Journal of Monetary Economics 158 (2026) 103895

Contents lists available at ScienceDirect

Journal of Monetary Economics

journal homepage: www.elsevier.com/locate/jme

# An endogenous gridpoint method for distributional dynamics[☆]

Christian Bayer[a,b,c], Ralph Luetticke[b,d,*], Maximilian Weiss[d], Yannik Winkelmann[d]

[a] Universität Bonn, Germany
[b] CEPR, France
[c] IZA, Germany
[d] Universität Tübingen, Germany

A R T I C L E  I N F O

*JEL classification:*
C46
C63
E32

*Keywords:*
Numerical methods
Distributions
Heterogeneous agent models
Linearization

A B S T R A C T

Modeling continuous choices in heterogeneous agent models as "lotteries" over a discretized state space is standard practice (Young, 2010), but renders the distributional dynamics linear in optimal policies. We present a novel, simple method that captures nonlinearities and solves the distributional dynamics with interpolation instead of integration using the idea of an endogenous grid. Our approach solves for a stationary equilibrium as quickly as the lottery method for a given precision, outperforms it for linear dynamics, and accommodates nonlinear dynamics and aggregate risk. We demonstrate its efficacy by studying a model with aggregate investment risk with a third-order perturbation solution.

## 1. Introduction

A large class of heterogeneous agent models has the evolution of the distribution of agents at its core. In this paper, we propose a novel method for implementing this evolution numerically. Our method exploits the fact that policy functions in many such models are monotone. This allows us to express the evolution of the distribution without relying on either linearized mappings (Young, 2010; Reiter, 2009) or full integration (Krusell and Smith, 1998). Instead, we extend the idea of endogenous gridpoints (Carroll, 2006) to distributional dynamics.

We show that our distributional endogenous gridpoint method, hereafter *DEGM*, is as fast and tractable as the "lottery method" proposed by Young (2010)—the standard in the literature. We also show that our method converges faster as the number of gridpoints increases, even when solving for a stationary equilibrium and studying linear dynamics. Importantly, it preserves nonlinearities and is thus suitable for higher-order perturbation solutions of macroeconomic models with heterogeneous agents.

We illustrate our method with two applications. We start with the Aiyagari (1994) economy and document the numerical efficiency gains over the lottery method when solving for stationary distributions. Both methods converge to the same solution as the number of gridpoints increases, but *DEGM* reaches this limit an order of magnitude faster. Our method works directly on the cumulative distribution function, parsimoniously capturing its curvature through shape preserving interpolation. Importantly, updating the distribution function is not costly because the novel endogenous gridpoint approach works without integration.

☆ We thank Anastasiia Antonova, Thomas Hintermaier, Herbert Koch and an anonymous referee for helpful comments. Christian Bayer gratefully acknowledges support by the DFG, Germany through CRC-TR224 EPoS, project C05, KFG50 CASFI, and in the framework of the German Excellence Strategy-EXC2126/1-390838866. Ralph Luetticke gratefully acknowledges funding by the European Union (ERC, AIRMAC, 101114991). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

* Corresponding author at: Universität Tübingen, Germany.

*E-mail addresses:* christian.bayer@uni-bonn.de (C. Bayer), ralph.luetticke@uni-tuebingen.de (R. Luetticke), maximilian.weiss@uni-tuebingen.de (M. Weiss), yannik.winkelmann@uni-tuebingen.de (Y. Winkelmann).

https://doi.org/10.1016/j.jmoneco.2026.103895
Received 29 May 2024; Received in revised form 9 January 2026; Accepted 9 January 2026

Available online 15 January 2026

We then propose a Krusell and Smith (1998) model with investment risk (depreciation shocks) as a new baseline model to study aggregate nonlinearities with household heterogeneity. This overcomes the approximate linearity in aggregate capital of the original Krusell and Smith (1998) model, while still being as parsimonious. We extend higher-order perturbation techniques to heterogeneous agent models, following Andreasen et al. (2018-01-01) and Levintal (2017-07). We solve our model up to third order and study asymmetric investment risk calibrated as in Barro (2006). The third-order solution becomes possible by combining the fast convergence of DEGM with state space reduction techniques developed in Bayer and Luetticke (2020), Bayer et al. (2024). For the second-order solution, we also solve the unreduced model and show that the reduction techniques yield identical results.

For first-order perturbations, *DEGM* gives the same solution as the lottery method in the limit, but again converges to the true impulse responses faster in terms of the number of gridpoints. However, there is a significant difference for higher-order perturbations, where the lottery method does not capture all nonlinear effects, as already argued by Bhandari et al. (2023). The lottery method *overstates*, in particular, the distributional responses to shocks. At the same time, it *understates* the average long-run increase in wealth inequality as a consequence of the presence of investment risk. Using a third-order perturbation solution with *DEGM*, we find that aggregate investment risk lowers the capital stock by 5 to 11 basis points and increases wealth inequality by up to 11 basis points, depending on the calibration of idiosyncratic income risk. Aggregate investment risk increases inequality by reducing the incentive to save, especially for poor households. The lottery method, by ignoring nonlinear terms in the distributional dynamics, predicts lower wealth inequality than *DEGM* in the presence of investment risk.

Similar to Angeletos (2007), the introduction of risky returns to savings reduces aggregate savings through a negative substitution effect that dominates over the income effect for the majority of households. Angeletos discusses these channels, but in a stylized model where all agents save the same proportion of their lifetime wealth but have different ex-post returns, so that the wealth distribution becomes non-stationary.[1] Since investment risk is not idiosyncratic in our model, the model does not generate enough inequality within the top 1%. However, we find that adding aggregate investment risk to the original Aiyagari economy does increase wealth inequality and in particular wealth holdings at the very top.

This finding complements recent work showing how the expectation of lower asset returns can increase wealth inequality through heterogeneous household portfolios (Fagereng et al., 2020; Gomez, 2024; Fernández-Villaverde and Levintal, 2024). While these studies emphasize the role of “passive” saving through rising asset prices, we show that aggregate uncertainty increases wealth inequality through “active” saving by the wealthy, as left-skewed investment risk has a stronger income effect for them.

Our method is directly implementable in the established sequence and state-space approaches for solving heterogeneous agent models with aggregate shocks (Auclert et al., 2021; Bayer et al., 2024). The idea of approximating the cumulative distribution function is not entirely new and can be found in Ríos-Rull (1997), Heer and Maussner (2009), and also in the special issue Den Haan et al. (2010). Our reformulation with an endogenous grid approach with nonlinear interpolation makes it differentiable, tractable and fast. The parsimonious representation of nonlinear distribution dynamics is key for higher-order perturbations.

The rest of the paper is organized as follows: Section 2 describes the distributional dynamics in terms of a difference equation of the distribution and policy functions, and presents our proposed method for solving this equation. Section 3 applies the method to the solution of an Aiyagari (1994) economy. Section 4 then uses an up to third order perturbation solution to the dynamic version of this economy with capital depreciation shocks as the source of aggregate risk. Section 5 concludes.

