A Toolbox for Solving and Estimating Heterogeneous Agent Macro Models*

Thomas Winberry†

April 2, 2016

Abstract

I develop a new method to solve and estimate heterogeneous agent macro models. The main challenge is that the state vector contains the distribution of microeconomic agents, which is typically infinite-dimensional. I approximate the distribution with a flexible parametric family, reducing the dimensionality to a finite set of parameters, and solve for the dynamics of these parameters by perturbation. I implement the method in `Dynare` and find that it is accurate and extremely efficient. As an illustration, I use the method to estimate a heterogeneous firm model with neutral and investment-specific productivity shocks using Bayesian techniques. The behavior of firms at the micro level matters quantitatively for inference about the aggregate shock processes, suggesting an important role for micro data in estimating macro models.

*This note elaborates on an algorithm developed in my dissertation. I thank my advisors, Richard Rogerson, Greg Kaplan, and Estaban Rossi-Hansberg for many helpful discussions, as well as Mark Aguiar, Jesús Fernández-Villaverde, Ezra Oberfield, Christian Vom Lehn, Michael Weber, and especially Stephen Terry for useful comments along the way. `Dynare` codes and a user guide are available on my website, currently at `http://faculty.chicagobooth.edu/thomas.winberry`.

†University of Chicago Booth School of Business; email: `Thomas.Winberry@chicagobooth.edu`.

1 Introduction

Heterogeneity is pervasive in microeconomic data: households vary tremendously by income, wealth, and consumption, for example, and firms vary by productivity, investment, and hiring. Accordingly, a rapidly growing literature has emerged in macroeconomics which asks how micro heterogeneity matters for our understanding of aggregate business cycles.[1] The models used in this literature are computationally challenging because the aggregate state of the economy contains the distribution of micro agents, generally an infinite-dimensional object. Most existing algorithms, following Krusell and Smith (1998), approximate the distribution with a finite number of moments, typically just the mean. This approximation works well if the mean accurately captures how the distribution affects aggregate dynamics, a condition known as "approximate aggregation." But by construction, approximate aggregation imposes sharp restrictions on how the distribution affects aggregate dynamics, leaving many important questions unanswerable.[2]

In this paper, I develop a new method to solve and estimate heterogeneous agent models that does not rely on approximate aggregation. Instead, I approximate the entire distribution with a finite-dimensional parametric family and include the parameters of that approximation in the state vector. A good approximation of the distribution may require a large number of parameters, however, leaving globally accurate approximation techniques infeasible due to the curse of dimensionality. Instead, I solve for the aggregate dynamics using locally accurate perturbation methods, which are computationally efficient even with a large state space. I show how to implement this perturbation step in `Dynare`, a `Matlab` toolbox designed to solve and estimate representative agent models in a user-friendly way. My website provides a example codes and a user guide for using `Dynare` to solve and estimate heterogeneous agent models as well.

Although the method is applicable to a wide range of heterogeneous agent models, for concreteness I demonstrate it in the context of a real business cycle model with heterogeneous firms and fixed capital adjustment costs, as in Khan and Thomas (2008). In the recursive equilibrium of this model, the aggregate state contains the distribution of firms over productivity and capital, which

[1] There are too many papers to provide a comprehensive list of citations. For recent papers on the household side, see Auclert (2015), Berger and Vavra (2015), or Kaplan, Moll, and Violante (2015); on the firm side, see Bachmann, Caballero, and Engel (2013), Khan and Thomas (2013), Clementi and Palazzo (2015), or Terry (2015a).

[2] A straightforward way to relax approximate aggregation is to extend the number of moments used to approximate the distribution. However, this quickly becomes infeasible due to the curse of dimensionality, as each new moment adds a new state variable.

evolves over time in response to aggregate shocks. The dynamics must satisfy a complicated fixed point problem: each firm's investment decision depends on its expectations of the dynamics of the distribution, and the dynamics of the distribution depend on firms' investment decisions. This type of fixed point problem is at the heart of the computational challenges faced by the heterogeneous agent literature.

My method solves this problem accurately and efficiently; depending on the degree of approximation of the distribution of firms (ranging from 5 to 20 parameters), computing a first order approximation of the aggregate dynamics takes between $50-140$ seconds in `Matlab`.[3] Degrees of approximation on the high end of this range are necessary to capture the shape of the distribution, which features positive skewness and excess kurtosis; however, degrees on the low end of this range are sufficient to capture the dynamics of aggregate variables. I also compute a second-order approximation of the dynamics, but consistent with the results in Khan and Thomas (2008), find quantitatively small nonlinearities in the aggregate. Finally, I consider a parameterization of the model in which approximate aggregation fails to hold and show that as expected my method continues to perform well in this case.

As an illustration, I then incorporate aggregate investment-specific productivity shocks in addition to the neutral shocks already in the model and estimate the parameters of the two shock processes using Bayesian techniques. Characterizing the posterior distribution of parameters using Markov Chain Monte Carlo takes less than 24 minutes using `Dynare`. To understand how micro behavior impacts this estimation of aggregate shock processes, I re-estimate the parameters conditional on different values of the fixed capital adjustment costs, which correspond to different patterns of firm investment behavior. For small adjustment costs, less volatile investment-specific shocks are needed to match the data, while for large adjustment costs, more volatile shocks are needed. Of course, the ideal estimation exercise would incorporate both micro- and macro-level data to jointly estimate the parameters of the model; these results show that doing so is feasible using my method, even with full-information Bayesian techniques.

The `Dynare` codes and user guide is designed to make solving and estimating heterogeneous agent models as simple as possible. Broadly speaking, the user provides two files: a `Matlab .m`

[3] This runtime overstates the time the algorithm spends on solving the model in each step of the estimation, because it includes the time spent processing the model and taking symbolic derivatives. Once these tasks are complete, they do not need to be performed again for different parameter values.

file that computes the stationary equilibrium of the model, and a `Dynare .mod` file that defines the approximate equilibrium conditions of the model. Thanks to the structure of `Dynare`, these conditions can be programmed almost as they would be written in a paper. `Dynare` then differentiates these conditions, computes the locally accurate dynamics around the stationary equilibrium, simulates the model, and if requested, estimates the model using either maximum likelihood or Bayesian techniques. Hence, if a researcher can solve for the stationary equilibrium of their model, these codes will compute the aggregate dynamics almost for free.

Related Literature My method builds heavily on two important papers in the computational literature. The first is Algan et al. (2008), who solve the Krusell and Smith (1998) model by parameterizing the distribution of households using the same parameterization I use in this paper. However, Algan et al. (2008) solve for the aggregate dynamics with globally accurate methods, which are extremely slow in their context. The second paper I build on is Reiter (2009) who, following an idea of Campbell (1998), solves the Krusell and Smith (1998) model using a mix of globally and locally accurate techniques. However, Reiter (2009) approximates the distribution with a fine histogram, which requires many parameters to achieve acceptable accuracy. This limits the approach to problems which have a low-dimensional individual state space because the size of the histogram grows exponentially in the number of individual states. Furthermore, neither Algan et al. (2008) or Reiter (2009) explore using their methods for formal estimation or implement their methods in `Dynare`.[4]

Road Map The rest of this note is organized as follows. I briefly describe the benchmark heterogeneous firm framework in Section 2. I then explain my solution method in the context of this benchmark model in Section 3. In Section 4, I add investment-specific shocks to the model, and estimate the parameters of the shock processes using Bayesian techniques. Section 5 concludes. Various appendices contain additional details not contained in the main text.

[4] Veracierto (2016) also proposes a method based on a mix of globally and locally accurate techinques that does not rely on any direct approximation of the distribution. Instead, Veracierto (2016) approximates the history of individual agents' decision rules, and simulates a panel of agents to compute the distribution at any point in time. He then linearizes the system with respect to the history of approximated decision rules and uses that to compute the evolution of the distribution. The advantage of his methodology is that it does not require any approximation of the distribution. However, the disadvantage is that it is extremely computationally intensive, precluding the possibility of estimation.

2 Benchmark Real Business Cycle Model with Firm Heterogeneity

Although I illustrate my method using the heterogeneous firm model from Khan and Thomas (2008), the method itself applies to a large class of heterogeneous agent models.[5] In the online codes and user guide, I use the method to solve the heterogeneous household model from Krusell and Smith (1998), and discuss how to generalize the method to solve other models as well.

2.1 Environment

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

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

where z_t is an aggregate productivity shock, ε_{jt} is an idiosyncratic productivity shock, k_{jt} is capital, n_{jt} is labor, θ is the elasticity of output with respect to capital, and ν with respect to labor. The aggregate shock z_t is common to all firms and follows the AR(1) process

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

The idiosyncratic shock ε_{jt} is independently distributed across firms, but within firms follows the AR(1) process

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

Each period, the firm j inherits its capital stock from previous periods' investments, observes the two productivity shocks, hires labor from a competitive market, and produces output.

After production, the firm invests in capital for the next period. Gross investment i_{jt} yields $k_{jt+1} = (1-\delta)\, k_{jt} + i_{jt}$ units of capital next period, where δ is the depreciation rate of capital. If $\frac{i_{jt}}{k_{jt}} \notin [-a, a]$, the firm must pay a fixed adjustment cost ξ_{jt} in units of labor. The parameter a governs a region around zero investment within which firms do not incur the fixed cost. The fixed cost ξ_{jt} is a random variable distributed $U[0, \overline{\xi}]$, independently over firms and time.

[5] Since the benchmark is directly taken from Khan and Thomas (2008), I keep my exposition brief, and refer the interested reader to their original paper for details.

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

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

where C_t is consumption, N_t is labor supplied to the market, β is the discount factor, σ is the coefficient of relative risk aversion, χ governs the disutility of labor supply, and α is the Frisch elasticity of labor supply. The total time endowment per period is normalized to 1, so $N_t \in [0,1]$. The household owns all the firms in the economy and markets are complete.

2.2 Firm Optimization

Following Khan and Thomas (2008), I directly incorporate the implications of household optimization into the firm's optimization problem by approximating the transformed value function

$$\begin{aligned}\widehat{v}(\varepsilon,k;\mathbf{s}) &= \lambda(\mathbf{s})\max_{n}\left\{e^{z}e^{\varepsilon}k^{\theta}n^{\nu}-w(\mathbf{s})\,n\right\} \\ &\quad +E_{\xi}\left[\max\left\{v^{a}(\varepsilon,k;\mathbf{s})-\xi\lambda(\mathbf{s})\,w(\mathbf{s}),v^{n}(\varepsilon,k;\mathbf{s})\right\}\right],\end{aligned} \tag{1}$$

where $\mathbf{s}$ is the aggregate state vector (defined in Section 2.3 below), $\lambda(\mathbf{s}) = C(\mathbf{s})^{-\sigma}$ is the marginal utility of consumption in equilibrium, and

$$v^{a}(\varepsilon,k;\mathbf{s}) = \max_{k'\in\mathbb{R}} -\lambda(\mathbf{s})\left(k'-(1-\delta)\,k\right)+\beta E\left[\widehat{v}(\varepsilon',k';\mathbf{s}'\left(z';\mathbf{s}\right)|\varepsilon,k;\mathbf{s}\right] \tag{2}$$

$$v^{n}(\varepsilon,k;\mathbf{s}) = \max_{k'\in[(1-\delta-a)k,(1-\delta+a)k]} -\lambda(\mathbf{s})\left(k'-(1-\delta)\,k\right)+\beta E\left[\widehat{v}(\varepsilon',k';\mathbf{s}'\left(z';\mathbf{s}\right)|\varepsilon,k;\mathbf{s}\right]. \tag{3}$$

Denote the unconstrained capital choice from (2) by $k^a(\varepsilon,k;\mathbf{s})$ and the constrained choice from (3) by $k^n(\varepsilon,k;\mathbf{s})$. The firm will choose to pay the fixed cost if and only if $v^a(\varepsilon,k;\mathbf{s}) - \xi\lambda(\mathbf{s})\,w(\mathbf{s}) \geq v^n(\varepsilon,k;\mathbf{s})$. Hence, there is a unique threshold which makes the firm indifferent between these two options,

$$\widetilde{\xi}(\varepsilon,k;\mathbf{s}) = \frac{v^{a}(\varepsilon,k;\mathbf{s})-v^{n}(\varepsilon,k;\mathbf{s})}{\lambda(\mathbf{s})\,w(\mathbf{s})}. \tag{4}$$

Denote $\widehat{\xi}(\varepsilon,k;\mathbf{s})$ as the threshold bounded to be within the support of ξ, i.e., $\widehat{\xi}(\varepsilon,k;\mathbf{s}) = \min\{\max\{0, \widetilde{\xi}(\varepsilon,k;\mathbf{s}),\bar{\xi}\}\}$.

