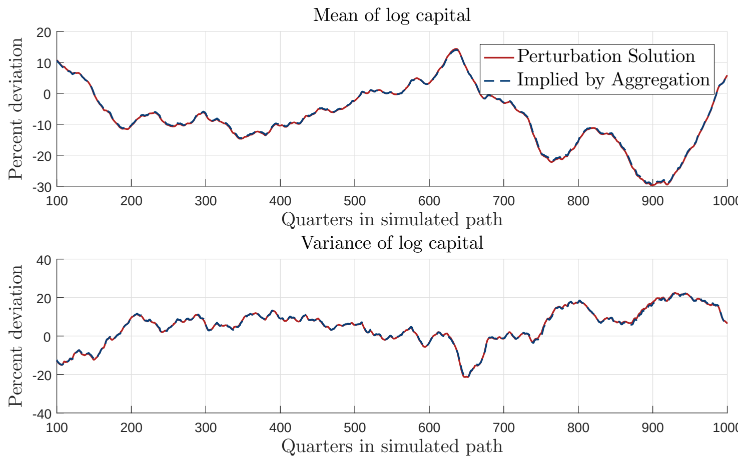


Notes: steady state distribution of firms. "Histogram" is the steady state distribution computed using a fine histogram instead of the parametric family. "High productivity" and "low productivity" correspond to approximately +/− two standard deviations of the distributions of idiosyncratic productivity shocks.

utility and the real wage) based on equations (7) - (9) would be extremely accurate. Following Den Haan (2010), I assess the forecasting power of this system by iterating (9) forward $T = 10,000$ periods to compute a path of capital and then using equations (7) and (8) to compute an implied path of prices. Table 4 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 III.A infeasible.

My method remains accurate even in the absence of approximate aggregation because it approximates the entire distribution of firms. Of course, the key restriction in this approximation is that the distribution is contained within the parametric family (6). Figure 2 shows that the parametric distribution in steady state is a tight fit to the true stationary distribution (computed using a fully non-

Figure 3. Accuracy of Distribution Dynamics



Notes: results from $T = 10,000$ quarter simulation. In each quarter, I use the state variable implied by the approximated solution and compute two objects: (i) the next quarter's moments $\mathbf{m}_{t+1}$ implied by the approximation ("perturbation solution") and (ii) the actual moments in the next quarter computed by aggregating the decision rules exactly ("implied by aggregation"). Lines are percentage deviation from steady state values.

parametric histogram, which is feasible in steady state).[4] In order to assess the accuracy of the dynamics of the distribution, I simulate the model for $T = 10,000$ quarters and, for each quarter t in the simulation, compute two objects: (i) the next quarter's moments $\mathbf{m}_{t+1}$ implied by the approximation and (ii) the actual moments in the next quarter computed by aggregating the decision rules exactly. If the method is not accurate, then the actual moments in step (ii) may fall outside the parametric family. Figure 3 shows that this is not the case; the series (i) and (ii) are nearly indistinguishable, indicating that the true distribution stays within the parametric family (6) in response to aggregate shocks. The correlation between the two series is over 0.997 in all cases.

VI. Business Cycle Analysis Appendix

Table 5 shows that the key results from Table 7 in the main text hold true in four alternative specifications of the model. First, the results hold if households

[4]I compute the histogram over a fine grid following the approach of Young (2010).

Table 5—Alternative Model Specifications

	Impact RI_t			Cumulative $\widehat{RI}_t$	
	95-5 ratio	90-10 ratio	$\rho(RI_t, Y_t)$	95-5 ratio	90-10 ratio
Full model	26.3%	20.3%	0.99	16.2%	12.7%
Separable preferences	15.2%	12.0%	0.95	8.3%	6.4%
No taxes	18.2%	13.9%	0.99	10.4%	8.6%
Lower returns to scale	19.7%	15.3%	0.99	7.1%	5.8%
No habit	25.5%	19.6%	1.00	15.7%	12.2%

Notes: impact responsiveness index RI_t defined in (15) and cumulative responsiveness index $\widehat{RI}_t$ defined in (16) of the main text. adj_t computes the fraction of firms which pay their fixed cost. "Full model" refers to the calibrated model. "Separable preferences" refers to the preference specification $\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\log(C_t - X_t) - \chi \frac{N_t^{1+\eta}}{1+\eta} \right)$. "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 $\lambda = 1600$.

have separable preferences over consumption and labor supply represented by[5]

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\log(C_t - X_t) - \chi \frac{N_t^{1+\eta}}{1+\eta} \right).$$

Second, the results hold if there are no taxes, which is more directly comparable to Khan and Thomas (2008). Third, the results are also similar 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 relatively unimportant. Fourth, the results are similar but slightly smaller in the version of the model without any habit formation at all ($\lambda = 0$).

VII. Policy Analysis Appendix

In this appendix, I show how common investment stimulus policies can be mapped into the investment stimulus shock ω 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

[5] I found that using the value of the habit formation parameter $\lambda = 0.75$ 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.

Table 6—Tax Depreciation Schedule

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

Notes: tax depreciation schedule for purchase of $1000 computer. Top panel: standard schedule absent stimulus policy. Middle panel: 50% bonus depreciation allowance. Bottom panel: 5% investment tax credit. Present value computed using 7% discount rate. Example drawn from Table 1 in Zwick and Mahon (2017).

