Quantitative Economics 9 (2018), 1123–1151 1759-7331/20181123

A method for solving and estimating heterogeneous agent macro models

Thomas Winberry
Booth School of Business, University of Chicago and NBER

I develop a computational method for solving and estimating heterogeneous agent macro models with aggregate shocks. The main challenge is that the aggregate state vector contains the distribution of agents, which is typically infinite-dimensional. I approximate the distribution with a flexible parametric family, reducing its dimensionality to a finite set of endogenous parameters, and solve for the dynamics of these endogenous parameters by perturbation. I implement the method in `Dynare` and show that it is fast, general, and easy to use. As an illustration, I use the method to perform a Bayesian estimation of a heterogeneous firm model with aggregate shocks to neutral and investment-specific productivity. I find that the behavior of investment at the firm level quantitatively shapes inference about the aggregate shock processes, suggesting an important role for micro data in estimating DSGE models.

Keywords. Heterogeneous agents, computational economics, estimation, lumpy investment.

JEL classification. C63, E22, E32.

1. Introduction

Heterogeneity is pervasive in microeconomic data; households vary tremendously in their income, wealth, and consumption, for example, while firms vary in their productivity, investment, and hiring. A rapidly growing literature has emerged which studies how this micro heterogeneity shapes our understanding of business cycle fluctuations.[1] The heterogeneous agent models studied in this literature are computationally challenging because the aggregate state vector contains the distribution of microeconomic agents,

Thomas Winberry: Thomas.Winberry@chicagobooth.edu
Previously circulated as "A Toolbox for Solving and Estimating Heterogeneous Agent Macro Models." This paper elaborates on an algorithm developed in my dissertation. I thank Jung Sakong and Yulia Zhestkova for excellent research assistance. I thank Mark Aguiar, Jesus Fernandez-Villaverde, Alejandro Justiniano, Greg Kaplan, Ezra Oberfield, Richard Rogerson, Esteban Rossi-Hansberg, Christian Vom Lehn, Michael Weber, and especially Stephen Terry for useful comments. I thank seminar participants at the NBER 2016 Summer Institute Workshop on Methods and Applications for Dynamic Equilibrium Models, the SITE 2016 workshop on Computational Methods for Dynamic Economies and Games, and the Bank of Canada. `Dynare` code and a user guide are available in the Supplemental Material (Winberry (2018)).

[1]There are too many papers to provide a comprehensive list of citations. For recent papers on the household side, see Auclert (2017), Berger and Vavra (2015), or Kaplan, Moll, and Violante (2018); on the firm side, see Bachmann, Caballero, and Engel (2013), Khan and Thomas (2013), Clementi and Palazzo (2016), Terry (2017b), or Ottonello and Winberry (2017).

Available at http://qeconomics.org. https://doi.org/10.3982/QE740

which is typically an infinite-dimensional object. Although many numerical algorithms have been developed to overcome this challenge, none are as general, efficient, or easy-to-use as the standard perturbation methods routinely employed to solve representative agent models using prepackaged toolboxes like `Dynare`. Due to this challenge, heterogeneous agent models have yet to reach widespread adoption, particularly among central banks and policy institutions.

In this paper, I develop a general, efficient, and easy-to-use computational method for solving heterogeneous agent models with aggregate shocks. I approximate the infinite-dimensional distribution with a finite-dimensional parametric family, so that the parameters of that family become endogenous state variables. An accurate approximation of the distribution may require a large number of parameters, so I solve for the aggregate dynamics of the model using locally accurate perturbation techniques (which are computationally efficient even when the aggregate state vector is large). In order to make the method accessible, I implement the perturbation step in `Dynare` and provide a detailed online code template which explains how to do so (see Winberry (2018)). Finally, to illustrate the strength of the method, I use it to estimate a heterogeneous firm model with full-information Bayesian techniques. The degree of frictions at the micro level has a quantitatively significant impact on the estimated aggregate shock processes, showing that micro behavior places important restrictions on the estimation of DSGE macro models.

Although the method is applicable to a wide range of models, for concreteness I demonstrate its use in the Khan and Thomas (2008) model, which extends the real business cycle framework to include firm heterogeneity and fixed capital adjustment costs. The aggregate state vector of the model contains the distribution of firms over idiosyncratic productivity and capital, which evolves over time in response to aggregate productivity shocks. The dynamics of the distribution must satisfy a complicated fixed-point problem: each firm's investment decision depends on its expectation of the dynamics of the distribution, and the dynamics of the distribution depend on firms' investment decisions. This infinite-dimensional fixed-point problem is at the heart of the computational challenges faced by the heterogeneous agent literature.

My computational method solves this problem accurately and efficiently. Depending on the degree of approximation of the distribution (ranging from 5 to 20 parameters), solving the model takes between 30 and 50 seconds using `Dynare`.[2] Although degrees of approximation on the low end of this range are sufficient to capture the dynamics of aggregate variables, degrees on the high end are necessary to capture changes in the shape of the distribution over time.

I extend this basic analysis in three main ways in order to illustrate the generality of the computational method. First, I compute a nonlinear approximation of the aggregate dynamics; consistent with the results in Khan and Thomas (2008), I find that nonlinearities are quantitatively unimportant in this model. Second, I add an investment-specific productivity shock to the model and show that the method continues to perform

[2]This running time overstates the time it takes to solve the model in each step of an estimation procedure because it includes the time spend processing the model and taking symbolic derivatives. Once these tasks are complete, they do not need to be performed again for different parameter values.

well. When the investment-specific shock is volatile enough, "approximate aggregation" breaks down, suggesting that approximating the distribution with a small number of moments as in Krusell and Smith (1998) would be impractical. Finally, I use the method to solve the heterogeneous household model of Krusell and Smith (1998), which features an occasionally binding borrowing constraint and mass points in the distribution of households.

Finally, to illustrate the power of the method, I estimate a version of the model with aggregate shocks to both investment-specific and total factor productivity using full-information Bayesian techniques. Specifically, I estimate the parameters of the aggregate shock processes conditional on different values of the fixed capital adjustment costs (corresponding to different patterns of micro-level investment behavior). Characterizing the posterior distribution of parameters with Markov Chain Monte Carlo takes less than 42 minutes using `Dynare`. The estimation procedure infers that investment-specific shocks are small when adjustment costs are small, but infers that investment-specific shocks are large when adjustment costs are large. 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 computational method.

Related literature

My method builds heavily on two strands of the computational literature. The first strand of literature approximates the cross-sectional distribution using a parametric family. I adapt the family from Algan, Allais, and Haan (2008), who use it to solve the Krusell and Smith (1998) model. Algan, Allais, and Haan (2008) solve for the dynamics of the parameters of the distribution using a globally accurate projection technique, which is computationally slower than the locally accurate perturbation method that I use.

The second strand of literature I build on uses a mix of globally accurate and locally accurate approximations to solve for the dynamics of heterogeneous agent models. The most closely related paper is Reiter (2009) who, himself building on an idea of Campbell (1998), solves the Krusell and Smith (1998) model using locally accurate approximations with respect to the aggregate state vector. Reiter (2009) approximates the distribution with a fine histogram, which requires many parameters to achieve acceptable accuracy. This choice limits Reiter (2009)'s approach to problems with a low-dimensional individual state space because the size of the histogram grows exponentially in the number of individual states.

Precursors to Reiter (2009)'s method can also be found in Dotsey, King, and Wolman (1999) and Veracierto (2002) in the context of (S,s) models.[3] Childers (2018) formalizes this class of methods in the function space and is able to prove certain properties of

[3] Veracierto (2017) extends Veracierto (2002) to a more general class of models. Unlike my method, Veracierto (2017) does not rely on any direct approximation of the distribution. Instead, Veracierto (2017) 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.

them. Ahn, Kaplan, Moll, Winberry, and Wolf (2018) adapt Reiter (2009)'s approach to continuous time and show how to use model reduction techniques to nonparametrically reduce the size of the distribution. My method reduces the distribution in a parametric way, which allows for a more straightforward extension to nonlinear approximations.[4]

More generally, the method developed in this paper is related to the large body of work which, following Krusell and Smith (1998), approximates the distribution with a small number of moments. This approximation works well if the moments accurately forecast the prices which occur in equilibrium. The advantage of Krusell and Smith (1998) is that it is globally accurate with respect to this reduced aggregate state and, therefore, can capture global nonlinearities more easily than the locally accurate approach pursued in this paper. However, Krusell and Smith (1998) relies on the fact that the distribution can be accurately summarized by a small number of moments. My method instead includes the entire distribution in the aggregate state vector, but relies on a local approximation of the aggregate dynamics.

Road map

I briefly describe the benchmark heterogeneous firm model in Section 2. I then explain how to use the solution method to solve this model in Section 3. In Section 4, I extend the method in various ways to illustrate its generality. In Section 5, I add investment-specific shocks to the model and estimate the parameters of the shock processes using Bayesian techniques. Section 6 concludes. Various Appendices contain additional details not included in the main text.

2. Benchmark RBC model with firm heterogeneity

For concreteness, I illustrate the method using the heterogeneous firm model from Khan and Thomas (2008); however, the method applies to a large class of heterogeneous agent models.[5] In Section 4 and the online code template, I also 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

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

[4] A number of papers pursue a pure perturbation approach with respect to both individual and aggregate state variables. Preston and Roca (2007) approximate the distribution with a particular set of moments and show how to derive the law of motion for those moments analytically given an approximation of the policy function. Mertens and Judd (2017) instead assume a finite but arbitrarily large number of agents and perturb around the point without idiosyncratic or aggregate shocks. They are able to leverage their perturbation approach to analytically prove properties of their method. Evans (2015) follows a related approach but updates the point of approximation around the actual cross-sectional distributions that arise in a simulation.