2.3 Equilibrium

In the recursive competitive equilibrium, the aggregate state $\mathbf{s}$ contains the current draw of the aggregate productivity shock, z, and the distribution of firms over (ε, k)-space, μ.

Definition 1 *A* ***recursive competitive equilibrium*** *for this model is a set* $\widehat{v}(\varepsilon, k; \mathbf{s})$, $n(\varepsilon, k; \mathbf{s})$, $k^a(\varepsilon, k; \mathbf{s})$, $k^n(\varepsilon, k; \mathbf{s})$, $\widehat{\xi}(\varepsilon, k; \mathbf{s})$, $\lambda(\mathbf{s})$, $w(\mathbf{s})$, *and* $\mathbf{s}'(z'; \mathbf{s}) = (z'; \mu'(z, \mu))$ *such that*

1. *(Firm optimization) Taking* $w(\mathbf{s})$, $\lambda(\mathbf{s})$, *and* $\mathbf{s}'(z'; \mathbf{s})$ *as given,* $\widehat{v}(\varepsilon, k; \mathbf{s})$, $n(\varepsilon, k; \mathbf{s})$, $k^a(\varepsilon, k; \mathbf{s})$, $k^n(\varepsilon, k; \mathbf{s})$, *and* $\widehat{\xi}(\varepsilon, k; \mathbf{s})$ *solve the firm's optimization problem (1) - (4).*

2. *(Implications of household optimization)*

   - $\lambda(\mathbf{s}) = C(\mathbf{s})^{-\sigma}$, *where* $C(\mathbf{s}) = \int [e^z e^\varepsilon k^\theta n(\varepsilon, k; \mathbf{s})^\nu + (1-\delta) k - \left(\frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\bar{\xi}}\right) k^a(\varepsilon, k; \mathbf{s}) - \left(1 - \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\bar{\xi}}\right) k^n(\varepsilon, k; \mathbf{s})\ ]\ d\mu(\varepsilon, k)$.
   - $w(\mathbf{s})$ *satisfies* $\int (n(\varepsilon, k; \mathbf{s}) + \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})^2}{2\bar{\xi}}) d\mu(\varepsilon, k) = \left(\frac{w(\mathbf{s})\lambda(\mathbf{s})}{\chi}\right)^{\frac{1}{\alpha}}$.

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

$$\begin{aligned} \mu'(z, \mu)(\Delta_\varepsilon \times \Delta_k) &= \int p(\rho_\varepsilon \varepsilon + \sigma_\varepsilon \omega^\varepsilon \in \Delta_\varepsilon) d\omega^\varepsilon \times [\left(\frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\bar{\xi}}\right) 1\{k^a(\varepsilon, k; \mathbf{s}) \\ &\in \Delta_k\} + \left(1 - \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\bar{\xi}}\right) 1\{k^n(\varepsilon, k; \mathbf{s}) \in \Delta_k\}] d\mu(\varepsilon, k), \end{aligned} \tag{5}$$

   where p *is the p.d.f. of idiosyncratic productivity shocks.*

4. *(Law of motion for aggregate shocks)* $z' = \rho_z z + \omega_z'$, *where* $\omega_z' \sim N(0, \sigma_z)$.

2.4 Baseline Parameterization

The baseline parameter values, reported in Table 1, are those chosen by Khan and Thomas (2008), adjusted to reflect the fact that my model does not feature trend growth. The model period is one year, and the utility function corresponds to indivisible labor. The firm-level adjustment costs and idiosyncratic shock process were chosen to match features of the investment rate distribution reported in Cooper and Haltiwanger (2006).

Table 1: Baseline Parameterization

Parameter	**Description**	**Value**	**Parameter**	**Description**	**Value**
β	Discount factor	.961	ρ_z	Aggregate TFP AR(1)	.859
σ	Utility curvature	1	σ_z	Aggregate TFP AR(1)	.014
α	Inverse Frisch	$\lim \alpha \to 0$	$\overline{\xi}$	Fixed cost	.0083
χ	Labor disutility	$N^* = \frac{1}{3}$	a	No fixed cost region	.011
ν	Labor share	.64	ρ_ε	Idiosyncratic TFP AR(1)	.859
θ	Capital share	.256	σ_ε	Idiosyncratic TFP AR(1)	.022
δ	Capital depreciation	.085			

Notes: Parameterization follows Khan and Thomas (2008), Table 1, adjusted for the fact that my model does not feature trend growth.

3 Using the Method to Solve the Benchmark Model

In this section, I show how to solve the benchmark model in three distinct steps. First, I approximate the equilibrium objects using finite-dimensional approximations. In particular, I approximate the distribution using a flexible parametric family, so that the approximation is pinned down by the parameters of that family, and I approximate the value function using a weighted sum of polynomials, so the approximation is pinned down by the coefficients on those polynomials. This yields a set of finite-dimensional approximate equilibrium conditions. Second, I compute the stationary equilibrium of the model with no aggregate shocks. Finally, I solve for the dynamics of these variables around their stationary values using locally accurate perturbation methods. This is completely analogous to solving a representative agent model by perturbation, except that the endogenous variables include the distribution parameters and polynomial coefficients. Throughout this section, I focus on the new features of the method, such as the approximation of the distribution and the perturbation for aggregate dynamics; for further details, see Appendix A.

3.1 Step 1: Approximate Equilibrium Conditions Using Finite-Dimensional Objects

Distribution Following Algan et al. (2008), I approximate the p.d.f. of the distribution of firms, denoted $g(\varepsilon, k)$, by

$$
\begin{aligned}
g(\varepsilon, k) \quad \cong \quad & g_0 \exp\{g_1^1 \left(\varepsilon - m_1^1\right) + g_1^2 \left(k - m_1^2\right) + \\
& \sum_{i=2}^{n_g} \sum_{j=0}^{i} g_i^j \left[\left(\varepsilon - m_1^1\right)^{i-j} \left(k - m_1^2\right)^j - m_i^j\right]\},
\end{aligned} \tag{6}
$$

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 parameters $\mathbf{g}$ and moments $\mathbf{m}$ must be consistent in the sense that the moments are actually implied by the parameters:[6]

$$
\begin{aligned}
m_1^1 &= \int \varepsilon g(\varepsilon, k)\, d\varepsilon dk, \\
m_1^2 &= \int k g(\varepsilon, k)\, d\varepsilon dk, \text{ and} \\
m_i^j &= \int \left(\varepsilon - m_1^1\right)^{i-j} \left(k - m_1^2\right)^j g(\varepsilon, k)\, d\varepsilon dk \text{ for } i = 2, ..., n_g,\ j = 0, ..., i.
\end{aligned} \tag{7}
$$

Hence, given the vector of moments $\mathbf{m}$, the parameters $\mathbf{g}$ are pinned down by (7). I therefore use the moments $\mathbf{m}$ as my characterization of the distribution, and approximate the infinite-dimensional aggregate state (z, μ) with $(z, \mathbf{m})$.[7]

To derive the law of motion for the approximate aggregate state, note that the current distribution $\mathbf{m}$ and decision rules pin down the p.d.f. of firms in the next period, $g'(\varepsilon', k'; z, \mathbf{m})$, through the identity:

$$
g'\left(\varepsilon', k'; z, \mathbf{m}\right) = \int \left[\begin{array}{c} 1\left\{\rho_\varepsilon \varepsilon + \sigma_\varepsilon \omega_\varepsilon' = \varepsilon'\right\} \times \left[\frac{\widehat{\xi}(\varepsilon,k;z,\mathbf{m})}{\overline{\xi}} 1\left\{k^a(\varepsilon, k; z, \mathbf{m}) = k'\right\}\right. \\ \left. + \left(1 - \frac{\widehat{\xi}(\varepsilon,k;z,\mathbf{m})}{\overline{\xi}}\right) 1\left\{k^n(\varepsilon, k; z, \mathbf{m}) = k'\right\}\right] \end{array}\right] p\left(\omega_\varepsilon'\right) g(\varepsilon, k; \mathbf{m})\, d\omega_\varepsilon' d\varepsilon dk, \tag{8}
$$

where p is the standard normal p.d.f. Because of the convolution with decision rules, the new p.d.f.

[6] The normalization g_0 is chosen so that the total mass of the p.d.f. is 1.

[7] The distribution in the benchmark model is continuous, but in the online user guide, I show how to extend this family to include mass points. Essentially, this adds one parameter specifying the location of the mass point and another specifying the mass itself.

$g'(\varepsilon', k'; z, \mathbf{m})$ is not necessarily in the parametric family (6). I therefore approximate the law of motion (8) by choosing $\mathbf{m}'(z, \mathbf{m})$ to match the moments of the true p.d.f. $g'(\varepsilon', k'; \mathbf{m})$:

$$
\begin{aligned}
m_1^{1\prime} &= \int (\rho_\varepsilon \varepsilon + \omega_\varepsilon') p\left(\omega_\varepsilon'\right) g\left(\varepsilon, k; \mathbf{m}\right) d\omega_\varepsilon' d\varepsilon dk \qquad (9) \\
m_1^{2\prime} &= \int \left[\begin{array}{c} \frac{\widehat{\xi}(\varepsilon,k;z,\mathbf{m})}{\bar{\xi}} k^a\left(\varepsilon, k; z, \mathbf{m}\right) \\ + \left(1 - \frac{\widehat{\xi}(\varepsilon,k;z,\mathbf{m})}{\bar{\xi}}\right) k^a\left(\varepsilon, k; z, \mathbf{m}\right) \end{array} \right] p\left(\omega_\varepsilon'\right) g\left(\varepsilon, k; \mathbf{m}\right) d\omega_\varepsilon' d\varepsilon dk \\
m_i^{j\prime}\left(z, \mathbf{m}\right) &= \int \left[\begin{array}{c} (\rho_\varepsilon \varepsilon + \omega_\varepsilon' - m_1^{1\prime})^{i-j} \{ \frac{\widehat{\xi}(\varepsilon,k;z,\mathbf{m})}{\bar{\xi}} \left(k^a\left(\varepsilon, k; z, \mathbf{m}\right) - m_1^{2\prime}\right)^j \\ + \left(1 - \frac{\widehat{\xi}(\varepsilon,k;z,\mathbf{m})}{\bar{\xi}}\right) \left(k^a\left(\varepsilon, k; z, \mathbf{m}\right) - m_1^{2\prime}\right)^j \} \end{array} \right] p\left(\omega_\varepsilon'\right) g\left(\varepsilon, k; \mathbf{m}\right) d\omega_\varepsilon' d\varepsilon dk.
\end{aligned}
$$

In practice, I compute this integral numerically using two-dimensional Gauss-Legendre quadrature, which replaces the integral with a finite sum.