## 2. Problem and method

Consider an economy in discrete time with a distribution of agents (of mass 1) over two variables $a$ and $y$. We assume that $\ln(y)$ follows a stationary AR(1) process with normally distributed innovations $\epsilon$ and persistence $0 < \rho < 1$:

$$\ln(y_{t+1}) = \rho \ln(y_t) + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim N(0, \sigma_\epsilon^2). \tag{1}$$

The continuous endogenous variable $a$ is determined by the agent's policy function $a^*(a, y)$, which we assume to be strictly monotone in $a$ (or composed of a constant part and strictly monotone part) and, when discussing perturbation solutions, that is continuously differentiable both in $a$ and $y$ up to at least the order of perturbation.[2] The cumulative joint distribution in $a$ and $y$ at time $t$ is given by $F_t(a, y) := P(x_t \leq a, z_t \leq y)$. We denote $f_t(a \mid y) = f_t(a, y)/dF_t(y)$ as the conditional density of non-mass point $a$ for a given income $y$.[3] Because the process for $y$ is exogenous and stationary, we focus on the case where the marginal density in $y$ is the time-constant stationary one, $dF_t(y) = dF(y)$, which we know in closed form.

### 2.1. Distributional dynamics in discrete time

The evolution of the distribution $F$ is then given by the time-discrete Kolmogorov forward equation (making use of $x$ and $z'$ being independent):

$$F_{t+1}(a', y') = \int_{z' \leq y'} \int_z \int_{\{x \mid a' \geq a^*(x,z)\}} f_t(x \mid z)\, \phi(z, z') dx\, dz\, dz', \tag{2}$$

[1] Benhabib et al. (2024) also study the effects of investment risk in a model with heterogeneous agents. By making a distributional assumption about idiosyncratic investment risk, their model generates a wealth distribution with a realistic (Pareto-like) right tail.

[2] This is the standard case in many economic models, see Carroll (2006).

[3] We prove in Appendix C.3 that the distribution in $a$ is indeed continuous above the borrowing constraint.

where $\phi(z, z')$ is the density of a transition from today's income, $z$, to tomorrow's income $z'$, so that $\pi(z \mid z') := \frac{\phi(z,z')}{dF(z')}$ is the conditional density of having been at $z$ in $t$ conditional on being in $t+1$ at $z'$.

A brute-force approach to solving the equation for $F_{t+1}$ would therefore require an integral approximation. This is computationally expensive. Originally, economists often used Monte Carlo methods to solve Eq. (2). To avoid this, Young (2010) suggests replacing the continuous distribution in $a$ and $y$ with a discrete counterpart. One defines a grid for $a$ and a grid for $y$ and represents the $y$-process by a discrete Markov chain and replaces the policy function $a^*(a, y)$ by lotteries over the gridpoints closest to $a^*(a, y)$. Therefore, this method is commonly known as the "lottery" method.[4] We denote the discrete probabilities from this "lottery" method (henceforth: *LM*) by the vector $\hat{f}(a, y)$. The simultaneous transitions along the $a$ and $y$ dimensions can then be summarized by a single transition matrix $\mathbf{A}^*$. This gives the discretized and stacked Kolmogorov forward equation:

$$\hat{f}_{t+1} = \hat{f}_t \mathbf{A}^*. \tag{3}$$

While this allows to calculate $\hat{f}_{t+1}$ very quickly, it not only creates an approximation error due to discretizing continuous densities, but also forces Eq. (3) to be piecewise linear in the optimal policies $a^*$.

### 2.2. The endogenous gridpoint method for distributions (DEGM)

Instead, we propose to use an endogenous gridpoint method analogous to the one proposed by Carroll (2006). To do so, we write the problem in terms of the distribution in *a conditional* on $y$, technically $F_t(a \mid y) = \left(\frac{\partial}{\partial y} F_t(a, y)\right) / dF_t(y)$, and divide each period into two sub-periods, a first one related to asset choices and a second one related to income changes. Because $y$ is exogenous and strictly stationary, we can assume its marginal distribution to be constant over time and equal to its ergodic distribution $dF_t(y) = dF(y)$. This renders working with the conditional distribution or the joint distribution equivalent in the following, but also renders the former easier to handle. Concretely, we obtain the split in sub-periods by defining the distribution at the end of period $t$, after asset choices but before income transitions, as

$$\tilde{F}_t(a' \mid y) := \int_{\{x \mid a^*(x,y) \le a'\}} f_t(x \mid y) dx, \tag{4}$$

such that we obtain the asset distribution at the beginning of period $t+1$, given new income level $y'$, as

$$F_{t+1}(a' \mid y') = \int_z \tilde{F}_t(a' \mid z) \pi(z \mid y') dz. \tag{5}$$

Importantly, there are standard ways, such as Tauchen and Hussey (1991), to calculate the latter integral numerically in the form of a discretization of the $y$ process by creating a partitioning $\{\mathbb{Y}_1, \dots \mathbb{Y}_{n_y}\}$ and evaluating $\tilde{F}$ at the conditional expectations $\mathcal{Y}_k := E(y \mid y \in \mathbb{Y}_k)$. We can then replace the conditional transition densities by discrete probability counterparts:

$$F_{t+1}(a' \mid \mathcal{Y}_k) \approx \sum_{j=1}^{n_y} \tilde{F}_t(a' \mid \mathcal{Y}_j) \Pi_{j,k}; \quad \Pi_{j,k} := P(y_t \in \mathbb{Y}_j \mid y_{t+1} \in \mathbb{Y}_k). \tag{6}$$

These discretizations converge to the theoretical integral in (5) as $n_y$ increases.

In addition to our partitioning in $y$, we now specify grids $\{\mathcal{A}_i\}_{i=1\dots n_a}$ for $a$ and calculate the associated policies $\mathcal{A}^*_{i,j} := a^*\left(\mathcal{A}_i, \mathcal{Y}_j\right)$. With these objects at hand, we first consider the case where $a^*(\cdot, y)$ is strictly monotone everywhere. This assumption allows us to simplify, for the *endogenous gridpoints*, $\mathcal{A}^*_{i,j}$, the set over which we integrate to $\{x \mid a^*(x, \mathcal{Y}_j) \le \mathcal{A}^*_{i,j}\} = \{x \mid a^*(x, \mathcal{Y}_j) \le a^*\left(\mathcal{A}_i, \mathcal{Y}_j\right)\} = \{x \mid x \le \mathcal{A}_i\}$, where the last equation results from the monotonicity of $a^*$. This allows us to move from $F_t$ to $\tilde{F}_t$ without explicit integration for these endogenous gridpoints as

$$\tilde{F}_t(\mathcal{A}^*_{i,j} \mid \mathcal{Y}_j) = \int_{\{x \mid a^*(x,\mathcal{Y}_j) \le \mathcal{A}^*_{i,j}\}} f_t(x \mid \mathcal{Y}_j) dx = \int_{x \le \mathcal{A}_i} f_t(x \mid \mathcal{Y}_j) dx = F_t(\mathcal{A}_i \mid \mathcal{Y}_j).$$