[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.

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 is the elasticity of output with respect to labor. The aggregate shock z_t is common to all firms and follows the AR(1) process:[6]

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

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

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

In each period, the firm j inherits its capital stock from previous periods' investments, observes the two productivity shocks, hires labor from a competitive labor 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.

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

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

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 $1/\alpha$ is the Frisch elasticity of labor supply. The total time endowment per period is normalized to 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

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

[6] An auxiliary assumption necessary for using perturbation methods is that the aggregate shock process stays in some bounded interval $[\underline{z}, \overline{z}]$. Since that assumption is not central to the analysis, I leave it implicit.

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

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

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

$\widehat{v}(\varepsilon, k; \mathbf{s}) = \mathbb{E}_\xi v(\varepsilon, k, \xi; \mathbf{s})$, and $\mathbf{s}'(z, \mathbf{s})$ is the law of motion for the aggregate state (defined below). 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 value of the fixed cost ξ 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, that is, $\widehat{\xi}(\varepsilon, k; \mathbf{s}) = \min\{\max\{0, \widetilde{\xi}(\varepsilon, k; \mathbf{s})\}, \overline{\xi}\}$.

It will be numerically convenient to approximate the ex ante value function $\widehat{v}(\varepsilon, k; \mathbf{s})$. Given that the extensive margin decision is characterized by the cutoff (4) and the fixed cost ξ is uniformly distributed, the expectation can be computed analytically:

$$\begin{aligned} \widehat{v}(\varepsilon, k; \mathbf{s}) = \lambda(\mathbf{s}) \max_n \big\{e^z e^\varepsilon k^\theta n^\nu - w(\mathbf{s})n\big\} \\ + \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\overline{\xi}}\left(v^a(\varepsilon, k; \mathbf{s}) - \lambda(\mathbf{s})w(\mathbf{s})\frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{2}\right) \\ + \left(1 - \frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\overline{\xi}}\right)v^n(\varepsilon, k; \mathbf{s}). \end{aligned} \tag{5}$$

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 density of firms over (ε, k)-space, $g(\varepsilon, k)$.[7]

Definition. 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 $\mathbf{s}'(z'; \mathbf{s}) = (z'; g'(z, g))$ such that:

(i) (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 (2)–(5).

(ii) (Implications of household optimization) For all $\mathbf{s}$,

[7] The computational method does not rely on the fact that the distribution is characterized by its density. In Section 4, I discuss how to extend the method to allow for mass points in the distribution.

TABLE 1. Baseline Parameterization.

Parameter	Description	Value	Parameter	Description	Value
β	Discount factor	0.961	ρ_z	Aggregate TFP AR(1)	0.859
σ	Utility curvature	1	σ_z	Aggregate TFP AR(1)	0.014
α	Inverse Frisch	$\lim \alpha \to 0$	$\overline{\xi}$	Fixed cost	0.0083
χ	Labor disutility	$\implies N^* = \frac{1}{3}$	a	No fixed cost region	0.011
ν	Labor share	0.64	ρ_ε	Idiosyncratic TFP AR(1)	0.859
θ	Capital share	0.256	σ_ε	Idiosyncratic TFP AR(1)	0.022
δ	Capital depreciation	0.085			

Note: Parameterization follows Khan and Thomas (2008), Table 1. Parameter values have been adjusted for the fact that my model does not contain trend growth. Disutility of labor supply χ chosen to ensure steady state labor supply $N^* = \frac{1}{3}$.

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

(iii) (Law of motion for distribution) For all (ε', k'),

$$
\begin{aligned}
g'(\varepsilon', k'; \mathbf{s}) &= \iiint \Bigg[\mathbb{1}\{\rho_\varepsilon \varepsilon + \sigma_\varepsilon \omega'_\varepsilon = \varepsilon'\} \times \Bigg[\frac{\widehat{\xi}(\varepsilon, k; \mathbf{s})}{\overline{\xi}} \mathbb{1}\{k^a(\varepsilon, k; \mathbf{s}) = k'\} \\
&\quad + \left(1 - \frac{\widehat{\xi}(\varepsilon, k; z, \mathbf{m})}{\overline{\xi}}\right) \mathbb{1}\{k^n(\varepsilon, k; \mathbf{s}) = k'\}\Bigg]\Bigg] \times p(\omega'_\varepsilon) g(\varepsilon, k)\, d\omega'_\varepsilon\, d\varepsilon\, dk,
\end{aligned} \tag{6}
$$

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

(iv) (Law of motion for aggregate shock) $z' = \rho_z z + \sigma_z \omega'_z$, where $\omega'_z \sim N(0, 1)$.

2.4 *Baseline parameterization*

I parameterize the model following Khan and Thomas (2008). The parameter values are reported in Table 1. 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 by Khan and Thomas (2008) to match features of the investment rate distribution reported in Cooper and Haltiwanger (2006).

3. THE COMPUTATIONAL METHOD

The computational method involves three main steps. The first step is to approximate the model's equilibrium objects—importantly, the value function and distribution—using finite-dimensional global approximations with respect to individual state variables. This step yields a finite-dimensional representation of the equilibrium for a given realization of the aggregate state. The second step of the method is to compute the stationary equilibrium of the finite-dimensional aproximation of the model in which there are no aggregate shocks but still idiosyncratic shocks. The final step is to compute the aggregate dynamics of the finite-dimensional model using a locally accurate Taylor expansion around the stationary equilibrium.

3.1 *Step 1: Approximate equilibrium using finite-dimensional objects*

The value function and cross-sectional distribution are infinite-dimensional functions of the individual state variable (ε, k) and the aggregate state variable (z, g). I approximate these functions with respect to individual state (ε, k) using globally accurate projection methods. I will approximate these functions with respect to the aggregate state (z, g) using locally accurate perturbation methods in Step 3.

Distribution Following Algan, Allais, and Haan (2008), I approximate the density $g(\varepsilon, k)$ with the functional form

$$g(\varepsilon, k) \approx g_0 \exp\Bigg\{ g_1^1(\varepsilon - m_1^1) + g_1^2(\log k - m_1^2) + \sum_{i=2}^{n_g} \sum_{j=0}^{i} g_i^j \big[(\varepsilon - m_1^1)^{i-j} (\log k - m_1^2)^j - m_i^j \big] \Bigg\}, \tag{7}$$

where n_g indexes the degree of approximation, g_0, g_1^1, g_1^2, $\{g_i^j\}_{i,j=(2,0)}^{(n_g,i)}$ are parameters, and m_1^1, m_1^2, $\{m_i^j\}_{i,j=(2,0)}^{(n_g,i)}$ are centralized moments of the distribution. The key difference from Algan, Allais, and Haan (2008) is that the distribution is over a two-dimensional vector rather than a univariate one.[8] A $n_g = 2$ degree polynomial in this family corresponds to a multivariate normal distribution; $n_g > 2$ polynomials allow for nonnormal features, such as skewness or excess kurtosis.

The parameter vector $\mathbf{g} = (g_0, \dots, g_{n_g}^{n_g})$ and moment vector $\mathbf{m} = (m_1^1, \dots, m_{n_g}^{n_g})$ must be consistent with each other in the sense that the moments are actually implied by the parameters:

$$\begin{aligned} m_1^1 &= \iint \varepsilon g(\varepsilon, k)\, d\varepsilon\, dk, \\ m_1^2 &= \iint \log k g(\varepsilon, k)\, d\varepsilon\, dk \quad \text{and} \\ m_i^j &= \iint (\varepsilon - m_1^1)^{i-j} (\log k - m_1^2)^j g(\varepsilon, k)\, d\varepsilon\, dk \quad \text{for } i = 2, \dots, n_g,\ j = 0, \dots, i. \end{aligned} \tag{8}$$

Given a vector of parameters $\mathbf{m}$, Algan, Allais, and Haan (2008) develop a simple and robust method for solving the system (8) for the associated parameters $\mathbf{g}$.[9] Hence, the vector of moments $\mathbf{m}$ completely characterizes the approximated density. I therefore

[8] It may be possible to further reduce the dimensionality of the approximation using copulas to link together two univariate distributions. I do not pursue that approach here but thank an anonymous referee for the suggestion.