Firm's Value Functions Given this approximation of the aggregate state, I approximate firms' value functions by

$$
v\left(\varepsilon, k; z, \mathbf{m}\right) \cong \sum_{i=1}^{n_\varepsilon} \sum_{j=1}^{n_k} \theta_{ij}\left(z, \mathbf{m}\right) T_i\left(\varepsilon\right) T_j\left(k\right),
$$

where n_ε and n_k define the order of approximation, $T_i(\varepsilon)$ and $T_j(k)$ are Chebyshev polynomials, and $\theta_{ij}(z, \mathbf{m})$ are coefficients on those polynomials.[8] I solve for the dependence of these coefficients on the aggregate state using perturbation in Section 3.3.

With this particular approximation of the value function, it is natural to approximate the Bellman equation (1) using collocation, which forces the equation to hold exactly at a set of grid points $\{\varepsilon_i, k_j\}_{i,j=1,1}^{n_\varepsilon, n_k}$:

$$
\widehat{v}\left(\varepsilon_i, k_j; z, \mathbf{m}\right) = \lambda\left(z, \mathbf{m}\right) \max_n \left\{ e^z e^{\varepsilon_i} k_j^\theta n^\nu - w\left(z, \mathbf{m}\right) n \right\} + \lambda\left(z, \mathbf{m}\right)\left(1 - \delta\right) k \qquad (10)
$$

[8] Technically, the Chebyshev polynomials are only defined on the interval $[-1, 1]$, so I rescale the state variables to this interval.

$$+\left(\frac{\widehat{\xi}\left(\varepsilon_i,k_j;z,\mathbf{m}\right)}{\bar{\xi}}\right)\left(\begin{array}{c}-\lambda\left(z,\mathbf{m}\right)\left(k^a\left(\varepsilon_i,k_j;z,\mathbf{m}\right)-w\left(z,\mathbf{m}\right)\frac{\widehat{\xi}\left(\varepsilon_i,k_j;z,\mathbf{m}\right)}{2}\right)\\ \\ +\beta E_{z'|z}\left[\int\widehat{v}\left(\rho_\varepsilon\varepsilon_i+\omega'_\varepsilon,k^a\left(\varepsilon_i,k_j;z,\mathbf{m}\right);z',\mathbf{m}'\left(z,\mathbf{m}\right)\right)p\left(\omega'_\varepsilon\right)d\omega'_\varepsilon\right]\end{array}\right)$$

$$+\left(1-\frac{\widehat{\xi}\left(\varepsilon_i,k_j;z,\mathbf{m}\right)}{\bar{\xi}}\right)\left(\begin{array}{c}-\lambda\left(z,\mathbf{m}\right)k^n\left(\varepsilon_i,k_j;z,\mathbf{m}\right)\\ \\ +\beta E_{z'|z}\left[\int\widehat{v}\left(\rho_\varepsilon\varepsilon_i+\omega'_\varepsilon,k^n\left(\varepsilon_i,k_j;z,\mathbf{m}\right);z',\mathbf{m}'\left(z,\mathbf{m}\right)\right)p(\omega'_\varepsilon)d\omega'_\varepsilon\right]\end{array}\right),$$

where the decision rules are computed from the value function via first order conditions.[9] Note that the conditional expectation of the future value function has been broken into its component pieces: the expectation with respect to idiosyncratic shocks is taken explicitly by integration, and the expectation with respect to aggregate shocks implicitly through the expectation operator. I compute the expectation with respect to idiosyncratic shocks using Gauss-Hermite quadrature, and will compute the expectation with respect to aggregate shocks using perturbation in Section 3.3.

Approximate Equilibrium Conditions With all of these approximations, the recursive equilibrium in Definition 1 becomes computable, replacing the true aggregate state (z,μ) with the approximate aggregate state $(z,\mathbf{m})$, the true Bellman equation (1) with the Chebyshev collocation approximation (10), and the true distribution law of motion with the approximation (9). I show in Appendix A that these approximate equilibrium conditions can be represented by a function $f:\mathbb{R}^{2n_\varepsilon n_k+n_g+2}\times\mathbb{R}^{2n_\varepsilon n_k+n_g+2}\times\mathbb{R}^{n_g+1}\times\mathbb{R}^{n_g+1}\rightarrow\mathbb{R}^{2n_\varepsilon n_k+n_g+2+n_g+1}$ which satisfies

$$E_{\omega'_z}\left[f\left(\mathbf{y}',\mathbf{y},\mathbf{x}',\mathbf{x}\right)=0\right], \tag{11}$$

where $\mathbf{y}=(\boldsymbol{\theta},\mathbf{k}^a,\mathbf{g},\lambda,w)$ are the control variables, $\mathbf{x}=(z,\mathbf{m})$ are the state variables, ψ is the perturbation parameter, and $\mathbf{k}^a$ denotes the target capital stock along the collocation grid. This is exactly the canonical form in Schmitt-Grohé and Uribe (2004), who show how to solve such systems

[9] The choice of Chebyshev collocation is not essential, and the online user guide explains how to use other approximations, such as splines.

using perturbation methods. A solution to this system is of the form

$$\begin{aligned}\mathbf{y} &= g(\mathbf{x};\psi)\\ \mathbf{x}' &= h(\mathbf{x};\psi)+\psi\times\eta\omega_z',\end{aligned}$$

where $\eta = \left(1, \mathbf{0}_{n_g\times 1}\right)'$.

Perturbation methods approximate the solution g and h using Taylor expansions around the point where $\psi = 0$, which corresponds to the stationary equilibrium with no aggregate shocks. For example, a first order Taylor expansion gives:

$$\begin{aligned}g(\mathbf{x};1) &\cong g_{\mathbf{x}}(\mathbf{x}^*;0)(\mathbf{x}-\mathbf{x}^*)+g_{\psi}(\mathbf{x}^*;0)\\ h(\mathbf{x};1) &\cong h_{\mathbf{x}}(\mathbf{x}^*;0)(\mathbf{x}-\mathbf{x}^*)+h_{\psi}(\mathbf{x}^*;0)+\eta\omega_z'.\end{aligned} \tag{12}$$

The unknowns in this approximation are the partial derivatives $g_{\mathbf{x}}$, g_{ψ}, $h_{\mathbf{x}}$, and h_{ψ}. Schmitt-Grohé and Uribe (2004) show how to solve for these partial derivatives from the partial derivatives of the equilibrium conditions, $f_{\mathbf{y}'}$, $f_{\mathbf{y}}$, $f_{\mathbf{x}'}$, $f_{\mathbf{x}}$, and f_{ψ}, evaluated at the stationary equilibrium with $\psi = 0$. Since this procedure is by now standard, I refer the interested reader to Schmitt-Grohé and Uribe (2004) for further details. Given values for $\mathbf{x}^*$ and $\mathbf{y}^*$, `Dynare` efficiently implements this procedure completely automatically. Hence, to solve for the aggregate dynamics, we just need to compute the stationary equilibrium $\mathbf{x}^*$ and $\mathbf{y}^*$, and plug this into `Dynare` to compute the Taylor expansion (12). An analogous procedure can be used to compute higher-order approximations of g and h, with no additional coding required.

3.2 Step 2: Compute Stationary Equilibrium with No Aggregate Shocks

In terms of Schmitt-Grohé and Uribe (2004)'s canonical form (11), the stationary equilibrium is represented by two vectors $\mathbf{x}^*=(0,\mathbf{m}^*)$ and $\mathbf{y}=(\boldsymbol{\theta}^*,\mathbf{k}^{a*},\mathbf{g}^*,\lambda^*,w^*)$ such that

$$f(\mathbf{y}^*,\mathbf{y}^*,\mathbf{x}^*,\mathbf{x}^*)=0.$$

In principle, this is a system of nonlinear equations that can be solved numerically; in practice, this system is large, so numerical solvers fail to converge. I instead solve this system using a stable iterative scheme described in Appendix A, similar to the algorithm developed in Hopenhayn and Rogerson (1993).

Figure 1 shows that a moderately high degree of approximation is necessary to capture the shape of the invariant distribution. The figure plots various slices of this invariant distribution for different degrees of approximation, and compares them to an "exact" histogram.[10] The marginal distribution is productivity in Panel (a) is normal, so a second degree approximation gives an exact match due to the functional form of the family (6). However, the marginal distribution of capital in Panel (b) features positive skewness and excess kurtosis, which requires a higher degree approximation. Additionally, a second degree approximation implies that the conditional distributions of capital by productivity in Panels (c) and (d) only vary in their location, while the true distribution varies in both location and scale. A $n_g = 6$ degree approximation captures these complicated shapes almost exactly.

In contrast, even low degree approximations of the distribution provide a good approximation of key aggregate variables. Table 2 computes various aggregates in the stationary equilibrium, using different degrees of approximation, and again compares them to the values obtained using an "exact" histogram. A second degree approximation, which fails to capture the non-normal features of the distribution, nevertheless yields aggregates which are virtually indistinguishable from higher degree approximations.

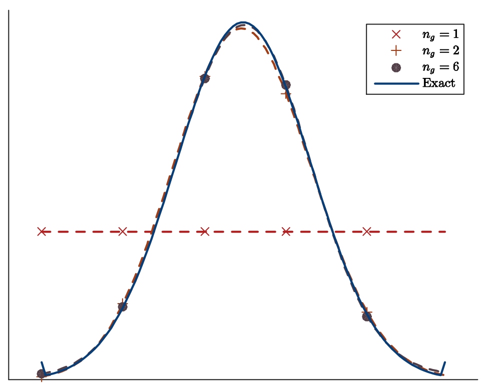
3.3 Step 3: Compute Aggregate Dynamics Using Perturbation

Given the values for $\mathbf{x}^*$ and $\mathbf{y}^*$ in the stationary equilibrium, it is straightforward to compute the Taylor expansions of g and h dynamics using `Dynare`. Solving for these dynamics involves two main steps: first, computing the derivatives of the equilibrium conditions (11), which gives a linear system of equations for the partial derivatives of the solution g and h, and second, solving the linear system for those partial derivatives. `Dynare` computes the derivatives of the equilibrium conditions using symbolic differentiation and solves the linear system using standard linear rational

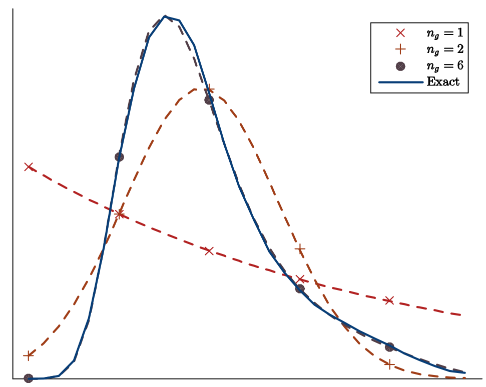
[10] In particular, I use the same iterative scheme, but approximate the distribution with a histogram which simply records the mass of firms at points along a fine grid as in Young (2010).