Therefore, the set of tuples $\left\{\left(\mathcal{A}^*_{i,j}, F_t(\mathcal{A}_i \mid \mathcal{Y}_j)\right)\right\}_{i=1\dots n_a}$ is on the graph of $\tilde{F}_t(\cdot \mid \mathcal{Y}_j)$. Consequently, we can construct an interpolant $\hat{\tilde{F}}_t$ for each $\mathcal{Y}_j$, since $\{\mathcal{A}^*_{i,j}\}_{i=1\dots n_y}$ is an ordered set, as $a^*$ is strictly monotone in $a$. Replacing $\tilde{F}_t$ by $\hat{\tilde{F}}_t$ in (5), as well as replacing the integration by a weighted sum over the finite income grid, then allows to evaluate $F_{t+1}$ at any $a'$ without integration:

$$F_{t+1}(a' \mid \mathcal{Y}_k) \approx \sum_j \quad \hat{\tilde{F}}_t(a' \mid \mathcal{Y}_j) \quad \Pi_{j,k}. \tag{7}$$

Finally, the definition of a cumulative distribution function implies two extrapolation rules. First, $\hat{\tilde{F}}_t(a' \mid \mathcal{Y}_j)$ is set to zero for any $a' < \min_i \left\{\mathcal{A}^*_{i,j}\right\}$. These are future endogenous states that are lower than the smallest optimal policy and hence are never reached. Second, $\hat{\tilde{F}}_t(a' \mid \mathcal{Y}_j)$ is set to $\lim_{a\to\infty} F_t(a \mid \mathcal{Y}_j) = 1$ for all $a' > \max_i \left\{\mathcal{A}^*_{i,j}\right\}$. The largest optimal policy is lower than these future endogenous states and hence the probability to observe an agent with a lower than this endogenous state, given income $\mathcal{Y}_j$, is one.

**Handling the Borrowing Constraint.** When $a^*$ is a savings function, however, it is typically only weakly monotone. It has a constant initial part at the borrowing constraint and is strictly monotone, for a given $\mathcal{Y}_j$, for $a$ greater than some threshold of asset

[4] Sometimes this is also referred to as "histogram method". We will use the latter term, however, for a related but not identical interpretation of the distribution as piece-wise linear interpolants over bins with transitions being uniformly distributed from one bin to the other, see Reiter (2009).

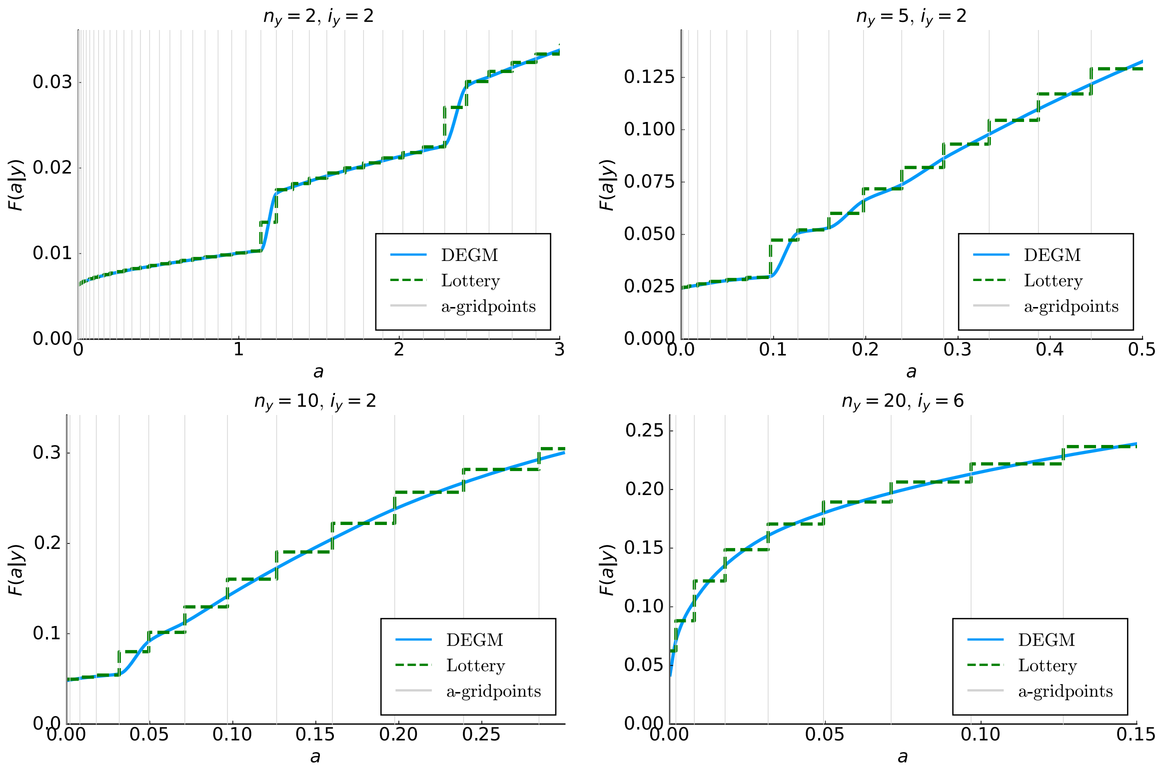



**Fig. 1.** Stationary distributions at selected income state.
*Notes*: Conditional CDFs ($F(a|y) = P(x \leq a, z = y)/P(z = y)$) for discrete approximations of the income process with $n_y = 2, 5, 10, 20$ states. We plot the income state $i_y$, where the policy maps away from the constraint for the first time. The CDF at this income state inherits the biggest "echoes". "Lottery" refers to the method which assumes point-masses (Young, 2010).

holdings $\underline{a}_j$. Our method can be easily adapted to account for this. Simply make the grid $\mathcal{A}_i$ start at $\underline{a}_j$, the EGM-solution (Carroll, 2006) corresponding to the borrowing constraint, to restore strict monotonicity. Because of weak monotonicity, and because we are working with cumulative distributions, evaluating $F_t(\underline{a}_j, \mathcal{Y}_j)$ gives the mass point at the borrowing constraint.