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

use the moments $\mathbf{m}$ as the characterization of the distribution, and approximate the infinite-dimensional aggregate state (z, g) with the finite-dimensional representation $(z, \mathbf{m})$.[10]

The fact that the distribution is completely characterized by its moments $\mathbf{m}$ suggests a convenient method for approximating the law of motion (6): use the law of motion for the moments themselves. Using (6), the law of motion for moments is given by

$$
\begin{aligned}
m_1^{1\prime} &= \iiint \left(\rho_\varepsilon \varepsilon + \omega'_\varepsilon\right) p\left(\omega'_\varepsilon\right) g(\varepsilon, k; \mathbf{m})\, d\omega'_\varepsilon\, d\varepsilon\, dk, \\
m_1^{2\prime} &= \iiint \left[\frac{\widehat{\xi}(\varepsilon, k; z, \mathbf{m})}{\overline{\xi}} \log k^a(\varepsilon, k; z, \mathbf{m}) \right. \\
&\quad \left. + \left(1 - \frac{\widehat{\xi}(\varepsilon, k; z, \mathbf{m})}{\overline{\xi}}\right) \log k^n(\varepsilon, k; z, \mathbf{m})\right] p\left(\omega'_\varepsilon\right) g(\varepsilon, k; \mathbf{m})\, d\omega'_\varepsilon\, d\varepsilon\, dk, \\
m_i^{j\prime}(z, \mathbf{m}) &= \iiint \left[\left(\rho_\varepsilon \varepsilon + \omega'_\varepsilon - m_1^{1\prime}\right)^{i-j} \left\{\frac{\widehat{\xi}(\varepsilon, k; z, \mathbf{m})}{\overline{\xi}} \left(\log k^a(\varepsilon, k; z, \mathbf{m}) - m_1^{2\prime}\right)^j \right.\right. \\
&\quad \left.\left. + \left(1 - \frac{\widehat{\xi}(\varepsilon, k; z, \mathbf{m})}{\overline{\xi}}\right) \left(\log k^n(\varepsilon, k; z, \mathbf{m}) - m_1^{2\prime}\right)^j \right\}\right] \\
&\quad \times p\left(\omega'_\varepsilon\right) g(\varepsilon, k; \mathbf{m})\, d\omega'_\varepsilon\, d\varepsilon\, dk.
\end{aligned} \tag{9}
$$

The system (9) provides a mapping from the current aggregate state into next period's moments $\mathbf{m}'(z, \mathbf{m})$ by integrating decision rules against the implied density.[11] I solve for the steady state values of the moments $\mathbf{m}^*$ by iterating on this mapping. Since the mapping is nonlinear there is no guarantee that this iteration converges, but I have found that it does in practice.[12]

Firm's value functions I approximate the firms' ex ante value function $\widehat{v}(\varepsilon, k; z, \mathbf{m})$ with respect to individual states using orthogonal polynomials:[13]

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

[10]This approximation requires that moments of the distribution exist up to the order of the approximation, which is restrictive for models in which the distribution features a fat tail. In those cases, a different parametric family should be used.

[11]I numerically compute these integrals using two-dimensional Gauss–Legendre quadrature, which replaces the integral with a finite sum. See the online codes and user guide for details of this procedure.

[12]In practice, higher degree approximations of the distribution n_g require more points in the quadrature in order to accurately compute the integrals.

[13]The choice of orthogonal polynomials is not essential to the method and is made primarily for numerical convenience. In the online codes and user guide, I show how to use linear splines instead. Furthermore, the choice of tensor product polynomials is not necessary, and I have used the set of complete polynomials and Smolyak polynomials as well.

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.[14] I solve for the dependence of these coefficients on the aggregate state in the perturbation step.

With this particular approximation of the value function, it is natural to approximate the Bellman equation (5) using collocation, which forces the equation to hold exactly at a set of grid points $\{\varepsilon_i, k_j\}_{i,j=1,1}^{n_\varepsilon, n_k}$:

$$\begin{aligned}
&\widehat{v}(\varepsilon_i, k_j; z, \mathbf{m}) \\
&\quad = \lambda(z, \mathbf{m}) \max_n \bigl\{ e^z e^{\varepsilon_i} k_j^\theta n^\nu - w(z, \mathbf{m}) n \bigr\} + \lambda(z, \mathbf{m})(1-\delta) k_j \\
&\qquad + \left(\frac{\widehat{\xi}(\varepsilon_i, k_j; z, \mathbf{m})}{\bar{\xi}} \right) \Biggl(-\lambda(z, \mathbf{m}) \biggl(k^a(\varepsilon_i, k_j; z, \mathbf{m}) \\
&\qquad + w(z, \mathbf{m}) \frac{\widehat{\xi}(\varepsilon_i, k_j; z, \mathbf{m})}{2} \biggr) \\
&\qquad + \beta \mathbb{E}_{z'|z} \biggl[\int \widehat{v}\bigl(\rho_\varepsilon \varepsilon_i + \sigma_\varepsilon \omega_\varepsilon', k^a(\varepsilon_i, k_j; z, \mathbf{m}); z', \mathbf{m}'(z, \mathbf{m})\bigr) p\bigl(\omega_\varepsilon'\bigr) \, d\omega_\varepsilon' \biggr] \Biggr) \\
&\qquad + \left(1 - \frac{\widehat{\xi}(\varepsilon_i, k_j; z, \mathbf{m})}{\bar{\xi}} \right) \Biggl(-\lambda(z, \mathbf{m}) k^n(\varepsilon_i, k_j; z, \mathbf{m}) \\
&\qquad + \beta \mathbb{E}_{z'|z} \biggl[\int \widehat{v}\bigl(\rho_\varepsilon \varepsilon_i + \sigma_\varepsilon \omega_\varepsilon', k^n(\varepsilon_i, k_j; z, \mathbf{m}); z', \mathbf{m}'(z, \mathbf{m})\bigr) p\bigl(\omega_\varepsilon'\bigr) \, d\omega_\varepsilon' \biggr] \Biggr),
\end{aligned} \tag{10}$$

where the decision rules are computed from the value function via first order conditions.[15] 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 is taken 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 in the perturbation step.

Approximate equilibrium conditions With all of these approximations, the recursive competitive equilibrium becomes computable, replacing the true aggregate state (z, g) with the approximate aggregate state $(z, \mathbf{m})$, the true Bellman equation (5) with the Chebyshev collocation approximation (10), and the true law of motion for the distribution (6) with the approximation (9). I show in Appendix A that these approximate equilibrium conditions can be represented by a residual function $f : \mathbb{R}^{2n_\varepsilon n_k + n_g + 2} \times \mathbb{R}^{2n_\varepsilon n_k + n_g + 2} \times \mathbb{R}^{n_g+1} \times \mathbb{R}^{n_g+1} \times \mathbb{R} \to \mathbb{R}^{2n_\varepsilon n_k + n_g + 2 + n_g + 1}$ that satisfies

$$\mathbb{E}_{\omega_z'}\bigl[f\bigl(\mathbf{y}', \mathbf{y}, \mathbf{x}', \mathbf{x}; \psi\bigr)\bigr] = 0, \tag{11}$$

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

[15] The use of first-order conditions exploits the fact that the first-order condition is a smooth function of k', due to taking the expectation over next periods draws of the fixed cost ξ' and idiosyncratic productivity shock ε'.

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.

The residual function (11) is exactly the canonical form studied by Schmitt-Grohe and Uribe (2004) who, following Judd and Guu (1997) and others, show how to solve such systems using perturbation methods. The remainder of the computational method simply follows the steps outlined in Schmitt-Grohe and Uribe (2004).

3.2 *Step 2: Compute stationary equilibrium with no aggregate shocks*

In terms of Schmitt-Grohe 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

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

In principle, this is a generic system of nonlinear equations which can be numerically solved using root-finding algorithms; in practice, the system is large so that numerical solvers often fail to converge. I instead solve the system using a stable root-finding problem in the wage w^* only, similar to the algorithm developed in Hopenhayn and Rogerson (1993). Details are provided in Appendix A.

Figure 1 shows that a moderate degree approximation is necessary to capture the shape of the steady state distribution of firms. The figure plots various slices of the stationary distribution for different degrees of approximation n_g in the parametric family (7) and compares them to a fine histogram.[16] The marginal distribution of productivity in Panel (a) is normal, so a $n_g = 2$ degree approximation gives an exact match to the histogram. In contrast, the marginal distribution of capital in Panel (b) features positive skewness and excess kurtosis, and the conditional distribution of capital varies in both location and shape as a function of productivity. A $n_g = 4$ degree approximation captures these complicated shapes. Even a $n_g = 2$ approximation also does quite well, indicating that the true distribution is close to log-normal.

Moderate degree approximations of the distribution are also sufficient to provide a good approximation of aggregate variables in steady state. Table 2 reports various aggregate variables using different degrees of approximation and again compares them to the values obtained using a fine histogram. An $n_g = 2$ degree approximation yields aggregates that are virtually indistinguishable from higher degree approximation or the histogram.

3.3 *Step 3: Compute aggregate dynamics using perturbation*

A solution to the dynamic problem (11) is of the form

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

[16] The fine histogram approximates the distribution of firms along a discrete of points, as in Young (2010). To compute the histogram, I use the same algorithm used to compute the steady state discussed in Appendix A, but approximate the distribution as a histogram instead of using the parametric family (7).