Figure 1: Invariant Distribution for Different Degrees of Distribution Approximation

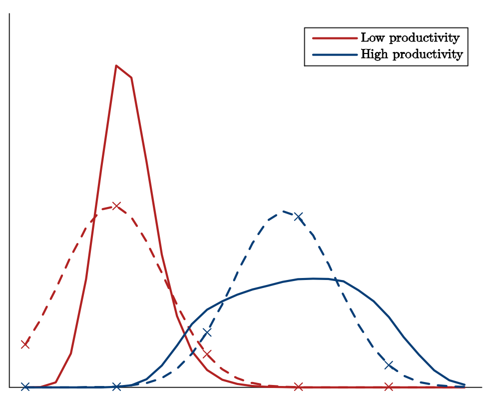
(a) Marginal distribution of productivity



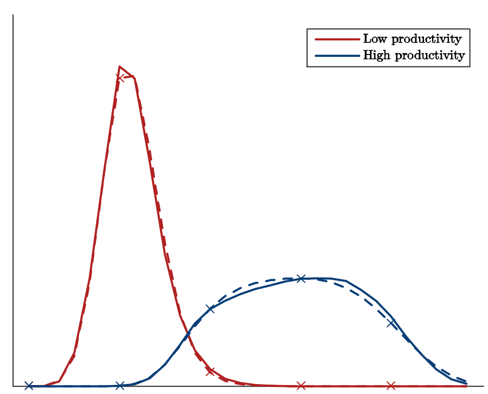
(b) Marginal distribution of capital



(c) Conditional distributions of capital, $n_g = 2$



(d) Conditional distributions of capital, $n_g = 6$



Notes: Slices of invariant distribution of firms over productivity ε and capital k. "Exact" refers to nonparametric histogram, following Young (2010). n_g refers to highest order moment used in parametric family (6). Marginal distributions computed by numerical integration of joint p.d.f. "High productivity" and "Low productivity" correspond to roughly + / - two standard deviations of the productivity distribution.

Table 2: Aggregates in Stationary Equilibrium, for Different Degrees of Distribution Approximation

Variable	$\mathbf{n}_g = 1$	$\mathbf{n}_g = 2$	$\mathbf{n}_g = 3$	$\mathbf{n}_g = 4$	$\mathbf{n}_g = 5$	**Exact**
Output	0.522	0.500	0.499	0.499	0.499	0.499
Consumption	0.420	0.413	0.413	0.412	0.412	0.412
Investment	0.100	0.087	0.086	0.086	0.086	0.086
Capital	1.178	1.023	1.015	1.014	1.013	1.015
Wage	0.945	0.962	0.961	0.961	0.961	0.961
Marginal Utility	2.382	2.422	2.423	2.426	2.426	2.427

Notes: Aggregates in stationary equilibrium computed using various orders of approximation. "Exact" refers to distribution approximated with fine histogram, as in Young (2008).

expectation model solvers, such as Anderson and Moore (1985) or Sims (2001); see Adjemian (2011) for more details.[11] `Dynare` will then, if requested, simulate the solution (if necessary, pruning as in Andreasen, Fernández-Villaverde, and Rubio-Ramírez (2014)), compute theoretical and/or empirical moments of simulated variables, or estimate the model using likelihood-based methods.

First Order Approximation Computing a first-order approximation of the dynamics takes between 50 – 140 seconds on a Dell workstation, depending on the degree of approximation for the distribution. Table 3 reports these run times and breaks them down into the fractions spent on various tasks. For all degrees of the distribution approximation, the majority of time is spent "pre-processing" the model, during which time `Dynare` reads in the model file and computes the symbolic derivatives. This is a fixed cost that does not need to be performed again for different parameter values. The remaining time is spent solving for the stationary equilibrium of the model ("stationary equilibrium") and computing the first-order approximation ("perturbation"). For higher degree approximations of the distribution, pre-processing takes up relatively more computation time, and the stationary equilibrium and perturbation steps less time.

The resulting dynamics of key aggregate variables are well in line with what Khan and Thomas (2008) and Terry (2015b) have reported using different algorithms to solve this model. Figure 2

[11] Moving to higher order approximations requires solving additional equations, but as described in Schmitt-Grohé and Uribe (2004) these additional systems are linear, and thus simple to solve.

Table 3: Computing Time for First Order Approximation

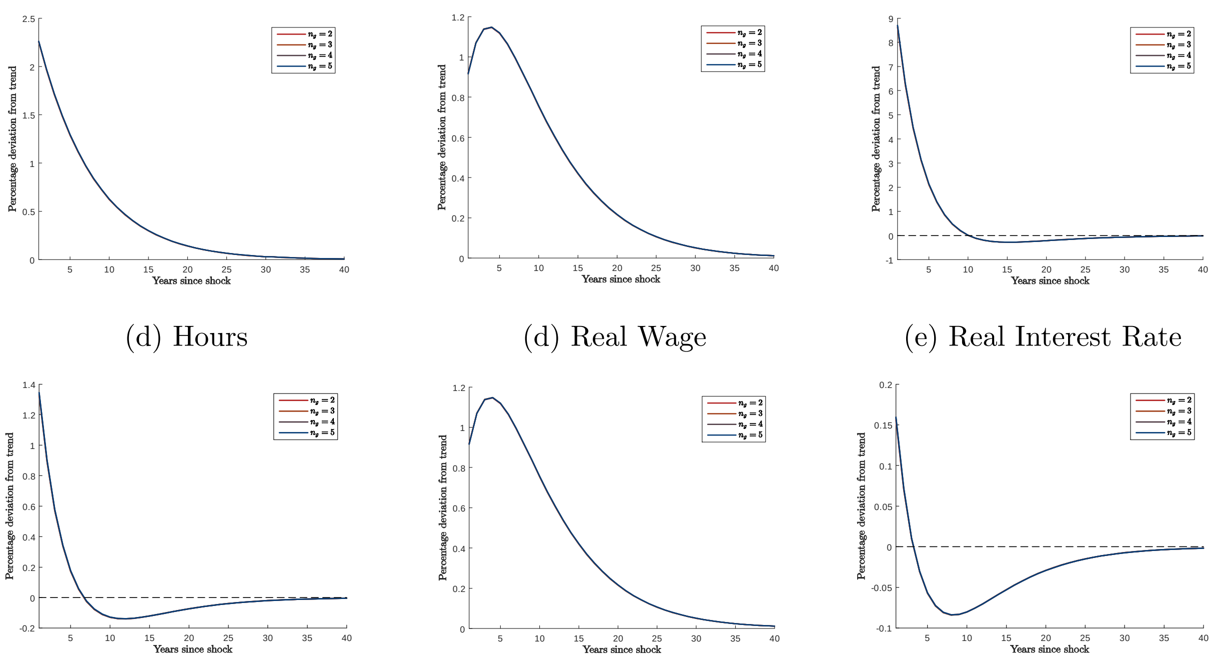
Task	$\mathbf{n}_g = 2$	$\mathbf{n}_g = 3$	$\mathbf{n}_g = 4$	$\mathbf{n}_g = 5$
Total time (in seconds)	53.04	72.89	101.14	136.01
Pre-processor	39%	48%	55%	58%
Stationary equilibrium	24%	19%	16%	15%
Perturbation	37%	33%	29%	27%

Notes: Computing time for computing a first order approximation of aggregate dynamics in seconds. "Pre-processor" refers to Dynare file processing, which parses the model file and symbolically differentiates the equilibrium conditions. "Stationary equilibrium" refers to computing the stationary equilibrium of the model with no aggregate shocks, as in Section 3.2. "Perturbation" refers to Dynare evaluating derivatives at the stationary equilibrium, and solving the linear system.

Figure 2: Impulse Responses of Aggregates, First Order Approximation

(a) Output (b) Consumption (c) Investment

(d) Hours (d) Real Wage (e) Real Interest Rate



Notes: Impulse respones of aggregate variables, for different degrees of approximation of the distribution. n_g refers to highest order moment used in parametric family (6).

Table 4: Business Cycle Statistics of Aggregates in First Order Approximation

SD (rel. to output)	$\mathbf{n}_g = 2$	$\mathbf{n}_g = 5$	**Corr. with Output**	$\mathbf{n}_g = 2$	$\mathbf{n}_g = 5$
Output	(2.13%)	(2.14%)	$\times$	$\times$	$\times$
Consumption	0.4695	0.4672	Consumption	0.9028	0.9013
Investment	3.8733	3.8925	Investment	0.9689	0.9687
Hours	0.6133	0.6121	Hours	0.9441	0.9444
Real wage	0.4695	0.4672	Real wage	0.9028	0.9013
Real interest rate	0.0845	0.0841	Real interest rate	0.7966	0.7978

Notes: Standard deviation of aggregate variables. All variables are HP-filtered with smoothing parameter $\lambda = 100$ and, with the exception of the real interest rate, have been logged. Standard deviations for variables other than output are expressed relative to that of output.

plots impulse response functions to an aggregate TFP shock, which completely characterize the dynamics in a first order approximation. An increase in aggregate TFP directly increases output, but also increases investment and labor demand, which further increases output but also factor prices. The resulting business cycle statistics are reported in Table 4. As usual in a real business cycle model, consumption is roughly half as volatile as output, investment is nearly four times as volatile, and labor is slightly less volatile. All variables are highly correlated with output because aggregate TFP is the only shock driving fluctuations in the model.

The aggregate dynamics are largely unaffected by the degree of approximation of the distribution. Visually, increasing the degree of approximation from $n_g = 2$ to $n_g = 5$ barely changes the impulse responses in Figure 2. Quantitatively, the business cycle statistics reported in Table 4 barely change as well. Hence, including high-degree approximation of the distribution would not significantly improve the accuracy of the algorithm for studying these aggregate dynamics.

Second Order Approximation In principle, the first order approximation of the model considered above could hide important nonlinearities in the aggregate dynamics. In a nonlinear model, impulse response functions depend on the size and sign of the shock, as well as the history of previous shocks. These nonlinearities could significantly alter the dynamics of the model. To investigate this possibility, in Appendix B I compute a second order approximation of the aggregate dynamics, which amounts to changing a `Dynare` option from `order=1` to `order=2`.[12] Quantitatively, the dynamics of the second order approximation closely resemble those of the first order approximation. This is consistent with Khan and Thomas (2008), who find little evidence of such nonlinearities using an alternative solution method.

Method Does Not Require Approximate Aggregation In this benchmark model, Khan and Thomas (2008) show that "approximate aggregation" holds, in the sense that the aggregate capital stock almost completely characterizes how the distribution influences aggregate dynamics. Following Krusell and Smith (1998), they solve the model by approximating the distribution with the aggregate capital stock, and find that their solution is extremely accurate. Hence, for this particular model, my method and the Krusell and Smith (1998) method are both viable. However, my method is significantly more efficient; in a comparison project, Terry (2015b) shows that my method solves and simulates the model in 0.098% the time of the Krusell and Smith (1998) method. This speed gain makes the full-information Bayesian estimation in Section 4 feasible.

Furthermore, because my method directly approximates the entire distribution, it can be applied to other models in which approximate aggregation fails. Appendix C makes this case concrete by adding investment-specific shocks to the model, and showing that for sufficiently volatile shocks the aggregate capital stock does not accurately approximate how the distribution affects dynamics. Extending Krusell and Smith (1998)'s algorithm would therefore require adding more moments to the forecasting rule. This quickly becomes infeasible, as each additional moment adds another state variable in a globally accurate solution method.

Hence, my method is not only significantly faster for models in which approximate aggregation holds, it applies equally well to models in which it fails.