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's Modified Accelerated Cost Recovery System, or MACRS.

Historically, there have been two main implementations of investment stimulus policies in the US: the investment tax credit, which was often 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 a firm which purchases $1000 in computer equipment.[6] Table 6 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.[7]

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

[6]This example draws heavily from Table 1 in Zwick and Mahon (2017).

[7]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.

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 $q(\mathbf{s}) = 1 - PV(\mathbf{s})$. Any changes in tax depreciation allowances can therefore be mapped into changes in this tax-adjusted price, which I defined as ω in the main text. The 50% bonus depreciation allowance is captured by $\omega = 0.5 \times (1 - PV(\mathbf{s}))$; it is as if the firm receives the baseline depreciation schedule on all investment, plus gets an extra subsidy on 50% of its investment. The extra subsidy ω is equal to how much the firm values output in the current period, 1, relative to a stream of output through the depreciation schedule, $PV(\mathbf{s})$. Similarly, the 5% investment tax credit is captured by $\omega = 0.05 \times \left(\frac{1}{\tau} - PV(\mathbf{s})\right)$; the implicit subsidy ω equals how much the firm values the tax writeoff $\frac{1}{\tau}$ relative to the baseline schedule $PV(\mathbf{s})$.

References

Board of Governors of the Federal Reserve System. 1954q1-2016q4. "3-Month Treasury Bill: Secondary Market Rate [DTB3]." retrieved from FRED Federal Reserve Bank of St. Louis `https://fred.stlouisfed.org/series/DTB3`, used June 1st, 2018 version.

Bureau of Economic Analysis. 1947-2016*a*. "NIPA Domestic Product and Income Table 1.1.5, Annual." `https://apps.bea.gov/iTable/index_nipa.cfm`, used December 21st, 2017 revision.

Bureau of Economic Analysis. 1947-2016*b*. "NIPA Domestic Product and Income Table 1.1.9, Quarterly." `https://apps.bea.gov/iTable/index_nipa.cfm`, used December 21st, 2017 revision.

Bureau of Economic Analysis. 1947-2016*c*. "NIPA Fixed Asset Table 1.1, Annual." `https://apps.bea.gov/iTable/index_nipa.cfm`, used August 23rd, 2017 revision.

Bureau of Economic Analysis. 1947-2016*d*. "NIPA Fixed Asset Table 1.3, Annual." `https://apps.bea.gov/iTable/index_nipa.cfm`, used August 23rd, 2017 revision.

Bureau of Economic Analysis. 1947q1-2016q4. "NIPA Domestic Product and Income Table 1.1.5, Quarterly." `https://apps.bea.gov/iTable/index_nipa.cfm`, used December 21st, 2017 revision.

Den Haan, Wouter. 2010. "Assessing the Accuracy of the Aggregate Law of Motion in Models with Heterogeneous Agents." *Journal of Economic Dynamics and Control*, 1: 79.

DiCecio, Riccardo. 1954q1-2016q4. "Relative Price of Investment Goods [PIRIC]." retrieved from FRED Federal Reserve Bank of St. Louis `https://fred.stlouisfed.org/series/PIRIC`, used June 1st, 2018 version.

Federal Reserve Bank of San Francisco. 1947q1-2016q4. "Total Factor Productivity Series." `https://www.frbsf.org/economic-research/indicators-data/total-factor-productivity-tfp/`, used December 6th, 2017 version.

Khan, Aubhik, and Julia K Thomas. 2008. "Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics." *Econometrica*, 76(2): 395–436.

Krusell, Per, and Anthony A Smith, Jr. 1998. "Income and wealth heterogeneity in the macroeconomy." *Journal of Political Economy*, 106(5): 867–896.

U.S. Bureau of Labor Statistics. 1954q1-2016q4*a*. "Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCSL]." retrieved from FRED Federal Reserve Bank of St. Louis `https://fred.stlouisfed.org/series/CPIAUCSL`, used June 1st, 2018 version.

U.S. Bureau of Labor Statistics. 1954q1-2016q4*c*. "Nonfarm Business Sector: Hours of All Persons [HOANBS]." retrieved from FRED Federal Reserve Bank of St. Louis `https://fred.stlouisfed.org/series/HOANBS`, used June 1st, 2018 version.

U.S. Bureau of Labor Statistics. 1954q1-2016q4*d*. "Unemployment Rate [UNRATE]." retrieved from FRED Federal Reserve Bank of St. Louis `https://fred.stlouisfed.org/series/UNRATE`, used June 1st, 2018 version.

U.S. Bureau of Labor Statistics. 1957q2-2016q4*b*. "Consumer Price Index for All Urban Consumers: All Items Less Food and Energy in U.S. City Average [CPILFESL]." retrieved from FRED Federal Reserve Bank of St. Louis `https://fred.stlouisfed.org/series/CPIFESL`, used June 1st, 2018 version.

Winberry, Thomas. 2018. "A method for solving and estimating heterogeneous agent macro models." *Quantitative Economics*, 9(3): 1123–1151.

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

Zwick, Eric, and James Mahon. 2017. "Tax policy and heterogeneous investment behavior." *American Economic Review*, 107(1): 217–48.