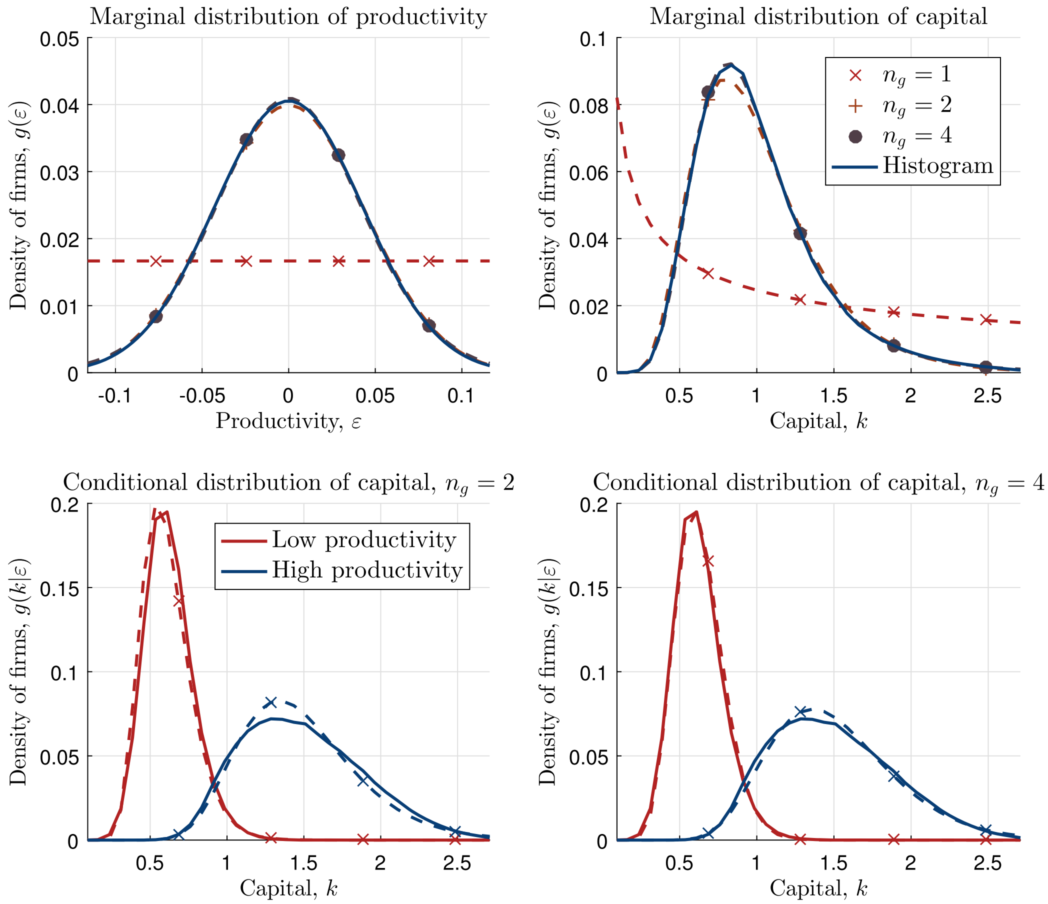

FIGURE 1. Stationary distribution of firms for different degrees of approximation. *Notes*: slices of the invariant distribution of firms over productivity ε and capital k. n_g is the order of the polynomial used in the parametric family (7). "Histogram" is the steady state distribution computed using a fine histogram rather than the parametric family. Marginal distributions are computed by numerical integration of the joint p.d.f. "High productivity" and "low productivity" correspond to approximately $+/-$ two standard deviations of the productivity distribution.

TABLE 2. Aggregate Variables in Steady State.

Variable	$\mathbf{n}_g = 1$	$\mathbf{n}_g = 2$	$\mathbf{n}_g = 3$	$\mathbf{n}_g = 4$	$\mathbf{n}_g = 6$	Histogram
Output	0.509	0.498	0.499	0.500	0.499	0.498
Consumption	0.555	0.412	0.413	0.414	0.413	0.413
Capital	1.200	1.006	1.008	1.010	1.008	1.007
Wage	0.984	0.958	0.958	0.958	0.958	0.958
Marginal utility	1.803	2.426	2.420	2.416	2.421	2.421

Note: Aggregates in stationary equilibrium computed using various orders of approximation. "Histogram" is the steady state distribution computed using a fine histogram rather than the parametric family.

where $\eta = (1, \mathbf{0}_{n_g \times 1})'$ and ψ is the perturbation parameter. Perturbation methods approximate the solution g and h using Taylor expansions around the point where $\psi = 0$, which corresponds to the steady state values $\mathbf{x}^*$ and $\mathbf{y}^*$. For example, a first-order Taylor expansion gives:

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

The unknowns in the approximation (12) are the partial derivatives $g_{\mathbf{x}}$, g_ψ, $h_{\mathbf{x}}$, and h_ψ. Schmitt-Grohe 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-Grohe and Uribe (2004) for further details.

`Dynare` implements this perturbation procedure completely automatically. First, it computes the derivatives of the equilibrium conditions (11), which gives a system of equations involving the partial derivatives of the solution g and h. Second, `Dynare` solves that system using standard linear rational expectation model solvers; see Adjemian, Bastani, Julliard, Karame, Mihoubi, Perendia, Pfeifer, Ratto, and Villemot (2011) for more details. An analogous procedure can be used to compute higher-order nonlinear approximations of g and h with no additional coding required.[17]

Online code template The online code template and user guides shows how to implement the method in `Dynare`. Broadly speaking, the user provides two sets of codes. First, the user provides `Matlab .m` files that compute the steady state of the model, that is, $\mathbf{x}^*$ and $\mathbf{y}^*$. Second, the user provides a `Dynare .mod` files that define the model's equilibrium conditions, that is, $f(\mathbf{y}', \mathbf{y}, \mathbf{x}', \mathbf{x}; \boldsymbol{\psi})$. In addition to the perturbation step described above, `Dynare` will, if requested, simulate the model, compute theoretical and/or empirical moments of simulated variables, and estimate the model using likelihood-based methods.

First order approximation Using `Dynare`, computing a first-order approximation of the model's dynamics takes between 35 seconds for a $n_g = 2$ degree approximation of the distribution and 47 seconds for a $n_g = 4$ degree approximation. The majority of time is spent computing the stationary equilibrium of the model. The remaining time is spent evaluating the derivatives of the model's equilibrium conditions f at steady state and solving the resulting dynamic system.[18]

The dynamics of aggregate variables are well in line with what Khan and Thomas (2008) and Terry (2017b) have reported using different algorithms to solve the model. Figure 2 plots the impulse responses of these variables to a positive aggregate TFP shock.

[17]Moving to higher-order approximations requires solving additional equations, but as described in Schmitt-Grohe and Uribe (2004), these systems are linear and thus in principle simple to solve. In practice, the system is large and dense, which places computational limitations. In the Khan and Thomas (2008) model, I have found that a second-order approximation is feasible using `Dynare` (see Section 4).

[18]Runtimes are computed using using `Dynare 4.5.3` in `Matlab R2016a` on a 3.10 GHz Windows desktop PC with 32 GB of RAM.

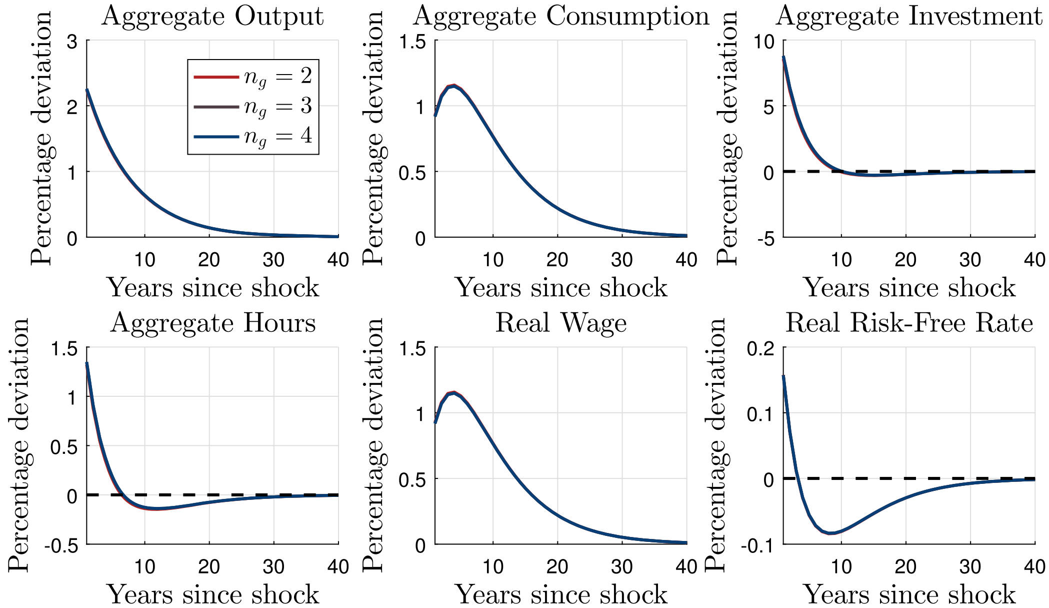

FIGURE 2. Impulse responses of aggregates, first-order perturbation. *Notes*: Impulse responses of aggregate variables for different degrees of approximation of the distribution. n_g refers to the degree of approximation used in the parametric family (7).

TABLE 3. Aggregate Business Cycle Statistics, First-Order Perturbation.

SD (rel. to output)	$\mathbf{n}_g = 2$	$\mathbf{n}_g = 4$	Corr. with Output	$\mathbf{n}_g = 2$	$\mathbf{n}_g = 4$
Output	(2.14%)	(2.16%)	×	×	×
Consumption	0.48	0.47	Consumption	0.90	0.90
Investment	3.86	3.93	Investment	0.97	0.97
Hours	0.61	0.61	Hours	0.95	0.94
Real wage	0.48	0.47	Real wage	0.90	0.90
Real interest rate	0.08	0.08	Real interest rate	0.80	0.79

Note: Business cycle fluctuations of key aggregate variables under a first-order perturbation. All variables have been 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 itself.

Higher TFP directly increases aggregate output, but also increases investment and labor demand, which further increase output and factor prices. Households respond to higher permanent income by increasing consumption and to higher wages by increasing labor supply.

The resulting business cycle statistics of aggregate variables are reported in Table 3. As usual in a real business cycle model, consumption is roughly half as volatile as output, investment is nearly four times as volatile as output, and labor is slightly less volatile than output. All variables are highly correlated with each other because aggregate TFP is the only shock driving fluctuations in the model.