4 Estimating Aggregate Shock Processes with Heterogeneous Firms

The goals of this section are to show that full-information Bayesian estimation of heterogeneous agent models is feasible using my method, and to illustrate how micro-level behavior can quan-

[12] In the interest of speed, for the second order approximation I use a slightly lower-order approximation of individual decision rules. As explained in the online user guide, this is because `Dynare` would otherwise run into issues with the size of the `Matlab` workspace. This can be overcome by instructing `Dynare` to instead use compiled `C++` code to compute the derivatives, which is relatively slow but still feasible. The slow `C++` compilation is during the "pre-processing" phase, which only must be performed once.

titatively impact estimation results. To do this, I extend the benchmark model to include aggregate investment-specific productivity shocks in addition to the neutral shocks and estimate the parameters of these two shock processes. As will become clear below, I include the additional investment-specific shock because this process is most directly shaped by micro-level investment behavior. The investment-specific shock only affects the capital accumulation equation, which becomes $k_{jt+1} = (1-\delta) k_{jt} + e^{q_t} i_{jt}$. The two aggregate shocks follow the joint process

$$\begin{aligned} z_t &= \rho_z z_{t-1} + \sigma_z \omega_t^z \\ q_t &= \rho_q q_{t-1} + \sigma_q \omega_t^q + \sigma_{qz} \omega_t^z, \end{aligned} \tag{13}$$

where ω_t^z and ω_t^q are i.i.d. standard normal random variables. I include a loading on neutral productivity innovations, σ_{qz}, to capture comovement between the two shocks. Without this loading factor, the investment-specific shocks would induce a counterfactual negative comovement between consumption and investment. Abusing notation introduced in Section 3, denote the vector of parameter values $\boldsymbol{\theta} = \left(\rho_z, \sigma_z, \rho_q, \sigma_q, \rho_{qz}\right)$.

I estimate the shock process parameters $\boldsymbol{\theta}$ conditional on four different parameterizations of the remaining parameters. The only parameters to vary across these parameterizations are $\overline{\xi}$, the upper bound on fixed cost draws, and σ_ε, the standard deviations of the innovations to idiosyncratic productivity. I vary the fixed costs from $\overline{\xi} = 0$, in which case the model exactly aggregates to a representative firm, to $\overline{\xi} = 1$; in the last case, I increase σ_ε from 0.02 to 0.04, because otherwise there would be little capital adjustment. These parameters vary the extent of micro-level adjustment frictions, and therefore micro-level investment behavior, from frictionless to extreme frictions. The remaining parameters are fixed at standard values, adjusting the model frequency to one quarter in order to match the frequency of the data. Table 5 collects all these parameter values.

The Bayesian approach combines a prior distribution of parameters, $p(\boldsymbol{\theta})$, with the likelihood function, $L(\mathbf{Y}|\boldsymbol{\theta})$ where $\mathbf{Y}$ is the observed time series of data, to form the posterior distribution of parameters, $p(\boldsymbol{\theta}|\mathbf{Y})$. The posterior is proportional to $p(\boldsymbol{\theta}) L(\mathbf{Y}|\boldsymbol{\theta})$, which is the object I characterize numerically. The data I use is $\mathbf{Y} = \left(\log \widehat{Y}_{1:T}, \log \widehat{I}_{1:T}\right)$, where $\log \widehat{Y}_{1:T}$ is the time series of log-linearly detrended real output and $\log \widehat{I}_{1:T}$ is log-linearly detrended real investment; for details, see Appendix D. I choose relatively standard prior distributions to form $p(\boldsymbol{\theta})$, also

Table 5: Parameterizations Considered in Estimation Exercises

Fixed Parameters	**Value**	**Fixed Parameters**	**Value**			
β (discount factor)	.99	a (no fixed cost)	.011			
σ (utility curvature)	1	ρ_ε (idioynscratic TFP)	.85			
α (inverse Frisch)	$\lim \alpha \to 0$	ν (capital share)	.21			
χ (labor disutility)	$N^* = \frac{1}{3}$	**Changing Parameters**	**Value 1**	**Value 2**	**Value 3**	**Value 4**
θ (labor share)	.64	σ_ε (idiosyncratic TFP)	.02	.02	.02	.04
δ (depreciation)	.025	$\bar{\xi}$	0	.01	.1	1

Notes: Calibrated parameters in the estimation exercises. "Fixed parameters" refer to those which are the same across the different estimations. "Changing parameters" are those which vary across estimations.

contained in Appendix D. I sample from $p(\boldsymbol{\theta})\, L(\mathbf{Y}|\boldsymbol{\theta})$ using the Metropolis-Hastings algorithm; since this procedure is now standard (see, for example, Fernández-Villaverde, Rubio-Ramírez, and Schorfheide (2015)), I omit further details. `Dynare` computes 10,000 draws from the posterior in 23 minutes, 57 seconds.[13]

Table 6 shows that as the upper bound of the fixed costs increases from $\bar{\xi} = 0$ to $\bar{\xi} = 1$, the estimated variance of investment-specific shocks significantly increases from 0.0058 to 0.0088. Intuitively, matching the aggregate investment data with large frictions requires more volatile shocks than with small frictions. Additionally, the factor loading of neutral shocks on investment-specific shocks shrinks, because larger frictions reduce the negative comovement of consumption and investment. The remaining parameters are broadly constant over the different specifications, indicating that micro-level adjustment frictions matter mainly for the inference of the investment-specific shock process.

Of course, the ideal estimation exercise would jointly estimate the adjustment frictions and shock processes using both micro and macro level data; these results show that, using my method, such exercises are now feasible. These exercises provide a potentially important step forward for formal inference in macroeconomics, which currently falls into two broad categories. The first is estimation of models with meaningful general equilibrium forces, as in the DSGE literature. To

[13] A key reason that this estimation is so efficient is that the parameters of the shock processes do not affect the stationary equilibrium of the model. Hence, the stationary equilibrium does not need to be recomputed at each point in the estimation process. Estimating parameters which affect the steady state would take longer, but is still feasible in `Dynare`.

Table 6: Posterior Distribution of Parameters, for Different Micro-Level Calibrations

Micro-Level	**TFP** ρ_z	**TFP** σ_z	**ISP** ρ_q	**ISP** σ_q	**ISP Loading** ρ_{zq}
$\bar{\xi} = 0$, $\sigma_\varepsilon = .02$	0.9811	0.0078	0.9718	0.0058	-0.00446
[90% HPD]	[0.9694, 0.9940]	[0.0072, 0.0083]	[0.954, 0.9912]	[0.0045, 0.0071]	[-0.0060, -0.0031]
$\bar{\xi} = .01$, $\sigma_\varepsilon = .02$	0.9808	0.0078	0.9747	0.0071	-0.0043
	[0.9687, 0.9924]	[0.0072, 0.0084]	[0.9591, 0.9901]	[0.0056, 0.0088]	[-0.0059, -0.0026]
$\bar{\xi} = .1$, $\sigma_\varepsilon = .02$	0.9796	0.0078	0.9732	0.0075	-0.0040
	[0.9670, 0.9922]	[0.0073, 0.0084]	[0.9581, 0.9902]	[0.0059, 0.0094]	[-0.0056, -0.0025]
$\bar{\xi} = 1$, $\sigma_\varepsilon = .04$	0.9786	0.0079	0.9730	0.0088	-0.0037
	[0.9659, 0.9913]	[0.0073, 0.0085]	[0.9549, 0.9924]	[0.0066,0.0111]	[-0.0054, -0.0020]

Notes: Posterior means and highest posterior density sets of parameters, conditional on micro-level parameterizations.

maintain tractability, these exercises generally rely upon (nearly) representative agent assumptions and ignore micro heterogeneity. The second category is estimation of models which focus on micro heterogeneity, but ignore meaningful general equilibrium, as in labor economics or industrial organization. Some recent work (such as Vavra (2014) or Bloom et al. (2014)) bridges this gap by estimating models with both micro heterogeneity and meaningful general equilibrium forces. Because of severe computational burden, these exercises use partial information, moment-based econometrics. My new solution method, and its `Dynare` implementation, instead bridges these two literatures in a tractable fashion, overcoming the extreme runtimes and restriction to partial information procedures in previous work.

5 Conclusion

In this note, I developed a general-purpose method for solving and estimating heterogeneous agent macro models. In contrast to most existing work, my method does not rely on the dynamics of the distribution being well-approximated by a small number of moments, substantially expanding the class of models which can be feasibly computed. Nevertheless, my method is straightforward to implement. I have provided codes and a user guide for solving a general class of models using `Dynare`, with the hope that it will help bring heterogeneous agent models into the fold of standard macroeconomic analysis. A particularly promising avenue for future research is incorporating micro data into the estimation of DSGE models. As I showed in Section 4, micro-level behavior places important restrictions on model parameters. In the current DSGE literature, such restrictions are either absent or imposed through ad-hoc prior beliefs; my method instead allows for the micro data to formally place these restrictions itself.

References

[1] Adjemian, Stéphane, Houtan Bastani, Michel Julliard, Frédéric Karamé, Ferhat Mihoubi, George Perendia, Johannes Pfeifer, Marco Ratto, and Sébastien Villemot (2011), "Dynare: Reference Manual, Version 4," *Dynare Working Papers*, 1, CEPREMAP.

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

[3] Andreasen, Martin, Jesús Fernández-Villaverde, and Juan Rubio-Ramírez (2014), "The Pruned State-Space System for Non-Linear DSGE Models: Theory and Applications," *Unpublished manuscript.*

[4] Auclert, Adrien (2015), "Monetary Policy and the Redistribution Channel," *Unpublished manuscript.*

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

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

[7] Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen Terry (2014), "Really Uncertain Business Cycles," *Unpublished manuscript.*

[8] Campbell, Jeffrey (1998), "Entry, Exist, Embodied Technology, and Business Cycles," *Review of Economic Dynamics* 1(2) 371-408.

[9] Clementi, Gian Luca and Beradino Palazzo (2015), "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations," *American Economic Journals: Macroeconomics (forthcoming).*

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

[11] Den Haan (2010), "Assessing the Accuracy of the Aggregate Law of Motion in Models with Heterogeneous Agents," *Journal of Economic Dynamics and Control*, 34(1), 79-99.

[12] Fernández-Villaverde, Jesús, Juan Rubio-Ramírez, and Frank Schorfheide (2015), "Solution and Estimation Methods for DSGE Models," *Unpblished manuscript,* prepared for Handbook of Macroeconomics, Volume 2.

[13] Hopenhayn, Hugo, and Richard Rogerson (1993), "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 915-938.

[14] Kaplan, Greg, Benjamin Moll, and Gianluca Violante (2015), "Monetary Policy According to HANK," *Princeton University Working Paper.*

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

[16] Khan, Aubhik and Julia Thomas (2013), "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*, 121(6), 1055-1107.

[17] Koop, Gary, M. Hashem Pesaran, and Simon Potter (1996), "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics*, 74(1), 119-147.

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

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

[20] Schmitt-Grohé, Stephanie, and Martín Uribe (2004), "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function," *Journal of Economic Dynamics and Control*, 28(4), 755-775.

[21] Stephen Terry (2015a), "The Macro Impact of Short-Termism," *Unpublished manuscript*.

[22] Stephen Terry (2015b), "Alternative Methods for Solving Heterogeneous Firm Models," *Unpublished manuscript*.

[23] Vavra, Joseph (2014), "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation," *The Quarterly Journal of Economics*, 129(1), 215-258.

[24] Veracierto, Marcelo (2016), "Adverse Selection, Risk Sharing and Business Cycles," *Federal Reserve Bank of Chicago Working Paper*.

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

A Details of the Method

This appendix provides additional details of the method referenced in Section 3 in the main text.