Another concern associated with a mass point at the borrowing constraint is that it can propagate through the distribution, especially when using a discrete approximation of the income process. Although there exist examples of policy functions that yield a distribution characterized by a finite number of mass points (see, e.g., Challe and Ragot (2016)), Appendix C demonstrates that this behavior is not generic. In fact, a discrete approximation of the income process produces a distribution function that is never truly "flat"—the set of points with nonzero mass is dense. The distribution function may still exhibit sharp increases—echoes of the borrowing constraint—which, however, diminish as the number of income states increases; see Fig. 1. This again highlights the importance of the assumption of a continuous income process. At the same time, Fig. 1 shows that a good approximation of (7) only requires a moderate number of income grid points. In this light, a smooth interpolant can be expected to approximate the true distribution more accurately than a gridded step function; even when such echoes occur, a sufficiently flexible smooth interpolant will still capture the "jump" effectively.

In the limit, the approximation converges to the true continuous income process, where these echoes are completely smoothed out. The approximated distribution function becomes differentiable everywhere in the interior, as does the true distribution function, which inherits this property from the differentiability of the policy function $a^*$. Importantly, this implies that an interpolant constructed from a finite set of income states approximates the wealth distribution well as we let the number of income states go to infinity, so that the numerical approximations to the integrals, $\Pi$, converge. In practice, this convergence with respect to the number of income gridpoints is fast, and as few as ten income gridpoints are sufficient to smooth the echoes of mass at the borrowing constraint.

### *2.3. Numerical implementation of DEGM*

To provide a practical guide to implementation, we conclude with a summary of the proposed algorithm, assuming that the dynamic programming problem leading to the policy function is solved w.l.o.g. on the grid $\{\mathcal{A}_i\}$. This means that $\left\{\mathcal{A}^*_{i,j}\right\}$ is readily available as a discretized representation of the policy function.

**Algorithm 1.** Start with the cumulative joint distribution (in $a$) at time $t$ given by $F_t(a \mid y)$ that is discretized on the grids $\{\mathcal{A}_i\}, \{\mathcal{Y}_j\}$, see Figure 2 (a). The following assumes the values of this are stored in the matrix $\mathbf{F}_t = \left[F_t(\mathcal{A}_i \mid \mathcal{Y}_j)\right]_i^j$.

1. For each exogenous state with index $j$, $y = \mathcal{Y}_j$, **create the interpolant** $\hat{\tilde{F}}_t^j(a)$.

   (a) Find the largest endogenous $\underline{a}_j$ s.t. $a^*(\underline{a}_j|y) = a_0$. This means, find the last asset state for which the policy $a^*$ is a constant (based on the end-of-period marginal value of $a_0$ and the budget constraint as in Carroll (2006)'s EGM).

   (b) Define

$$\bar{\mathcal{A}}_{i,j} = \begin{cases} \left\{\underline{a}_j\right\} \cup \left\{a \in \mathcal{A}_i \,:\, a > \underline{a}_j\right\} & \text{if } \underline{a}_j > a_0 \\ \mathcal{A}_i & \text{else} \end{cases}$$

   and the set of corresponding choices $\bar{\mathcal{A}}_{i,j}^*$

   (c) Find the set of according points on the graph of $\tilde{F}_t^j$: $G := \left\{\left(\bar{\mathcal{A}}_{i,j}^*, \mathbf{F}_t(i,j)\right)\right\}$, see Figure 2 (b).

   (d) Create an interpolant $\hat{\tilde{F}}_t^j$ based on the set $G$, see Figure 2 (c).

2. Loop through all $i, j$ to **evaluate the interpolant** for each $\mathcal{A}_i$ from the fixed grid $\{\mathcal{A}_i\}$ and each $\mathcal{Y}_j \in \{\mathcal{Y}_j\}$ to calculate:

$$\hat{\tilde{\mathbf{F}}}_t(i,j) = \begin{cases} 0 & \text{if } \mathcal{A}_i < \bar{\mathcal{A}}_{1,j}^* \\ \mathbf{F}_t(end,j) & \text{if } \mathcal{A}_i > \bar{\mathcal{A}}_{end,j}^* \\ \hat{\tilde{F}}_t^j(\mathcal{A}_i) & \text{else} \end{cases}$$

   and collect this in a matrix $\hat{\tilde{\mathbf{F}}}_t$. This yields the CDF in $a$ on the fixed grid $\{\mathcal{A}_i\}$ prior to the exogenous Markov transitions, see Figure 2 (d).

3. Apply the **exogenous** Markov **transition** matrix $\Pi$ to obtain $\mathbf{F}_{t+1}$ as:

$$\mathbf{F}_{t+1} = \hat{\tilde{\mathbf{F}}}_t \Pi'$$

Practical implementation requires the choice of an interpolation routine. Since cumulative distribution functions are monotone, the interpolant should preserve this property. Both linear interpolation and piecewise cubic hermitian splines do. However, the linear interpolant does not preserve differentiability everywhere. Note that the linear interpolant is neither equivalent to the histogram nor to the lottery method, since we interpolate the CDF with optimal policy choices being the interpolation nodes.[5]

### 2.4. Nonlinear distributional dynamics

Bhandari et al. (2023) have highlighted the fact that *LM* fails to fully capture the nonlinear dynamics of the distribution. Returning to Eq. (4), the nonlinear effects of the distribution $f_t$ on its transition dynamics derive entirely from its effect on the policy function $a^*$. Holding the policy function constant, the transition is a linear operator.[6]

We can thus sufficiently characterize the missing nonlinearities of the *LM* by analyzing the effects of changes in the optimal policy function. Let them be caused by some generic perturbation $\xi_t$, for example, aggregate shocks or changes in the mean of the distribution. The second-order derivative of the transition matrix $\mathbf{A}^*$ of the Kolmogorov forward equation to such a shock is generally composed of two terms and is given by

$$\frac{\partial \mathbf{A}^*(k,l)}{\partial a_k^*}\frac{\partial^2 a_k^*}{\partial \xi_t^2} + \frac{\partial^2 \mathbf{A}^*(k,l)}{\partial a^{*2}_k}\left[\frac{\partial a_k^*}{\partial \xi_t}\right]^2, \tag{8}$$