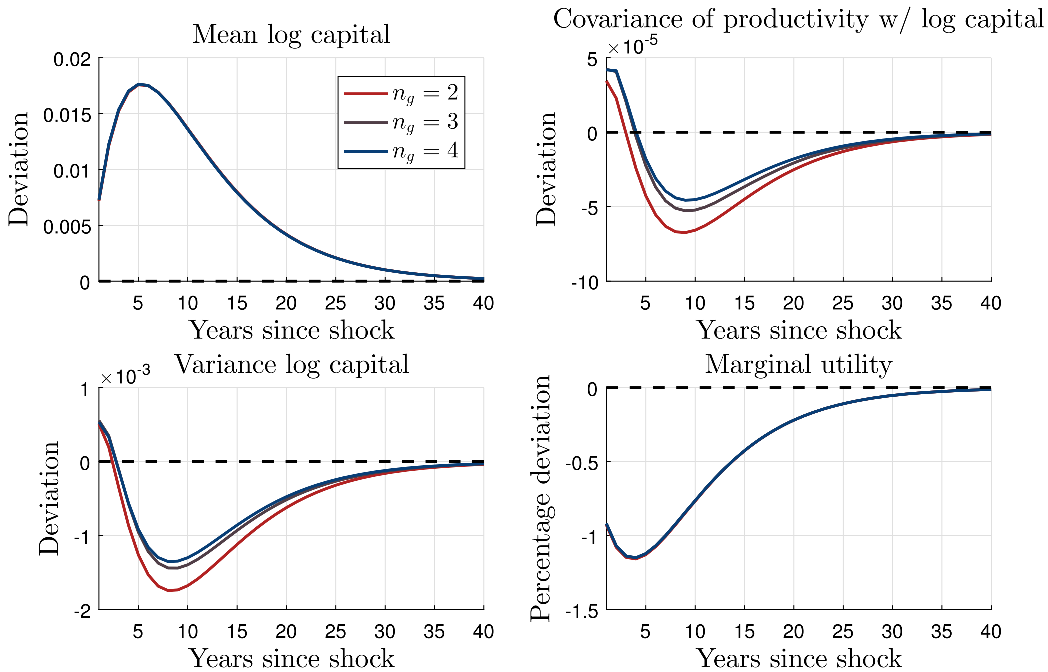

FIGURE 3. Impulse responses of distributional variables, first-order approximation. *Notes*: Impulse responses of distributional variables for different degrees of approximation of the distribution. n_g refers to the degree of approximation used in the parametric family (7).

The aggregate dynamics are largely unaffected by the degree n_g of the parametric family (7) approximating the distribution. Visually, increasing the degree of approximation from $n_g = 2$ to $n_g = 4$ barely changes the impulse response functions in Figure 2. Quantitatively, the business cycle statistics in Table 3 barely change as well. Terry (2017a) finds that the aggregate dynamics implied by my method are quantitatively close to the results of other methods used in the literature, such as Krusell and Smith (1998).

The dynamics of cross-sectional features of the distribution are more sensitive to the degree of approximation n_g than are the dynamics of aggregate variables. Figure 3 plots the impulse responses of the cross-sectional $\mathbb{E}[\log k]$, $\mathbb{C}\text{ov}(\varepsilon, \log k)$, and $\mathbb{V}\text{ar}(\log k)$ with respect to a TFP shock. Although the response of mean log capital is virtually identical across the degrees of approximation n_g, the responses of the other two statistics change with n_g. However, these differences are small and do not generate differences in the dynamics of aggregate variables described above. When computing their decisions, firms must forecast the level of marginal utility λ; Figure 3 shows that the impulse response of marginal utility is virtually identical for the different degrees of approximation.[19]

Table 4 confirms these graphical results by computing time-series properties of the dynamics of the cross-sectional statistics plotted in Figure 3; although the cyclicality

[19] Given the linear disutility of labor, the real wage is an explicit function of the marginal utility of consumption.

TABLE 4. Distributional Business Cycle Statistics, First-Order Perturbation.

$\mathbb{E}[\log k]$	$\mathbf{n}_g = \mathbf{2}$	$\mathbf{n}_g = \mathbf{3}$	$\mathbf{n}_g = \mathbf{4}$	$\mathbb{C}\text{ov}(\varepsilon, \log k)$	$\mathbf{n}_g = \mathbf{2}$	$\mathbf{n}_g = \mathbf{3}$	$\mathbf{n}_g = \mathbf{4}$
Mean	−0.0899	−0.0822	−0.0824	Mean	0.0123	0.0121	0.0122
SD	0.0125	0.0126	0.0127	SD	6.7e−5	7.2e−5	6.9e−5
Corr w/ Y	0.6017	0.6095	0.6117	Corr w/ Y	0.7157	0.8276	0.8432
Autocorr	0.8280	0.8268	0.8264	Autocorr	0.7472	0.7339	0.7240
$\mathbb{V}\text{ar}(\log k)$	$\mathbf{n}_g = \mathbf{2}$	$\mathbf{n}_g = \mathbf{3}$	$\mathbf{n}_g = \mathbf{4}$	Marginal Utility	$\mathbf{n}_g = \mathbf{2}$	$\mathbf{n}_g = \mathbf{3}$	$\mathbf{n}_g = \mathbf{4}$
Mean	0.1529	0.1476	0.1485	Mean	0.8995	0.8934	0.8931
SD	0.0014	0.0013	0.0012	SD	0.0103	0.0102	0.0101
Corr w/ Y	0.5752	0.6608	0.6539	Corr w/ Y	−0.8999	−0.9001	−0.8999
Autocorr	0.7980	0.7823	0.7782	Autocorr	0.6704	0.6712	0.6715

Note: Business cycle fluctuations of key features of the distribution under a first-order perturbation. All variables have been HP-filtered with smoothing parameter $\lambda = 100$. Marginal utility has been logged.

of $\mathbb{C}\text{ov}(\varepsilon, \log k)$ and $\mathbb{V}\text{ar}(\log k)$ differ with the degree of approximation n_g, their overall variation is limited.

4. Generalizations of the method

In order to be concrete, Section 3 developed the computational method in the context of the Khan and Thomas (2008) model. However, this model contains a number of special features which are not central to the application of the method. In this section, I discuss how to generalize the method to cases where these special features do not hold.

4.1 *Aggregate nonlinearities*

Second-order approximation I incorporate aggregate nonlinearities by computing a second-order approximation of the model, which amounts to changing a `Dynare` option from `order = 1` to `order = 2`. Quantitatively, the dynamics of the second-order approximation closely resemble those of the first-order approximation. Furthermore, Figure 4 show that the impulse responses in a second-order approximation feature quantitatively small sign-dependence and state-dependence. These results are consistent with Khan and Thomas (2008), who find little evidence of such nonlinearities using the nonlinear Krusell and Smith (1998) algorithm.

Limitations of a local approximation Although the method can capture aggregate nonlinearities with a local approximation, it is not well suited for capturing nonlinearities that are global in nature. For example, portfolio choice problems in which assets differ only in their aggregate risk characteristics cannot be directly solved using the `Dynare` code template because the stationary distribution of portfolios is not unique.[20] In addition, the method is ill-suited to solve models which feature the economy transitioning between multiple steady states, such as Brunnermeier and Sannikov (2014).

[20] In order to solve such models, one could potentially adapt the technique in Devereux and Sutherland (2011).

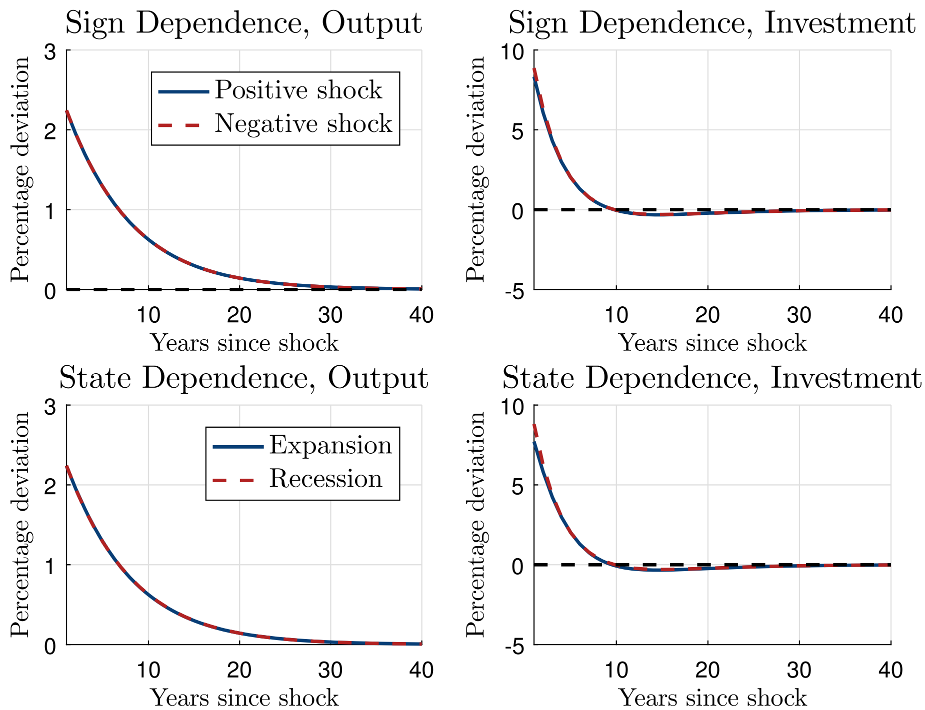

FIGURE 4. Sign and size dependence in impulse responses, second-order approximation. *Notes*: Nonlinear features of the impulse responses of aggregate output (left column) and investment (right column). "Sign dependence" refers to the impulse response to a one standard deviation positive versus negative shock. Negative shocks have been multiplied by -1. "State dependence" refers to the impulse response after positive one standard deviation shocks versus negative one standard deviation shocks in the previous two years. All impulse responses computed nonlinearly as in Koop, Pesaran, and Potter (1996).

It is important to note that these limitations apply to aggregate dynamics only; the method does compute a fully global approximation of individual behavior. Hence, the method can be used to solve models in which aggregate shocks affect the distribution of idiosyncratic shocks, such as the uncertainty shocks in Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018).