A.1 Approximate Equilibrium Conditions

I first show that the approximate equilibrium conditions can be written as a system of $2n_\varepsilon n_k + n_g + 2 + n_g + 1$ equations, of the form (11). To that end, let $\{\tau_i^g, (\varepsilon_i, k_i)\}_{i=1}^{m_g}$ denote the weights and nodes of the two-dimensional Gauss-Legendre quadrature used to approximate the integrals with respect to the distribution, and let $\{\tau_i^\varepsilon, \omega_i^\varepsilon\}_{i=1}^{m_\varepsilon}$ denote the weights and nodes of the one-dimensional Gauss-Hermite quadrature used to approximate the integrals with respect to the idiosyncratic shock innovations. In my numerical implementation, I use the degree of approximation for value functions $n_\varepsilon = 3$ and $n_k = 5$, for the Gauss-Legendre quadrature $m_g = 64$ (from the tensor product of two 8^{th} order, one dimensional Gauss-Legendre quadrature nodes and weights), and $m_\varepsilon = 3$ for the degree of the Gauss-Hermite quadrature.

With this notation, and the notation defined in the main text, the approximate Bellman equation (10) can be written as

$$0 = E[\sum_{k=1}^{n_\varepsilon} \sum_{l=1}^{n_k} \theta_{ij} T_k(\varepsilon_i) T_l(k_j) - \lambda \left(e^z e^{\varepsilon_i} k_j^\theta n(\varepsilon_i, k_j)^\nu - wn(\varepsilon_i, k_j) \right) - \lambda (1-\delta) k_j \qquad (14)$$

$$-\left(\frac{\widehat{\xi}\left(\varepsilon_i,k_j\right)}{\overline{\xi}}\right)\left(\begin{array}{c}-\lambda\left(k^a\left(\varepsilon_i,k_j\right)-w\frac{\widehat{\xi}\left(\varepsilon_i,k_j\right)}{2}\right)\\ \\ +\beta\sum_{o=1}^{m_\varepsilon}\tau_o^\varepsilon\sum_{k=1}^{n_\varepsilon}\sum_{l=1}^{n_k}\theta_{ij}'T_k\left(\rho_\varepsilon\varepsilon_i+\sigma_\varepsilon\omega_o^\varepsilon\right)T_l\left(k^a\left(\varepsilon_i,k_j\right)\right)\end{array}\right)$$

$$-\left(1-\frac{\widehat{\xi}\left(\varepsilon_i,k_j\right)}{\overline{\xi}}\right)\left(\begin{array}{c}-\lambda k^n\left(\varepsilon_i,k_j\right)\\ \\ +\beta\sum_{o=1}^{m_\varepsilon}\tau_o^\varepsilon\sum_{k=1}^{n_\varepsilon}\sum_{l=1}^{n_k}\theta_{ij}'T_k\left(\rho_\varepsilon\varepsilon_i+\sigma_\varepsilon\omega_o^\varepsilon\right)T_l\left(k^n\left(\varepsilon_i,k_j\right)\right)\end{array}\right)],$$

for the $n_\varepsilon n_k$ collocation nodes $i=1,...,n_\varepsilon$ and $j=1,...n_k$. The optimal labor choice is defined through the first order condition

$$n\left(\varepsilon_i,k_j\right)=\left(\frac{\nu e^z e^{\varepsilon_i}k_j^\theta}{w}\right)^{\frac{1}{1-\nu}}.$$

The policy functions $k^a\left(\varepsilon_i,k_j\right)$, $k^n\left(\varepsilon_i,k_j\right)$, and $\widehat{\xi}\left(\varepsilon_i,k_j\right)$ are derived directly from the approximate value function $\boldsymbol{\theta}$ as follows. First, the adjust capital decision rule $k^a\left(\varepsilon_i,k_j\right)$ must satisfy the first order condition

$$0=E\left[\lambda-\beta\sum_{o=1}^{m_\varepsilon}\tau_o^\varepsilon\sum_{k=1}^{n_\varepsilon}\sum_{l=1}^{n_k}\theta_{ij}'T_k\left(\rho_\varepsilon\varepsilon_i+\sigma_\varepsilon\omega_o^\varepsilon\right)T_l'\left(k^a\left(\varepsilon_i,k_j\right)\right)\right]. \tag{15}$$

Conditional on this choice, the constrained capital decision is

$$k^n\left(\varepsilon_i,k_j\right)=\left\{\begin{array}{c}\left(1-\delta+a\right)k_j \text{ if } k^a\left(\varepsilon_i,k_j\right)>\left(1-\delta+a\right)k_j\\ \\ k^a\left(\varepsilon_i,k_j\right) \text{ if } k^a\left(\varepsilon_i,k_j\right)\in\left[\left(1-\delta-a\right)k_j,\left(1-\delta+a\right)k_j\right]\\ \\ \left(1-\delta-a\right)k_j \text{ if } k^a\left(\varepsilon_i,k_j\right)<\left(1-\delta-a\right)k_j\end{array}\right\}.$$

Finally, the capital adjustment threshold $\widetilde{\xi}\left(\varepsilon_i,k_j\right)$ is defined as

$$\widetilde{\xi}\left(\varepsilon_i,k_j\right)=\frac{1}{w\lambda}\left[\begin{array}{c}-\lambda\left(k^a\left(\varepsilon_i,k_j\right)-k^n\left(\varepsilon_i,k_j\right)\right)\\ \\ +\beta\sum_{o=1}^{m_\varepsilon}\tau_o^\varepsilon\sum_{k=1}^{n_\varepsilon}\sum_{l=1}^{n_k}\theta_{ij}'T_k\left(\rho_\varepsilon\varepsilon_i+\sigma_\varepsilon\omega_o^\varepsilon\right)\left(T_l\left(k^a\left(\varepsilon_i,k_j\right)\right)-T_l\left(k^n\left(\varepsilon_i,k_j\right)\right)\right)\end{array}\right],$$

and the bounded threshold is given by $\widehat{\xi}\left(\varepsilon_i, k_j\right) = \min\{\max\{0, \widetilde{\xi}\left(\varepsilon_i, k_j\right), \bar{\xi}\}\}$. To evaluate the decision rules off the grid, I interpolate the adjust capital decision rule k^a using Chebyshev polynomials, and derive k^n and $\widehat{\xi}$ from the above formulae.

Given the firm decision rules, the implications of household optimization can be written as

$$0 = \lambda - \left(\sum_{i=1}^{m_g} \tau_i^g \begin{pmatrix} e^z e^{\varepsilon_i} k_i^\theta n\left(\varepsilon_i, k_i\right) + \left(1-\delta\right) k_i \\ -\frac{\widehat{\xi}\left(\varepsilon_i, k_i\right)}{\bar{\xi}} k^a\left(\varepsilon_i, k_i\right) - \left(1 - \frac{\widehat{\xi}\left(\varepsilon_i, k_i\right)}{\bar{\xi}}\right) k^n\left(\varepsilon_i, k_i\right) \end{pmatrix} g\left(\varepsilon_i, k_i | \mathbf{m}\right) \right)^{-\sigma} \tag{16}$$

$$0 = \left(\frac{w\lambda}{\chi}\right)^{\frac{1}{\alpha}} - \sum_{i=1}^{m_g} \tau_i^g \left(n\left(\varepsilon_i, k_i\right) + \frac{\widehat{\xi}\left(\varepsilon_i, k_i\right)^2}{2\bar{\xi}} \right) g\left(\varepsilon_i, k_i | \mathbf{m}\right), \tag{17}$$

where

$$\begin{aligned} g\left(\varepsilon_l, k_l | \mathbf{m}\right) &= g_0 \exp\{ g_1^1 \left(\varepsilon_l - m_1^1\right) + g_1^2 \left(k_l - m_1^2\right) + \\ &\quad \sum_{i=2}^{n_g} \prod_{j=0}^{i} g_i^j \left[\left(\varepsilon_l - m_1^1\right)^{i-j} \left(k_l - m_1^2\right)^j - m_i^j \right] \}. \end{aligned}$$

The approximate law of motion for the distribution (9) can be written

$$\begin{aligned} 0 &= m_1^{1\prime} - \sum_{l=1}^{m_g} \tau_l^g \sum_{k=1}^{m_\varepsilon} \tau_k^\varepsilon \left(\rho_\varepsilon \varepsilon_l + \sigma_\varepsilon \omega_k^\varepsilon\right) g\left(\varepsilon_l, k_l | \mathbf{m}\right) \qquad (18) \\ 0 &= m_1^{2\prime} - \sum_{l=1}^{m_g} \tau_l^g \sum_{k=1}^{m_\varepsilon} \tau_k^\varepsilon \left[\frac{\widehat{\xi}\left(\varepsilon_l, k_l\right)}{\bar{\xi}} \left(k^a\left(\varepsilon_l, k_l\right) - m_1^{2\prime}\right) + \left(1 - \frac{\widehat{\xi}\left(\varepsilon_l, k_l\right)}{\bar{\xi}}\right) \left(k^a\left(\varepsilon_l, k_l\right) - m_1^{2\prime}\right) \right] g\left(\varepsilon_l, k_l | \mathbf{m}\right) \\ 0 &= m_i^{j\prime} - \sum_{l=1}^{m_g} \tau_l^g \sum_{k=1}^{m_\varepsilon} \tau_k^\varepsilon \begin{bmatrix} \frac{\widehat{\xi}\left(\varepsilon_l, k_l\right)}{\bar{\xi}} \left(\left(\rho_\varepsilon \varepsilon_l + \omega_k^\varepsilon - m_1^{1\prime}\right)^{i-j} \left(k^a\left(\varepsilon_l, k_l\right) - m_1^{2\prime}\right)^j \right) \\ + \left(1 - \frac{\widehat{\xi}\left(\varepsilon_l, k_l\right)}{\bar{\xi}}\right) \left(\left(\rho_\varepsilon \varepsilon_l + \omega_k^\varepsilon - m_1^{1\prime}\right)^{i-j} \left(k^a\left(\varepsilon_l, k_l\right) - m_1^{2\prime}\right)^j \right) \end{bmatrix} g\left(\varepsilon_l, k_l | \mathbf{m}\right). \end{aligned}$$

Consistency between the moments $\mathbf{m}$ and parameters $\mathbf{g}$ requires

$$
\begin{aligned}
m_1^1 &= \sum_{l=1}^{m_g} \tau_l^g \varepsilon_l g\left(\varepsilon_l, k_l | \mathbf{m}\right), \\
m_1^2 &= \sum_{l=1}^{m_g} \tau_l^g k_l g\left(\varepsilon_l, k_l | \mathbf{m}\right), \text{ and} \\
m_i^j &= \sum_{l=1}^{m_g} \tau_l^g \left[\left(\varepsilon_l - m_1^1\right)^{i-j} \left(k_l - m_1^2\right)^j - m_i^j\right] g\left(\varepsilon_l, k_l | \mathbf{m}\right) \text{ for } i = 2, ..., n_g,\ j = 0, ..., i.
\end{aligned}
\tag{19}
$$

Finally, the law of motion for the aggregate productivity shock is

$$
0 = E\left[z' - \rho_z z\right] \tag{20}
$$

With all these expressions, we can define $f\left(\mathbf{y}', \mathbf{y}, \mathbf{x}', \mathbf{x}; \psi\right)$ which outputs (14), (15), (16), (17), (18), (19), and (20).

A.2 Solving for Stationary Equilibrium

Following Hopenhayn and Rogerson (1993), I solve for the stationary equilibrium by iterating on the wage w^*:

1. Guess a value for the wage w^*

2. Given w^*, compute the firm's value function $\boldsymbol{\theta}^*$ by iterating on the Bellman equation (10). Note that λ^* does not enter the stationary Bellman equation because it is a multiplicative constant.

3. Using the firm's decision rules, compute the invariant distribution $\mathbf{m}^*$ by iterating on the law of motion (9).