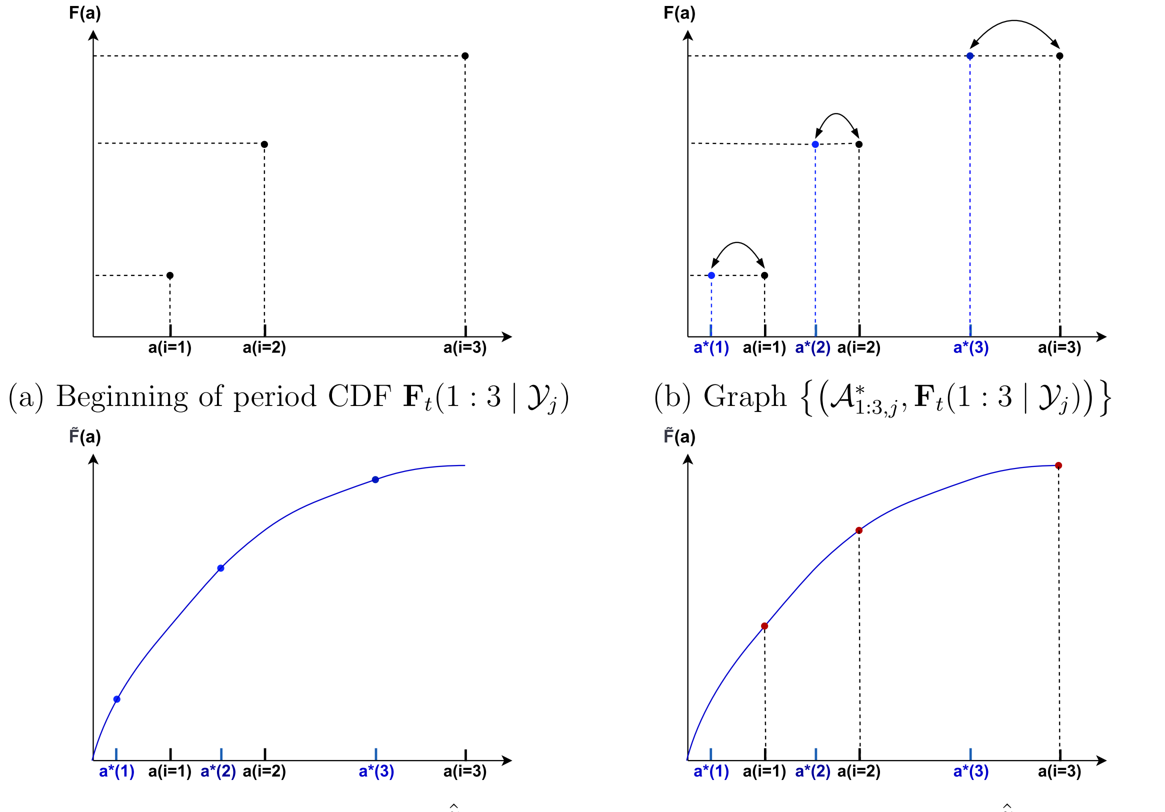
where $a_k^*$ denotes the optimal policy at wealth level $\mathcal{A}_k$. The first effect captures the direct nonlinearity of the policy function. The second effect reflects that the Kolmogorov forward equation is in principle nonlinear in policies. However, *LM* constructs $\mathbf{A}^*$ as (ignoring the exogenous state transitions for simplicity of notation)

$$\mathbf{A}^*(k,l) = \begin{cases} 1 - \frac{a_k^* - \mathcal{A}_l}{\mathcal{A}_{l+1} - \mathcal{A}_l} & \text{if } a_k^* \in [\mathcal{A}_l, \mathcal{A}_{l+1}) \\ \frac{a_k^* - \mathcal{A}_{l-1}}{\mathcal{A}_l - \mathcal{A}_{l-1}} & \text{if } a_k^* \in [\mathcal{A}_{l-1}, \mathcal{A}_l) \\ 0 & \text{else,} \end{cases} \tag{9}$$

which is piecewise linear in $a_k^*$. Therefore $\frac{\partial^2}{\partial a^{*2}_k}\mathbf{A}^*(k,l) = 0$ wherever it is defined. What is more, it is not defined (or infinite) on the grid points $\mathcal{A}_l$. Both properties of the lottery method create inaccuracies when perturbing to higher orders. In Appendix A, we extend this analysis to the third-order derivative of Eq. (4).

[5] See Appendix A.1 for details.

[6] The effect of the distribution on the policy function works through a market clearing condition, where higher aggregate demand for an asset, say, increases the market price.




(a) Beginning of period CDF $\mathbf{F}_t(1:3 \mid \mathcal{Y}_j)$ (b) Graph $\left\{\left(\mathcal{A}^*_{1:3,j}, \mathbf{F}_t(1:3 \mid \mathcal{Y}_j)\right)\right\}$

(c) Construction of interpolant $\hat{\tilde{F}}_t(\cdot \mid \mathcal{Y}_j)$ (d) Evaluation of interpolant $\hat{\tilde{\mathbf{F}}}_t(1:3 \mid \mathcal{Y}_j)$

**Fig. 2.** Illustration of DEGM with an interpolation over an endogenous grid.

Our method, on the other hand, can capture all nonlinearities up to the order of the splines used to interpolate the CDF. Again, the second-order derivative, now of our interpolant $\hat{\tilde{F}}_t^j(\mathcal{A}_i)$, has the general form:

$$\frac{\partial^2 \hat{\tilde{\mathbf{F}}}_t(i,j)}{\partial \xi_t^2} = \frac{\partial \hat{\tilde{\mathbf{F}}}_t(i,j)}{\partial \mathcal{A}^*_{.j}} \frac{\partial^2 \mathcal{A}^*_{.j}}{\partial \xi_t^2} + \left[\frac{\partial \mathcal{A}^*_{.j}}{\partial \xi_t}\right]' \frac{\partial^2 \hat{\tilde{\mathbf{F}}}_t(i,j)}{\partial {\mathcal{A}^*_{.j}}^2} \left[\frac{\partial \mathcal{A}^*_{.j}}{\partial \xi_t}\right], \tag{10}$$

where, unlike (8), the second term is nonzero because $\mathcal{A}^*_{.j}$ are the vectors of the interpolation nodes (and the derivatives are vector-valued). Therefore, the Hessian $\frac{\partial^2 \hat{\tilde{\mathbf{F}}}_t(i,j)}{\partial {\mathcal{A}^*_{.j}}^2}$ is generally nonzero. As also described in Bhandari et al. (2023), the second term in Eq. (10) reflects second-order responses of the distributional dynamics to first-order changes in the optimal policy. What is more, if the continuous distribution has *curvature* at these pre-images, $\mathcal{A}^*_{.j}$, approximation of $\frac{\partial^2 \hat{\tilde{\mathbf{F}}}_t(i,j)}{\partial {\mathcal{A}^*_{.j}}^2}$ requires a shape-preserving interpolation method, as illustrated in Fig. 2 using cubic splines.[7]

## 3. Solving for stationary distributions

Our first application is the solution of an Aiyagari (1994) economy, where Equation (2) takes the special form of $F_t = F\ \forall t$ as an equilibrium condition. Specifically, we consider an economy with a continuum of households facing idiosyncratic risk in their human capital, $h_t$, which they rent out to firms at the wage rate, $w_t$. Households can self-insure by accumulating non-negative amounts of capital, $k_t$, which they rent out to firms at rate $r_t$. Capital depreciates at the rate $\delta_t$. Human capital evolves according to a log-normal AR-1 process as in (1). We discretize this process in order to obtain the transitions in form of the matrix $\Pi$ using the Tauchen and Hussey (1991) algorithm. Households enjoy utility from consumption, $c_t$, and solve the dynamic program:

$$\max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} \mathbb{E} \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{11}$$

[7] Appendix A.3 shows that the continuous limit counterpart to $\frac{\partial^2 \hat{\tilde{\mathbf{F}}}_t(i,j)}{\partial {\mathcal{A}^*_{.j}}^2}$ is typically non-zero.