4.2 *Method does not require approximate aggregation*

Khan and Thomas (2008) show that "approximate aggregation" holds in this model in the sense that the aggregate capital stock almost completely characterizes how the distribution of firms 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 than Krusell and Smith (1998); in a comparison project, Terry (2017a) shows that his implementation of my method solves and simulates the model in 0.098% of the time of the Krusell and Smith (1998) method. This gain in speed makes the full-information Bayesian estimation in Section 5 feasible.

Furthermore, because my method directly approximates the distribution, it can be easily applied to other models in which approximate aggregation fails. I make this case concrete in Appendix B by adding aggregate 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 aggregate dynamics. Extending Krusell and Smith (1998)'s algorithm would therefore require adding more moments to the forecasting rule, which quickly becomes infeasible since each additional moment adds another state variable. Hence, my method is not only significantly faster for models in which approximate aggregation holds, but also applies to models in which it fails.

4.3 *Occasionally binding constraints and mass points*

Another special feature of the Khan and Thomas (2008) model is that the distribution is characterized by its density $g(\varepsilon, k)$, which makes a smooth polynomial approximation of the distribution straightforward. In other models, the distribution may not be characterized by its density; for example, in the Krusell and Smith (1998) model, the occasionally binding borrowing constraint introduces mass points into the distribution of households. The online code template and user guide shows how to extend the method to solve the Krusell and Smith (1998) model. I simply add an additional parameter to the distribution approximation—the mass of households at the borrowing constraint—and approximate the distribution away from the constraint with the polynomial family.[21] In principle, one could extend this procedure to incorporate multiple mass points as well.

5. 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 impact inference. To do so, I extend the benchmark Khan and Thomas (2008) model to include aggregate investment-specific productivity shocks in addition to the neutral shocks already in the model, and estimate the parameters of the two aggregate shock processes. I include investment-specific shocks because inference about this process is most directly shaped by micro-level investment behavior. The investment-specific shock only affects the capital accumulation equation, which becomes $k_{jt+1} = (1-\delta)k_{jt} + e^{q_t} i_{jt}$, where q_t is the investment-specific productivity shock. 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 in the investment-specific process, σ_{qz}, in order to capture

[21] Strictly speaking, the distribution in the Krusell and Smith (1998) model is a collection of mass points, and I approximate the collection of points away from the borrowing constraints with a smooth polynomial.

TABLE 5. Parameterizations Considered in Estimation Exercises.

Fixed Parameters	Value	Fixed Parameters	Value		
β (discount factor)	0.99	a (no fixed cost)	0.011		
σ (utility curvature)	1	ρ_ε (idioynscratic TFP)	0.859		
α (inverse Frisch)	$\lim \alpha \to 0$	ν (capital share)	0.256		
χ (labor disutility)	$\implies N^* = \frac{1}{3}$	σ_ε (idiosyncratic TFP)	0.015		
θ (labor share)	0.64	Changing Parameter	Value 1	Value 2	Value 3
δ (depreciation)	0.025	$\overline{\xi}$	0.01	0.1	1

Note: 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.

comovement between the two shocks. Without this loading factor, investment-specific shocks would induce a counterfactually negative comovement between consumption and investment and, therefore, be immediately rejected by the data. Denote the vector of parameter values $\Theta = (\rho_z, \sigma_z, \rho_q, \sigma_q, \sigma_{qz})$.

I estimate the parameters of the shock processes Θ conditional on three different values for the remaining parameters. The only parameter to vary across these parameterizations is $\overline{\xi}$, the upper bound on fixed cost draws. I vary the fixed costs from $\overline{\xi} = 0.01$ to $\overline{\xi} = 1$. These parameter values 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, $p(\Theta)$, with the likelihood function, $L(\mathbf{Y}_{1:T}|\Theta)$ where $\mathbf{Y}_{1:T}$ is the observed time series of data, to form the posterior distribution of parameters, $p(\boldsymbol{\Theta}|\mathbf{Y}_{1:T})$. The posterior is proportional to $p(\Theta)L(\mathbf{Y}_{1:T}|\Theta)$, which is the object I characterize numerically. The data I use is $\mathbf{Y}_{1:T} = (\log \widehat{Y}_{1:T}, \log \widehat{I}_{1:T})$, 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; see Appendix C for details. I choose relatively standard prior distributions to form $p(\Theta)$, also contained in Appendix C. I sample from $p(\Theta)L(\mathbf{Y}_{1:T}|\Theta)$ using the Metropolis–Hastings algorithm; since this procedure is now standard (see, e.g., Fernandez-Villaverde, Rubio-Ramirez, and Schorfheide (2016)), I omit further details. `Dynare` computes 10,000 draws from the posterior in 44 minutes, 32 seconds.[22]

Table 6 reports the estimated aggregate shock processes for the three different micro-level parameterizations considered in Table 5. As the upper bound of the fixed costs increases from $\overline{\xi} = 0.01$ to $\overline{\xi} = 1$, the estimated variance of investment-specific shocks significantly increases from $\sigma_q = 0.0056$ to $\sigma_q = 0.0077$; hence, matching the aggregate investment data with larger adjustment frictions requires more volatile shocks. Additionally, the loading on neutral shocks in the investment-specific shock process

[22]A key reason the 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.

TABLE 6. Posterior Distribution of Parameters.

Micro-Level	TFP ρ_z	TFP σ_z	ISP ρ_q	ISP σ_q	ISP Loading σ_{zq}
$\overline{\xi} = 0.01$	0.9793	0.0080	0.9694	0.0056	−0.0045
[90% HPD]	[0.9695, 0.9925]	[0.0074, 0.0086]	[0.9528, 0.9871]	[0.0047, 0.0067]	[−0.0058, −0.0031]
$\overline{\xi} = 0.1$	0.9797	0.0080	0.9695	0.0060	−0.0043
	[0.9668, 0.9933]	[0.0074, 0.0086]	[0.9500, 0.9852]	[0.0049, 0.0072]	[−0.0057, −0.0030]
$\overline{\xi} = 1$	0.9789	0.0080	0.9711	0.0077	−0.0037
	[0.9650, 0.9919]	[0.0074, 0.0086]	[0.9524, 0.9894]	[0.0059, 0.0093]	[−0.0052, −0.0021]

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

TABLE 7. Estimated Variance Decomposition.

Micro-Level	Share from ω_t^z (Neutral)	Share from ω_t^q (Investment-Specific)
$\overline{\xi} = 0.01$		
Output	84.73%	15.33%
Consumption	91.21%	8.79%
Investment	35.08%	64.92%
$\overline{\xi} = 1$		
Output	84.08%	15.92%
Consumption	90.56%	9.44%
Investment	34.27%	65.73%

Note: Variance decomposition of aggregate output, consumption, and investment under two of the estimations reported in Table 6.

shrinks because larger frictions reduce the negative comovement between consumption and investment. The remaining parameters are broadly constant over the different specifications, indicating that micro-level adjustment frictions mainly matter for the inference of the investment-specific shock process.

Table 7 reports the variance decomposition of aggregate output, consumption, and investment under the estimated parameters corresponding to two different parameterizations of micro-level behavior. With relatively flexible capital adjustment ($\overline{\xi} = 0.01$), most of the variation in output and consumption is driven by TFP, due to the fact that the estimated investment-specific shock process is relatively unimportant under this parameterization. Although the investment-specific shock accounts for a slightly larger share of fluctuations with large adjustment frictions ($\overline{\xi} = 1$), the differences are quantitatively small. Of course, the ideal estimation exercise would jointly estimate the capital adjustment frictions and aggregate shock processes using both micro- and macro-level data. The illustrative results in this section show that, using my method, such exercises are now feasible.

6. Conclusion

This paper has developed a general-purpose computational method for solving and estimating heterogeneous agent models. In contrast to much of the existing literature, the method does not rely on the dynamics of the distribution being well-approximated by a small number of moments, expanding the class of models which can be feasibly computed. I have provided codes and a user guide to implement the method using `Dynare` with the hope that it will bring heterogeneous agent models into the fold of standard quantitative macroeconomic analysis.

A particularly promising avenue for future research is incorporating micro data into the estimation of DSGE models. As I showed in Section 5, micro-level behavior places important restrictions on estimated model parameters. In the current representative agent DSGE literature, these restrictions are either completely absent or imposed only indirectly through prior beliefs. The computational method I developed in this paper instead allows the micro data to formally place these restrictions itself.

Appendix 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 approximate the value functions with degree $n_\varepsilon = 3$ and $n_k = 5$ polynomials. I approximate integrals with respect to the distribution using a tensor-product Gauss–Legendre quadrature of order 8 with respect to ε and 10 with respect to k, so that the total number of points is $m_g = 80$. Finally, I approximate integrals with respect to idiosyncratic shock using a $m_\varepsilon = 3$ degree Gauss–Hermite quadrature.