4. Compute aggregate labor demand using this invariant distribution, and compute aggregate labor supply using the household's first order condition.

5. Update the guess of w^* appropriately.[14]

[14] Although I describe this as an iteration, it is actually more efficient numerically to view this as a root-finding problem, solving for the wage which sets excess labor demand to 0.

Table 7: Business Cycle Statistics, First vs. Second Order Approximation

SD (rel. to output)	**Order 1**	**Order 2**	**Corr. with Output**	**Order 1**	**Order 2**
Output	(2.11%)	(2.11%)	×	×	×
Consumption	0.4739	0.4739	Consumption	0.9078	0.9078
Investment	3.7299	3.7299	Investment	0.9698	0.9698
Hours	0.6066	0.6066	Hours	0.9445	0.9445
Real wage	0.4739	0.4739	Real wage	0.9078	0.9078
Real interest rate	0.0806	0.0806	Real interest rate	0.7864	0.7864

Notes: All variables are HP-filtered with smoothing parameter $\lambda = 100$ and, with the exception of the real interest rate, have been logged. Standard deviations for variables other than output are expressed relative to that of output.

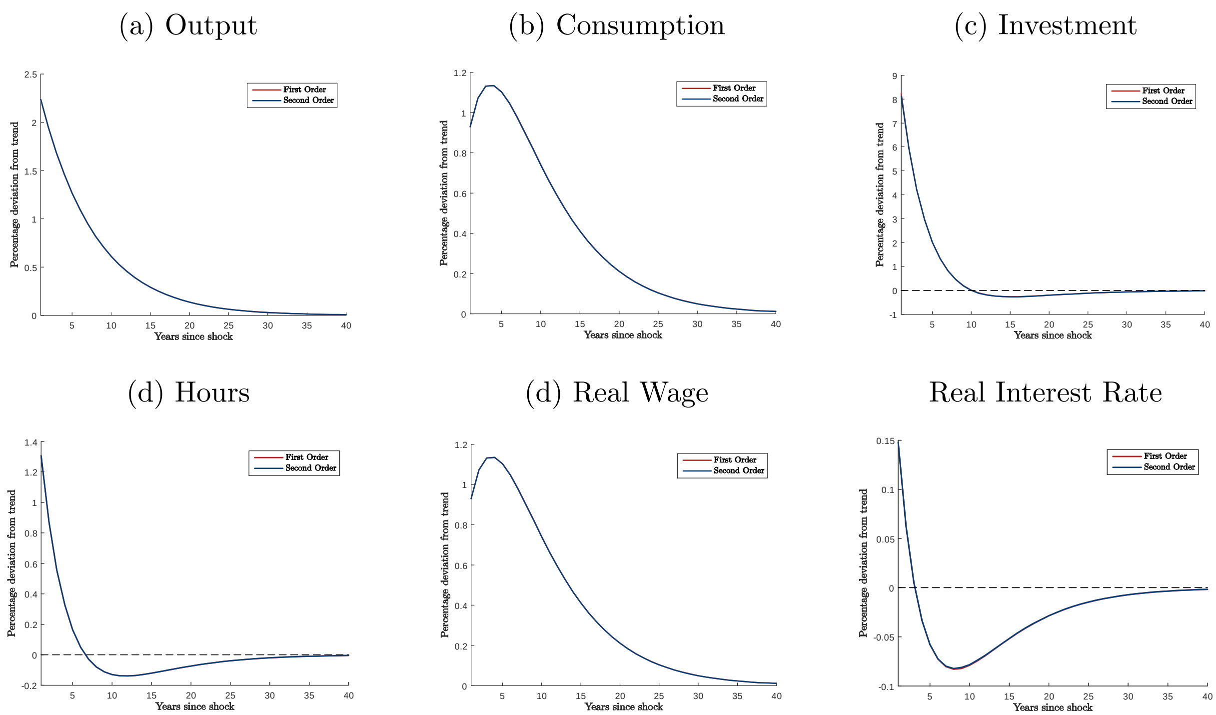
For the "exact" histogram comparisons in the main text, I follow these steps, except I approximate the distribution with a histogram recording the mass of firms along a fine grid following Young (2010).

B Second Order Approximation of Aggregate Dynamics

In this appendix, I document properties of the second order approximation of aggregate dynamics referenced in Section 3.3 of the main text. Figure 3 plots the impulse response to a positive, one standard deviation aggregate TFP shock starting from the stationary equilibrium, in a first order and a second order approximation of the model. The two responses are nearly indistinguishable. In fact, the resulting business cycle statistics reported in Table 7 are quantitatively identical up to four decimal places.

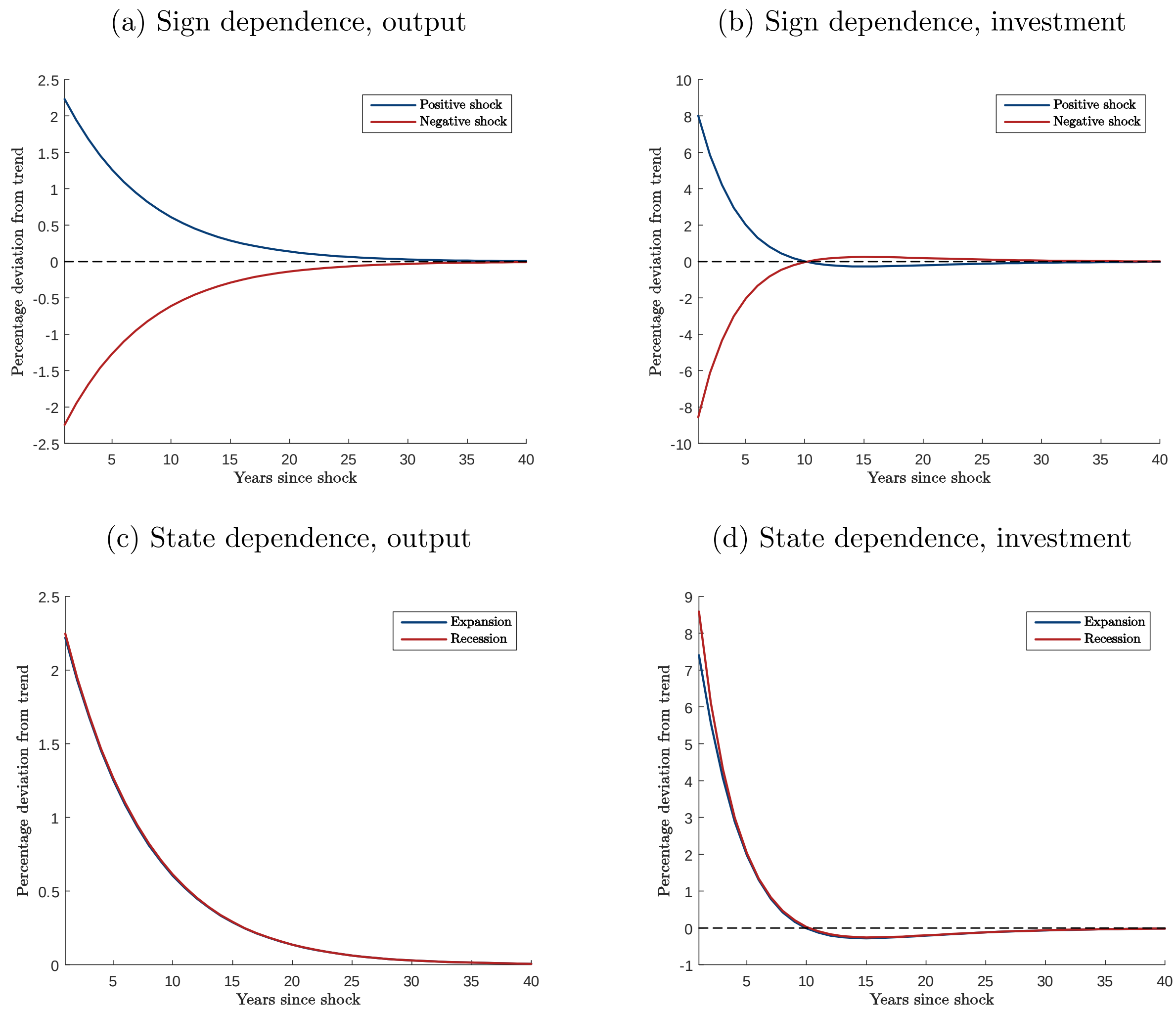
To further investigate the potential for nonlinearities, the top row of Figure 4 plots the response to a positive vs. a negative shock in the second order expansion, and finds that the absolute responses are almost the same. The bottom row of Figure 4 plots the response to a one standard deviation positive shock, starting from a "recession" (negative one standard deviation shocks in the previous two periods) compared to an "expansion" (positive shocks in the previous periods). Although there is a slight history dependence in the response of investment, it is quantitatively small.

Figure 3: Aggregate Impulse Responses, First vs. Second Order Approximation



Notes: Impulse respones of aggregate variables, for different orders of approximation. "First order" refers to linear impulse response. "Second order" refers to nonlinear generalized impulse response function, as in Koop et al. (1996).

Figure 4: Sign and Size Dependence in Impulse Responses, Second Order Approximation



Notes: Nonlinear features of the impulse responses of aggregate output (left column) and investment (right column). "Sign dependence" refers to the impulse response to a one standard deviation positive vs. negative shock. "State dependence" refers to the impulse response after positive one standard deviation shocks vs. negative one standard deviation shocks in the previous two years. All impulse responses computed nonlinearing as in Koop et al. (1996).

Table 8: Standard Deviation of Aggregates in First Order Approximation

Forecasting Equation	$\mathbf{R}^2$	**RMSE**
Marginal Utility	0.998185	0.00145
Future capital	0.999855	0.00061

Notes: Results from running forecasting regressions (21) on data simulated from first order model solution.

C Method Does Not Require Approximate Aggregation

In this appendix, I show that my method continues to perform well even when approximate aggregation fails to hold. To do this, I modify the benchmark model, because as Khan and Thomas (2008) show approximate aggregation holds in this case. In the benchmark model, the distribution impacts firms' decisions through two channels: first, by determining the marginal utility of consumption $\lambda(z,\mu)$, and second, by determining the law of motion of the distribution, $\mu'(z,\mu)$.[15] Table 8 shows that the aggregate capital stock K_t captures both of these channels very well, by estimating the forecasting equations

$$\begin{aligned}\log \lambda_t &= \alpha_0 + \alpha_1 z_t + \alpha_2 \log K_t \\ \log K_{t+1} &= \gamma_0 + \gamma_1 z_t + \gamma_2 \log K_t\end{aligned} \tag{21}$$

on data simulated using my solution. The R^2 of these forecasting equations are high, and the root mean-squared error low, indicating that a Krusell and Smith (1998) algorithm using the aggregate capital stock performs well in this environment.