**Table 1**
Calibration.

| Parameters | | Value | Parameters | | Value |
|---|---|---|---|---|---|
| Calibration with persistent human capital process | | | | | |
| $\beta$ | Discount factor | 0.98 | $\sigma_\epsilon$ | Std. of log-income shocks | 0.14 |
| $\gamma$ | Rel. risk aversion | 1.00 | $\rho$ | Persistence of log-income | 0.98 |
| $\alpha$ | Capital share | 0.32 | $\delta$ | Depreciation rate | 0.02 |
| Implied wealth distribution | | | | | |
| Mass at $k = 0$ | | 0.04 | Wealth Gini | | 0.66 |
| Calibration with more transitory human capital process | | | | | |
| $\beta$ | Discount factor | 0.99 | $\sigma_\epsilon$ | Std. of log-income shocks | 0.18 |
| $\gamma$ | Rel. risk aversion | 1.00 | $\rho$ | Persistence of log-income | 0.88 |
| $\alpha$ | Capital share | 0.32 | $\delta$ | Depreciation rate | 0.02 |
| Implied wealth distribution | | | | | |
| Mass at $k = 0$ | | 0.01 | Wealth Gini | | 0.42 |
| Grid size used for calibration | | | | | |
| $n_k$ | Gridpoints for $k$ | 160 | $n_h$ | Gridpoints for $h$ | 20 |

$$\text{s.t. } c_t + k_{t+1} = \left(1 + r_t - \delta_t\right) k_t + h_t w_t \tag{12}$$

$$k_{t+1} \geq 0. \tag{13}$$

The wage and capital rates are given by the marginal products of labor and capital, respectively, where the production function is given by

$$Y_t = K_t^\alpha N^{1-\alpha}, \tag{14}$$

where $N$ is the total amount of human capital supplied by households.

We seek an equilibrium in which prices are constant such that households form optimal policies given $r, w, \delta$. These optimal policies are continuous choices of $k_{t+1}$. They depend on the continuous states $k_t$ (endogenous) and $h_t$ (exogenous). It is easy to show that for strictly concave felicity functions $u(c)$ the optimal policies are weakly monotone in $k_t$ and strictly monotone outside the borrowing constraint. The problem thus fits the setup of Section 2.

We use this workhorse model as a laboratory to present our novel method and compare it to the widely used lottery method (*LM*). To do so, we follow in principle the calibration idea of Den Haan et al. (2010), see Table 1, but deviate by having a continuous human capital process with sufficient persistence to generate a significant fraction of credit-constrained households to show that DEGM is robust to the challenges depicted in Fig. 1. Our baseline calibration features 4% of households exactly at the constraint.[8] It is important to note that this number is not identical to the share of hand-to-mouth households, defined as liquid assets below one-half of monthly labor income following (Weidner et al., 2014), see Appendix D.3 for details. For our baseline calibration this share is 16% (not reported in the table). We also consider a version with larger income shocks and lower persistence, closer to the original Den Haan et al. (2010) setup.

Conceptually, *DEGM* involves iterating the cumulative distribution function to find the stationary distribution.[9] We solve the model for $n_k = 160$ gridpoints in capital $k$ and $n_h = 20$ gridpoints in human capital h.[10] As a baseline, we use our novel method to find the equilibrium. Going beyond 160 gridpoints for capital and 20 for income had no significant effect on the equilibrium (using the *DEGM* method), so we consider the distribution at $n_k = 160$ and up to $n_h = 20$ to be the "true" distribution.[11]

We perform two exercises. First, we isolate the quality of the approximation to the distribution by keeping prices and optimal policies fixed at the benchmark solution for the stationary equilibrium, that is, we use the *DEGM* solution with $n_k = 160$. We then select a subset of gridpoints for the sparser grid and use the associated policies to iterate the distribution to convergence for both the established *LM* and our new *DEGM*.[12] Second, we solve for the stationary equilibrium, including prices and policies, which more closely resembles the actual use case. *LM* finds the stationary distribution via eigendecomposition, while our method uses iteration. For the first exercise, we use the uniform distribution as a starting guess, which we update for the second exercise in each iteration on the equilibrium prices with the last converged distribution.

Table 2 shows the distance of two moments of the stationary distribution, average capital holdings and the Gini coefficient of capital holdings, for $n_k = 40, 80$, and 160 gridpoints in the asset dimension relative to the baseline solution using *DEGM* with 160 gridpoints for capital. All deviations are expressed by varying the asset grid but keeping the income grid fixed. For the income dimension, we consider three variants with $n_h = 5, 10$, and 20 gridpoints.

[8] We find that the lottery method overstates the probability mass at the borrowing constraint, in particular for small $n_k$. In Appendix D.3, we discuss this further.

[9] We compute the aggregate capital stock as $E[X] = b \cdot F(b) - a \cdot F(a) - \int_a^b F(x)\,dx$.

[10] The grid is defined as $k_i = k_{min} + u_i^2$ with $u_i$ uniformly distributed on $[0, \sqrt{200}]$

[11] Beyond 160 gridpoints, the *DEGM* solution no longer changes. For *LM*, convergence is achieved at about 320 gridpoints. Then there is no difference between the two methods in the solution of the stationary equilibrium. Also, going beyond 20 gridpoints for human capital does not significantly change the results.

[12] We use piecewise cubic Hermite splines to interpolate the cumulative distribution function.

**Table 2**
Convergence of stationary equilibria under *LM* and *DEGM* in $n_h, n_k$.

| | $n_h$ \ $n_k$ | LM 40 | LM 80 | LM 160 | DEGM 40 | DEGM 80 |
|---|---|---|---|---|---|---|
| **Panel A**: Stationary distribution [relative deviations in percent] | | | | | | |
| Capital stock | 5 | 1.72 | 0.56 | 0.12 | 0.09 | 0.03 |
| | 10 | 1.21 | 0.43 | 0.10 | 0.12 | 0.03 |
| | 20 | 0.85 | 0.32 | 0.09 | 0.17 | 0.03 |
| Wealth Gini | 5 | 2.24 | 0.78 | 0.15 | 0.00 | −0.01 |
| | 10 | 1.30 | 0.45 | 0.11 | 0.03 | −0.00 |
| | 20 | 0.84 | 0.30 | 0.08 | −0.02 | −0.00 |
| **Panel B**: Stationary equilibrium [relative deviations in percent] | | | | | | |
| Capital stock | 5 | 0.31 | 0.11 | 0.03 | −0.03 | −0.00 |
| | 10 | 0.20 | 0.07 | 0.02 | −0.01 | −0.00 |
| | 20 | 0.12 | 0.05 | 0.02 | 0.01 | −0.00 |
| Wealth Gini | 5 | 2.66 | 0.83 | 0.17 | 0.20 | 0.02 |
| | 10 | 1.60 | 0.51 | 0.13 | 0.17 | 0.02 |
| | 20 | 1.09 | 0.36 | 0.09 | 0.11 | 0.02 |
| **Panel C**: Computation times [in s] | | | | | | |
| Time (s) | 5 | 0.12 | 0.20 | 0.29 | 0.19 | 0.32 |
| | 10 | 0.22 | 0.35 | 0.59 | 0.37 | 0.60 |
| | 20 | 0.45 | 0.71 | 1.49 | 0.82 | 1.39 |