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

$$
\begin{aligned}
0 = \mathbb{E}\Bigg[&\sum_{k=1}^{n_\varepsilon}\sum_{l=1}^{n_k}\theta_{kl}T_k(\varepsilon_i)T_l(k_j) - \lambda\big(e^z e^{\varepsilon_i}k_j^\theta n(\varepsilon_i, k_j)^\nu - wn(\varepsilon_i, k_j)\big) - \lambda(1-\delta)k_j \\
&- \left(\frac{\widehat{\xi}(\varepsilon_i, k_j)}{\overline{\xi}}\right)\Bigg(-\lambda\left(k^a(\varepsilon_i, k_j) - w\frac{\widehat{\xi}(\varepsilon_i, k_j)}{2}\right) \\
&+ \beta\sum_{o=1}^{m_\varepsilon}\tau_o^\varepsilon\sum_{k'=1}^{n_\varepsilon}\sum_{l'=1}^{n_k}\theta'_{k'l'}T_{k'}\big(\rho_\varepsilon\varepsilon_i + \sigma_\varepsilon\omega_o^\varepsilon\big)T_{l'}\big(k^a(\varepsilon_i, k_j)\big)\Bigg)
\end{aligned}
\tag{14}
$$

$$- \left(1 - \frac{\widehat{\xi}(\varepsilon_i, k_j)}{\overline{\xi}}\right)$$
$$\times \left(-\lambda k^n(\varepsilon_i, k_j) + \beta \sum_{o=1}^{m_\varepsilon} \tau_o^\varepsilon \sum_{k'=1}^{n_\varepsilon} \sum_{l'=1}^{n_k} \theta'_{k'l'} T_{k'}\big(\rho_\varepsilon \varepsilon_i + \sigma_\varepsilon \omega_o^\varepsilon\big) T_{l'}\big(k^n(\varepsilon_i, k_j)\big)\right)\Bigg]$$

for the collocation nodes $i = 1, \dots, n_\varepsilon$ and $j = 1, \dots, n_k$ and where $\theta'_{k'l'}$ denotes next period's values of the value function coefficients. The optimal labor choice is defined through the first-order condition

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

The policy functions $k^a(\varepsilon_i, k_j)$, $k^n(\varepsilon_i, k_j)$, and $\widehat{\xi}(\varepsilon_i, k_j)$ are derived from the approximate value function as follows. First, the capital decision rule conditional on adjusting $k^a(\varepsilon_i, k_j)$ satisfies the first-order condition

$$0 = \mathbb{E}\left[\lambda - \beta \sum_{o=1}^{m_\varepsilon} \tau_o^\varepsilon \sum_{k'=1}^{n_\varepsilon} \sum_{l'=1}^{n_k} \theta'_{k'l'} T_{k'}\big(\rho_\varepsilon \varepsilon_i + \sigma_\varepsilon \omega_o^\varepsilon\big) T'_{l'}\big(k^a(\varepsilon_i, k_j)\big)\right], \tag{15}$$

where $T'_{l'}$ denotes the first derivative of the Chebyshev polynomial. Conditional on this unconstrained choice, the constrained capital decision is

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

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

$$\widetilde{\xi}(\varepsilon_i, k_j)$$
$$= \frac{1}{w\lambda}\Bigg[-\lambda\big(k^a(\varepsilon_i, k_j) - k^n(\varepsilon_i, k_j)\big)$$
$$+ \beta \sum_{o=1}^{m_\varepsilon} \tau_o^\varepsilon \sum_{k'=1}^{n_\varepsilon} \sum_{l'=1}^{n_k} \theta'_{k'l'} T_{k'}\big(\rho_\varepsilon \varepsilon_i + \sigma_\varepsilon \omega_o^\varepsilon\big)\Big(T_{l'}\big(k^a(\varepsilon_i, k_j)\big) - T_{l'}\big(k^n(\varepsilon_i, k_j)\big)\Big)\Bigg],$$

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

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

$$0 = \lambda - \left(\sum_{i=1}^{m_g} \tau_i^g \Big(e^z e^{\varepsilon_i} k_i^\theta n(\varepsilon_i, k_i) + (1-\delta)k_i\right. \tag{16}$$

$$- \frac{\widehat{\xi}(\varepsilon_i, k_i)}{\overline{\xi}} k^a(\varepsilon_i, k_i) - \left(1 - \frac{\widehat{\xi}(\varepsilon_i, k_i)}{\overline{\xi}}\right) k^n(\varepsilon_i, k_i)\Big) g(\varepsilon_i, k_i | \mathbf{m})\Big)^{-\sigma},$$

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

where

$$g(\varepsilon_i, k_j | \mathbf{m}) = g_0 \exp\Bigg\{ g_1^1 (\varepsilon_i - m_1^1) + g_1^2 (\log k_j - m_1^2) + \sum_{k=2}^{n_g} \prod_{l=0}^{k} g_k^l \big[(\varepsilon_i - m_1^1)^{k-l} (\log k_j - m_1^2)^l - m_k^l \big] \Bigg\}.$$

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

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

Consistency between the moments **m** and parameters **g** requires

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

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

$$0 = \mathbb{E}\big[z' - \rho_z z\big]. \tag{20}$$

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

A.2 *Solving for stationary equilibrium*

Following Hopenhayn and Rogerson (1993), I compute the steady state by solving for the wage w^* which sets labor supply equal to labor demand. Given a value of the wage w^*, I compute labor demand in the following three steps:

(i) 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.

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

(iii) Aggregate individual firms' labor demand according to $N_{\text{d}} = \int (n(\varepsilon, k) + \frac{\widehat{\xi}(\varepsilon,k)^2}{2\bar{\xi}}) \times g(\varepsilon, k)\, d\varepsilon\, dk$.

I then compute labor supply using the household's first order condition $N_{\text{s}} = (\frac{w^*\lambda^*}{\chi})^{1/\alpha}$. I solve for the market-clearing wage w^* using a root-finding algorithm. The marginal utility of consumption λ^* is computed from firm's behavior using $C^* = Y^* - I^*$.

For the comparisons to the histogram approximation in the main text, I follow the same steps, but approximate the distribution using a fine histogram as in Young (2010).

Appendix B: 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. In order to do this, I modify the benchmark model, because as Khan and Thomas (2008) show approximate aggregation holds in this model. To see this result, note that the distribution impacts firms' decisions through two channels: first, by determining the marginal utility of consumption $\lambda(z, g)$, and second, by determining the law of motion of the distribution $g'(z, g)$.[23] 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^2s of these forecasting equations are high, indicating that a Krusell and Smith (1998) algorithm using the aggregate capital stock performs well in this environment.

Den Haan (2010) notes that the R^2 is a poor 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 can hide occasionally large

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

Table 8. Forecasting With Aggregate Capital, Benchmark Model.

Forecasting Equation	$\mathbf{R}^2$	DH Statistic
Marginal utility	0.9993505	0.0085866%
Future capital	0.999968	0.0029358%

Note: Results from running forecasting regressions (21) on data simulated from first order model solution. "DH statistic" refers to Den Haan (2010)'s suggestion of iterating on the forecasting equations with updating K_t from simulated data. Expressed as a percent of mean marginal utility or log capital.

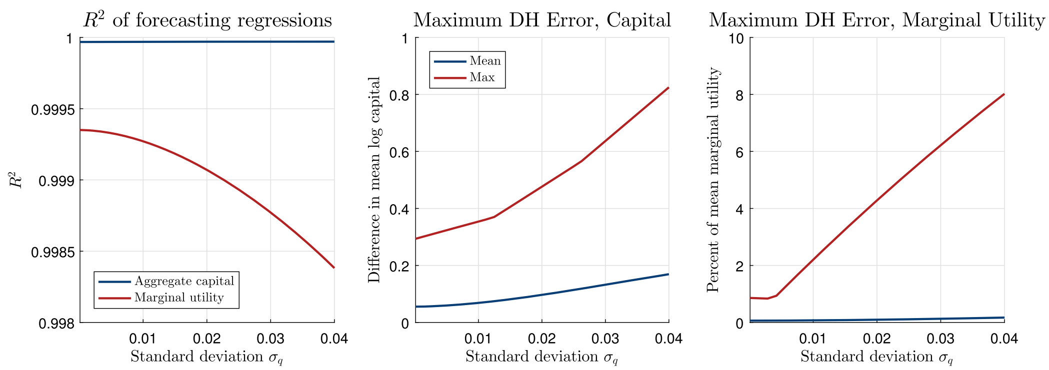

Figure 5. Forecasting power of aggregate capital, as a function of investment-specific shock variance. *Notes*: Results from running forecasting regressions (22) on data simulated from the first-order model solution. "DH statistic" refers to Den Haan (2010)'s suggestion of iterating on forecasting equations without updating K_t from simulated data.

errors. To address these concerns, Den Haan (2010) proposes iterating on the forecasting equations (22) without updating the capital stock, and computing the maximum deviations of these forecasts from the actual values in a simulation. Table 8 shows that this more stringent error metric is also small in the benchmark model.

To break this approximate aggregation result, I add an aggregate 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 productivity shock becomes more important. The left panel 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. The R^2 of these marginal utility forecast falls as the volatility of investment-specific shocks σ_q increase.

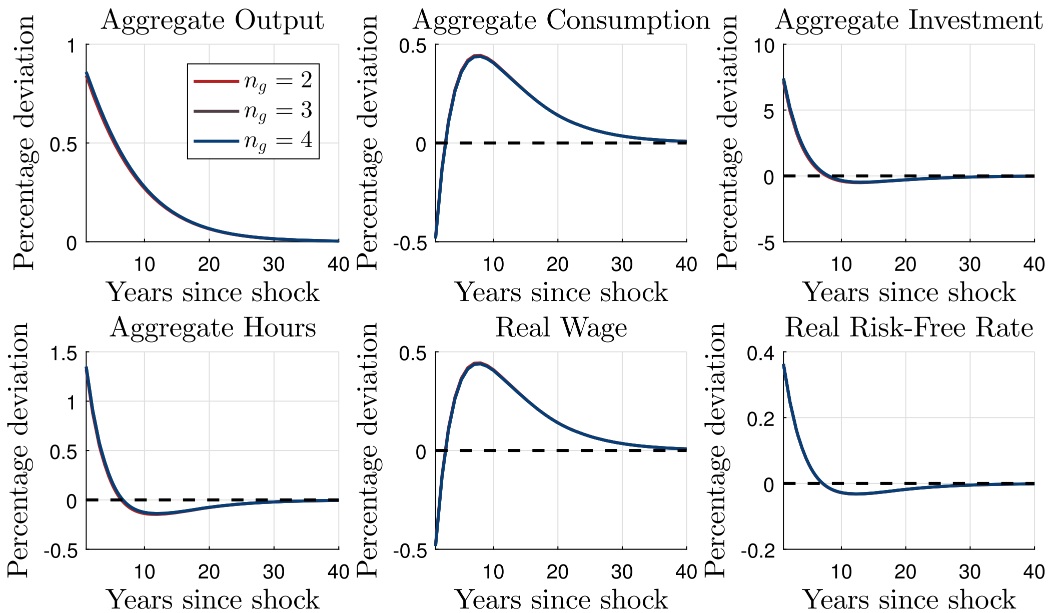

Figure 6. Aggregate impulse responses to investment-specific productivity shock. *Notes*: Impulse responses of aggregate variables, for different orders of approximation of the distribution. n_g refers to highest order moment used in parametric family (7).