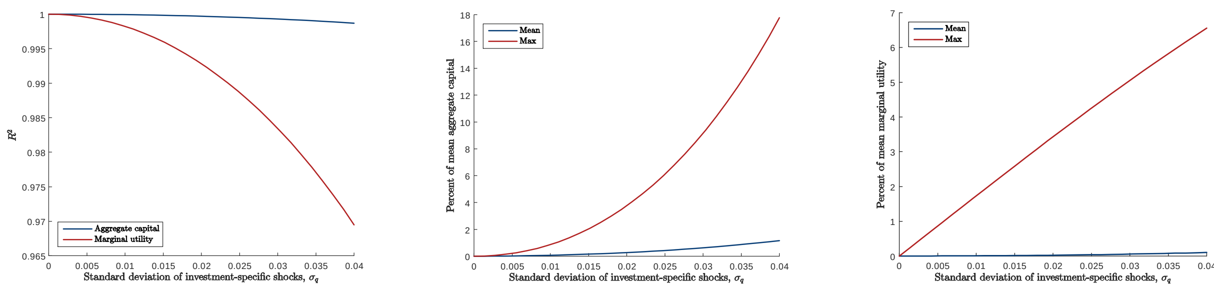
To break this approximate aggregation result, I add an investment-specific productivity shock q_t to the benchmark model. In this case, the capital accumulation equation becomes $k_{jt+1} = (1-\delta)k_{jt} + e^{q_t} i_{jt}$, but the remaining equations are unchanged. I assume the investment-specific shock follows the AR(1) process $q_t = \rho_q q_{t-1} + \sigma_q \omega_t^q$, where $\omega_t^q \sim N(0,1)$, independently of the aggregate TFP shock.

Figure 5 shows that approximate aggregation becomes weaker as the investment-specific pro-

[15] Given the linear disuility of labor supply, the wage is purely a function of the marginal utility of consumption.

Figure 5: Forecasting Power of Aggregate Capital, as a Function of Investment-Specific Shock Variance

(a) R^2 of forecasting regs (b) DH statistic, capital (c) DH statistic, mar. utility



Notes: Results from running forecasting regressions (22) on data simulated from first order model solution. "DH statistic" refers to Den Haan (2010)'s suggestion of iterating on forecasting equations without updating K_t from simulated data.

ductivity shock becomes more important. Panel (a) plots the R^2s from the forecasting equations

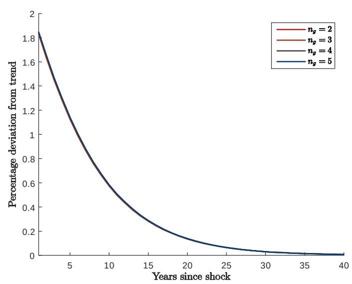
$$\begin{aligned} \log \lambda_t &= \alpha_0 + \alpha_1 z_t + \alpha_2 q_t + \alpha_3 \log K_t \\ \log K_{t+1} &= \gamma_0 + \gamma_1 z_t + \gamma_2 q_t + \gamma_3 \log K_t \end{aligned} \tag{22}$$

as a function of the shock volatility σ_q, keeping $\rho_q = 0.859$ throughout. However, as Den Haan (2010) notes, the R^2 is a loose error metric for two reasons: first, it only measures one period ahead forecasts, whereas agents must forecast into the infinite future; and second, it only measures average deviations, which potentially hide occasionally large errors. To address these concerns, Den Haan (2010) proposes iterating on the forecasting equations (22) without updating the capital stock, and computing both average and maximum deviations of these forecasts from the actual values in a simulation. Panels (b) and (c) of Figure 5 shows that these more stringent metrics grow even more sharply as a function of the volatility σ_q. Hence, Krusell and Smith (1998) algorithms which approximate the distribution with only the aggregate capital stock will fail in these cases.

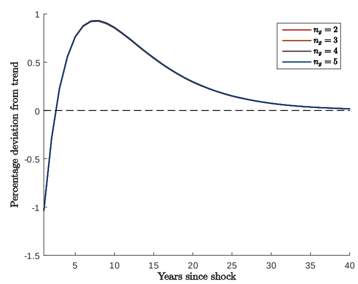
Because my method directly approximates the distribution, rather than relying on these low-dimensional forecasting rules, it continues to perform well as investment-specific shocks become more important. Figure 6 plots the impulse responses of key aggregate variables to an investment-specific shock for $\sigma_q = 0.02$, a value for which approximate aggregation fails. As with neutral

Figure 6: Aggregate Impulse Responses to Investment-Specific Productivity Shock

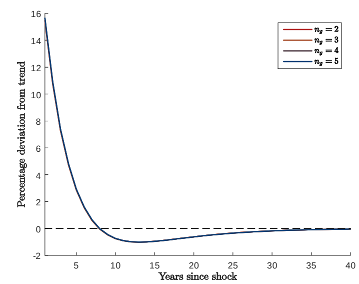
(a) Output



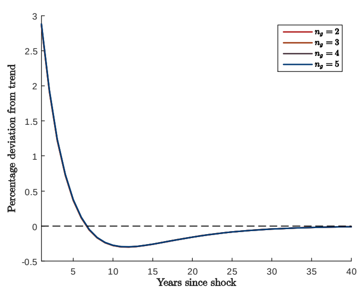
(b) Consumption



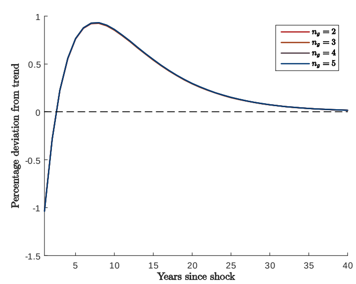
(c) Investment



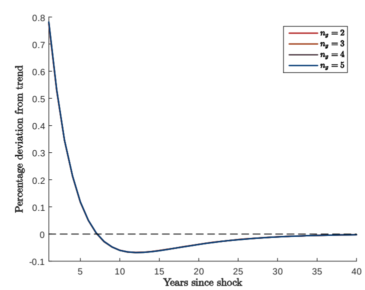
(d) Hours



(d) Real Wage



(e) Real Interest Rate



Notes: Impulse respones of aggregate variables, for different orders of approximation of the distribution. n_g refers to highest order moment used in parametric family (6).

Table 9: Prior Distributions of Parameters

Parameter	**Role**	**Prior Distribution**
ρ_z	TFP autocorrelation	Beta (0.9, 0.07)
σ_z	TFP innovation sd	Inverse Gamma (0.01, 1)
ρ_q	ISP autocorrelation	Beta (0.9, 0.07)
σ_q	ISP innovation sd	Inverse Gamma (0.01, 1)
σ_{qz}	ISP loading on TFP innovation	Uniform (-0.05, 0.05)

Notes: Prior distributions for estimated parameters.

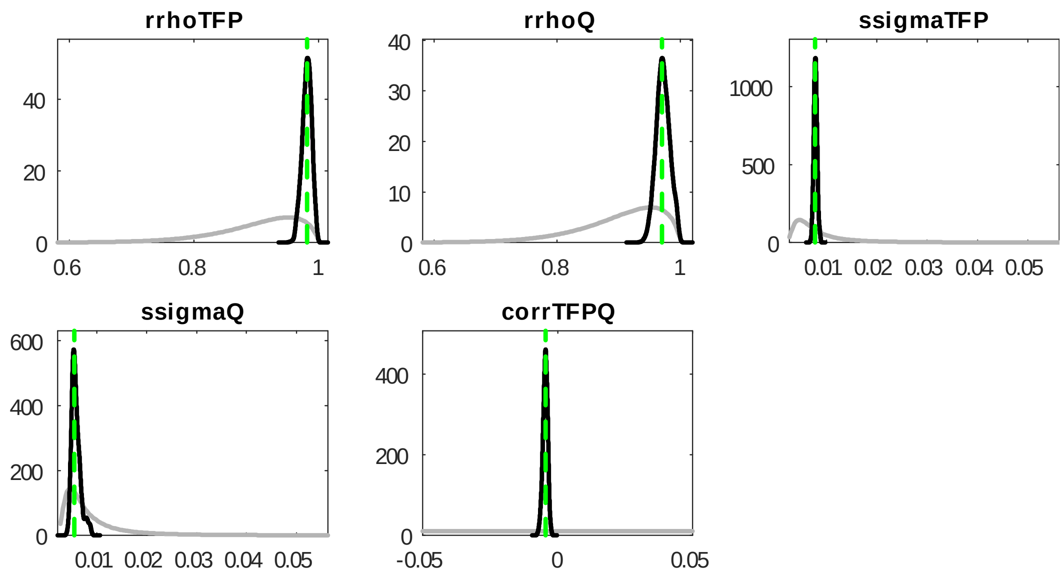
productivity shocks in Figure 2, even relatively low degree approximations of the distribution are sufficient to capture the dynamics of these variables. However, there is a slightly greater difference between $n_g = 2$ and $n_g = 3$ degree approximations, indicating that the shape of the distribution varies more in response to the investment-specific shocks.

D Estimation Details

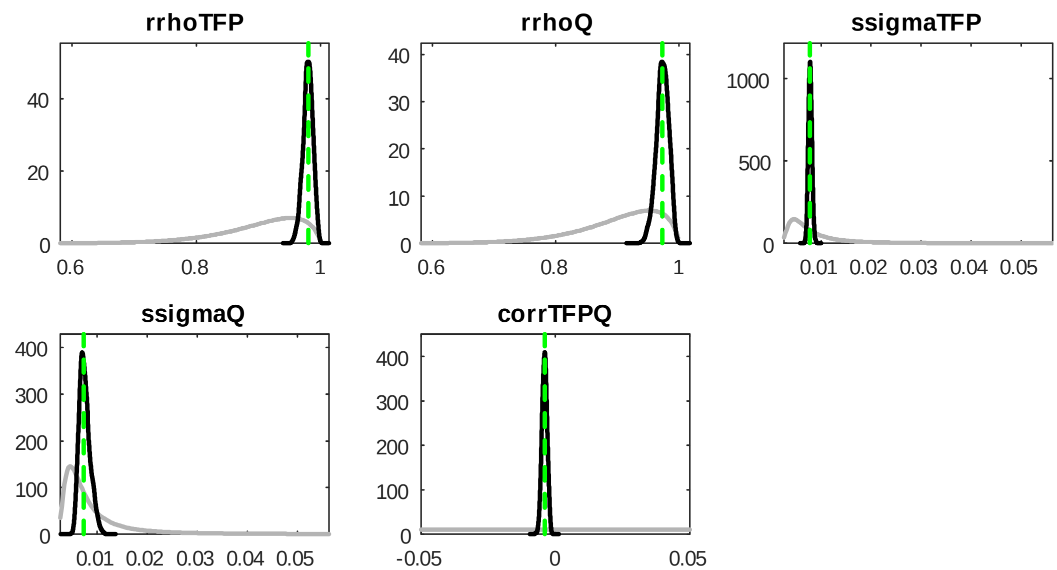
In this appendix, I provide additional details of the estimation exercise described in Section 4 of the main text. The particular data sets I use are (1) Real Gross Private Domestic Investment, 3 Decimal (series ID: `GPDIC96`), quarterly 1954-01-01 to 2015-07-01, and (2) Real Personal Consumption Expenditures: Nondurable Goods (chain-type quantity index) (series ID: `DNDGRA3Q086SBEA`), seasonally adjusted, quarterly 1954-01-01 to 2015-07-01. I log-linearly detrend both series and match them to log-deviations from stationary equilibrium in the model. The prior distributions of parameters are independent of each other, and given in Table 9. To sample from the posterior distribution, I use Markov Chain Monte Carlo with 10,000 draws, and drop the first 5,000 draws as burn-in. Figure 7 plots the prior and estimated posterior distributions of parameters under two micro-level calibrations. Increasing the capital adjustment frictions from Panel (a) to Panel (b), the posterior distribution of σ_q is shifted rightward and is slightly more dispersed.

Figure 7: Estimated Distribution of Investment-Specific Shock Variance

(a) Representative Firm ($\bar{\xi} = 0$, $\sigma_\varepsilon = 0.02$)



(b) High Frictions ($\bar{\xi} = 1$, $\sigma_\varepsilon = 0.04$)



Notes: Estimation results for different micro-level parameterizations. Grey lines are the prior distribution of parameters. Dashed greens lines are the posterior mode. Black lines are the posterior distribution.