*Notes*: For each row, values represent percent deviations of the solutions with $n_k$ gridpoints to the reference solution (*DEGM* with $n_k = 160$). For *LM* we use discrete aggregation methods, while we use continuous integration methods for *DEGM*. Values for the "persistent" income calibration from Table 1.
**Panel A:** Calculating stationary distribution using policies from the reference solution.
**Panel B:** Solving the stationary equilibrium including prices and policies.
**Panel C:** Time in seconds for solutions of Panel B. CPU with 16-cores, 3.3 GHz.

Panel A does this for the first exercise with constant prices and policies. Panel B compares the two methods for the second stationary equilibrium exercise. Regardless of the size of the income grid $n_h$, we find that our method converges to the "true" distribution much faster, especially for cross-sectional moments.[13] For a given number of asset gridpoints, *LM* is faster in terms of computational time, mainly because it does not require iterations when updating the distribution. However, for a given accuracy ($n_k = 40$ *DEGM* $\approx n_k = 160$ *LM*), our method is faster in solving for the stationary equilibrium (Panel C). The faster convergence reflects the fact that the distribution function is typically nowhere flat and the set of asset states that can be reached is dense, as discussed in Appendix C. *LM* approximates this necessarily by a step function for the distribution and for a coarse grid this approximation lacks precision.

## 4. Higher-order perturbations of distributional dynamics

Our second application is a setup with aggregate risk. As explained in Section 2, our method is able to capture nonlinearities in such setups. Specifically, we study an up to third-order perturbation solution of the Aiyagari model outlined above with capital depreciation shocks. Since we use third-order splines to interpolate the distribution, our method captures all nonlinear effects in both the distribution and the policy function.

To do this, we extend the state-space perturbation techniques for heterogeneous agent models to higher orders. Following Reiter (2009), we include the distribution and value functions in the state space. We then define a nonlinear differential equation on these objects, see Appendix B. The higher-order perturbation solution of this difference equation is then the same as for any differential equation reflecting a state-space system. Therefore, we can rely on established methods and solve the system with the algorithms developed by Andreasen et al. (2018-01-01) and Levintal (2017-07).[14] We also repeat the details of Levintal (2017-07)'s algorithm applied to heterogeneous agent models in Appendix B.

In addition, we show for the second-order perturbation that the Bayer and Luetticke (2020) reduction and its refinement in Bayer et al. (2024) yield the same results as solving the unreduced Reiter (2009) system. This reduces the state space of the model by writing the distribution in the form of a copula and marginals, and representing the value functions as sparse combinations of basis functions. For the third-order perturbation, we then exploit the sparseness of the reduced system. Currently, it is not feasible to solve the full system beyond second order due to the increased memory requirements.

Solving the model up to third order allows us to implement asymmetric shocks. We follow Levintal in approximating the binomial distribution of a depreciation shock with a continuous distribution that has the same higher-order moments. Agents internalize

[13] This mirrors the findings in Den Haan et al. (2010), which compares the approximation of the Kolmogorov forward equation by Monte Carlo simulation, the lottery method, or direct integration using a spline for the cumulative distribution function.

[14] The calculation of the first moments of the higher-order approximations using the Andreasen et al. (2018-01-01) method is equivalent to the method developed earlier by Rudebusch and Swanson (2012).