The center and right panels of Figure 5 show that the more stringent Den Haan (2010) metrics grow even more sharply as a function of the volatility σ_q.

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 $\sigma_q = 0.014$, a value for which approximate aggregation fails. As with neutral productivity shocks in Figure 2, even relatively low degree approximations of the distribution are sufficient to capture the dynamics of these variables.

Appendix C: Estimation details

In this Appendix, I provide additional details of the estimation exercise described in Section 5 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 steady state in the model. The prior distributions of parameters are independent of each other and presented in Table 9. To sample from the posterior distribution, I use Markov Chain Monte Carlo with 10,000 draws, and drop the first 5000 draws as burn-in. Figure 7 plots the prior and estimated posterior distributions of parameters in the estimation with $\overline{\xi} = 1$.

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)

Note: Prior distributions for estimated parameters.

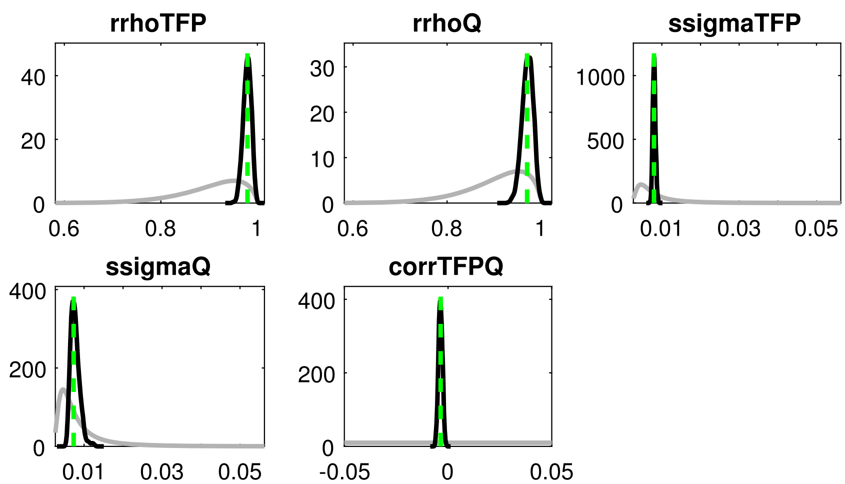

Figure 7. Estimated distribution of aggregate shock processes. *Notes*: Estimation results for $\overline{\xi} = 1$. Grey lines are the prior distribution of parameters. Dashed lines are the posterior mode. Black lines are the posterior distribution.

References

Adjemian, S., H. Bastani, M. Julliard, F. Karame, F. Mihoubi, G. Perendia, J. Pfeifer, M. Ratto, and S. Villemot (2011). Dynare working papers, 1, CEPREMAP. [1135]

Ahn, S., G. Kaplan, B. Moll, T. Winberry, and C. Wolf (2018), "When inequality matters for macro and macro matters for inequality." In *National Bureau of Economic Research Macroeconomics Annual*, Vol. 32 (M. Eichenbaum and J. A. Parker, eds.), 1–75, University of Chicago Press, Chicago. [1126]

Algan, Y., O. Allais, and W. D. Haan (2008), "Solving heterogeneous-agent models with parameterized cross-sectional distributions." *Journal of Economic Dynamics and Control*, 32, 875–908. [1125, 1130]

Auclert, A. (2017), "Monetary policy and the redistribution channel." NBER working paper 23451. [1123]

Bachmann, R., R. Caballero, and E. Engel (2013), "Aggregate implications of lumpy investment: New evidence and a DSGE model." *American Economic Journals: Macroeconomics*, 5, 29–67. [1123]

Berger, D. and J. Vavra (2015), "Consumption dynamics during recessions." *Econometrica*, 83, 101–154. [1123]

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. Terry (2018), "Really Uncertain Business Cycles." *Econometrica*, 86 (3), 1031–1065. [1139]

Brunnermeier, M. and Y. Sannikov (2014), "A macroeconomic model with a financial sector." *The American Economic Review*, 104 (2), 379–421. [1138]

Campbell, J. (1998), "Entry, exit, embodied technology, and business cycles." *Review of Economic Dynamics*, 1, 371–408. [1125]

Childers, D. (2018), "On the solution and application of rational expectations models with function-valued states." Working paper. [1125]

Clementi, G. L. and B. Palazzo (2016), "Entry, exit, firm dynamics, and aggregate fluctuations." *American Economic Journals: Macroeconomics*, 8, 1–41. [1123]

Cooper, R. and J. Haltiwanger (2006), "On the nature of capital adjustment costs." *The Review of Economic Studies*, 73, 611–633. [1129]

Den Haan, W. (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. [1146, 1147, 1148]

Devereux, M. and A. Sutherland (2011), "Country portfolios in open economy macro models." *Journal of the European Economic Association*, 9, 337–369. [1138]

Dotsey, M., R. King, and A. Wolman (1999), "State-dependent pricing and the general equilibrium dynamics of money and output." *Quarterly Journal of Economics*, 114 (2), 655–690. doi:10.1162/003355399556106. [1125]

Evans, D. (2015), "Optimal taxation with persistent idiosyncratic investment risk." Working paper. [1126]

Fernandez-Villaverde, J., J. Rubio-Ramirez, and F. Schorfheide (2016), "Solution and estimation methods for DSGE models." *Handbook of Macroeconomics*, 2, 526–724. [1141]

Hopenhayn, H. and R. Rogerson (1993), "Job turnover and policy evaluation: A general equilibrium analysis." *Journal of Political Economy*, 101, 915–938. [1133, 1146]

Judd, K. and S.-M. Guu (1997), "Asymptotic methods for aggregate growth models." *Journal of Economic Dynamics and Control*, 21, 1025–1042. [1133]

Kaplan, G., B. Moll, and G. L. Violante (2018), "Monetary policy according to HANK." *The American Economic Review*, 108 (3), 697–743. doi:10.1257/aer.20160042. [1123]

Khan, A. and J. Thomas (2013), "Credit shocks and aggregate fluctuations in an economy with production heterogeneity." *Journal of Political Economy*, 121, 1055–1107. [1123]

Khan, A. and J. K. Thomas (2008), "Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics." *Econometrica*, 76 (2), 395–436. [1124, 1126, 1127, 1129, 1135, 1138, 1139, 1140, 1146]

Koop, G., M. H. Pesaran, and S. Potter (1996), "Impulse response analysis in nonlinear multivariate models." *Journal of Econometrics*, 74, 119–147. [1139]

Krusell, P. and A. A. Smith (1998), "Income and wealth heterogeneity in the macroeconomy." *Journal of Political Economy*, 106 (5), 867–896. [1125, 1126, 1137, 1138, 1139, 1140, 1146]

Mertens, T. and K. Judd (2018), "Solving an incomplete markets model with a large cross-section of agents." *Journal of Economic Dynamics and Control*, 91, 349–368. [1126]

Ottonello, P. and T. Winberry (2017), "Financial heterogeneity and the investment channel of monetary policy." NBER working paper 24221. [1123]

Preston, B. and M. Roca (2007), "Incomplete markets, heterogeneity and macroeconomic dynamics." NBER Working Papers 13260, National Bureau of Economic Research, Inc. [1126]

Reiter, M. (2009), "Solving heterogeneous-agent models by projection and perturbation." *Journal of Economic Dynamics and Control*, 33 (3), 649–665. [1125, 1126]

Schmitt-Grohe, S. and M. Uribe (2004), "Solving dynamic general equilibrium models using a second-order approximation to the policy function." *Journal of Economic Dynamics and Control*, 28, 755–775. [1133, 1135]

Terry, S. (2017a), "Alternative methods for solving heterogeneous firm models." *Journal of Money, Credit and Banking*, 49, 1081–1111. [1137, 1139]

Terry, S. (2017b), "The macro impact of short-termism." Working paper. [1123, 1135]

Veracierto, M. (2002), "Plant level irreversible investment and equilibrium business cycles." *American Economic Review*, 92, 181–197. [1125]

Veracierto, M. (2017), "Adverse selection, risk sharing and business cycles." Chicago Fed working paper 2014-10. [1125]

Winberry, T. (2018), "Supplement to 'A method for solving and estimating heterogeneous agent macro models'." *Quantitative Economics Supplemental Material*, 86, https://doi.org/10.3982/QE740. [1123, 1124]

Young, E. (2010), "Solving the incomplete markets model with aggregate using the Krusell–Smith algorithm and non-stochastic simulations." *Journal of Economic Dynamics and Control*, 34, 36–41. [1133, 1146]

Co-editor Karl Schmedders handled this manuscript.

Manuscript received 27 June, 2016; final version accepted 15 January, 2018; available online 22 February, 2018.