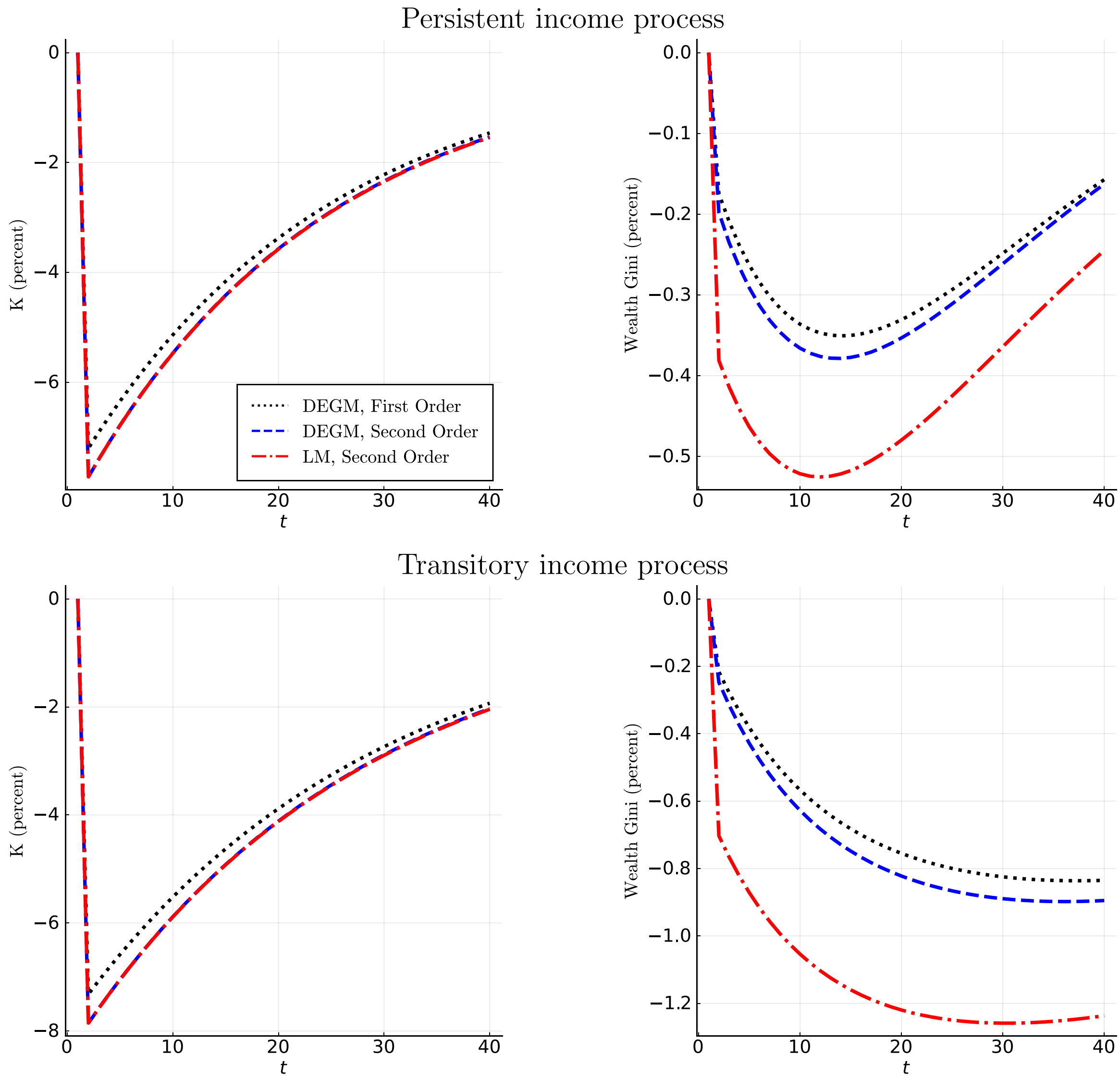



**Fig. 3.** Impulse responses to a capital depreciation shock.
*Notes*: Impulse responses of capital (left panels) and wealth Gini (right panels) to the shock $\nu_1 = 7.5$ *p.p.* from first-order (dotted black) and second-order (dashed blue) solutions using *DEGM*, and from second-order solution using *LM* (dash-dotted red), evaluated at the non-stochastic steady state ($n_y = 10, n_k = 160$). Top row: persistent calibration. Bottom row: transitory calibration. Y axis: Percent deviation from the non-stochastic steady state. X-axis: Quarter.

the positive third moment of the shock, which acts as an investment risk. In particular, capital depreciation $\delta_t$ deviates from its steady-state value by following the process:

$$\delta_t = \delta + \nu_t,\ \nu_t \sim F^{\nu}(0, \sigma_\delta, \tau_\delta), \tag{15}$$

where $\sigma_\delta^2$ and $\tau_\delta^3$ are the second and third moments of the depreciation shock distribution, respectively. In our calibration, $\sigma_\delta = 0.005$ and $\tau_\delta = 0.012$, which corresponds to a 0.4% chance that a disaster destroys 7.5% of the capital stock in a given period (quarter) and causes a 10% drop in annual GDP, consistent with the evidence in Barro (2006).

We propose this Krusell and Smith (1998)-style model with investment risk as a baseline model for studying aggregate nonlinearities with household heterogeneity. It overcomes the approximate linearity in aggregate capital of the original Krusell and Smith (1998) model while being equally parsimonious. Fig. 3 compares the impulse responses to a one-time 7.5% destruction of the capital stock using *DEGM* to compute the dynamics with first- and second-order perturbation solutions. The first-order solution slightly understates the decline in aggregate capital and also the decline in the Gini coefficient of wealth in response to the capital depreciation shock. The difference in the second-order solution across methods is small for the response of the aggregate capital

**Table 3**
Convergence of 1st- and 2nd-order perturbation for *LM* and *DEGM*.

| | $n_k$ / $n_h$ | *LM* | | | *DEGM* | |
|---|---|---|---|---|---|---|
| | | **40** | **80** | **160** | **40** | **80** |
| **Panel A**: IRF statistic as in Bayer et al. (2024) | | | | | | |
| Capital stock (FO) | 5 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 10 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Capital stock (SO) | 5 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 10 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wealth Gini (FO) | 5 | 0.89 | 0.99 | 1.00 | 1.00 | 1.00 |
| | 10 | 0.97 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wealth Gini (SO) | 5 | 0.52 | 0.37 | 0.23 | 0.99 | 1.00 |
| | 10 | 0.91 | 0.85 | 0.79 | 1.00 | 1.00 |
| **Panel B**: Second-order moments [relative deviations in basis points] | | | | | | |
| Capital stock | 5 | −0.47 | −0.46 | 1.68 | −0.09 | −0.03 |
| | 10 | −1.08 | −0.79 | 1.38 | −0.96 | −0.07 |
| Wealth Gini | 5 | −3.94 | −4.18 | −10.35 | 0.50 | 0.12 |
| | 10 | −0.53 | −1.32 | −8.59 | 2.32 | −0.25 |

*Notes*: In each row, the values represent the basis point deviations of the $n_k$ gridpoints solutions from the reference (*DEGM* with $n_k = 160$). Parameters of the "persistent" calibration, Table 1.
**Panel A** shows the $R^2$-like statistics from Bayer et al. (2024) for an impulse response following a 7.5 p.p. shock to $\delta$ (over 100 periods) with a first order (FO) and second order (SO) perturbation solution. The $R^2$-like statistic is $1 - \sum_{h=1}^{H} \frac{(IRF_{DEGM,n_k=160,n_h}(h) - IRF_{method,n_k,n_h}(h))^2}{\sum_{h=1}^{H} IRF_{DEGM,n_k=160,n_h}(h)^2}$.
**Panel B** shows the ergodic moment solving the model with a second order perturbation. The deviation shown is the difference in basis point deviation from each steady state over the basis point deviation in the baseline solution (*DEGM*, $n_k = 160$). The aggregates under *LM* are derived using discrete aggregation methods, while continuous integration methods are used for *DEGM*.

stock, but significant for the response of the Gini coefficient. This is true for both income processes considered. Taken together, this suggests that the distributional dynamics in this model are nonlinear with respect to aggregate shocks, but that the feedback from inequality to aggregates and equilibrium prices is modest.

Table 3 compares the solutions (without state-space reduction) for different grid sizes. *LM* and *DEGM* converge to the same first-order perturbation solution, just as they converge to the same stationary equilibrium (consistent with Bhandari et al., 2023, showing that *LM* has no bias up to first-order perturbations). In terms of IRFs of aggregates, convergence is fast in $n_k$ for both methods. In terms of the first-order dynamics of the wealth distribution, we find no significant differences in the $n_k = 40$ solution with DEGM to the 160-grid-point benchmark. For *LM*, the dynamics of the wealth distribution, in terms of IRFs of the wealth Gini, become the same as in our baseline only when we use $n_k = 160$ gridpoints. In other words, *DEGM* converges to the true solution much faster for the first-order solution than *LM*. For the second-order solution, as expected, we find no convergence of the *LM* to the *DEGM* benchmark solution, neither in terms of ergodic means nor in terms of distributional IRFs.

The computing time (not reported) is now nearly identical for both methods for the same number of gridpoints. Both methods require one iteration of the Kolmogorov forward equation when solving for the perturbation solution, and these differ slightly, but compared to the total computational cost of solving the difference equation, this one iteration has a negligible computational cost so that variations therein are irrelevant. This gives *DEGM* a clear advantage in solving for aggregate dynamics over *LM*, because it requires less gridpoints for the same precision and computing time increases sharply with the number of gridpoints.

All of the above results do not use dimensionality reduction. Next, we check whether the Bayer and Luetticke (2020) method, which reduces the dimensionality of the difference equation, can be extended to higher-order perturbations. The method replaces the distribution function by a copula and marginals, and considers deviations from the steady state in copula and value functions only on a coarse grid/based on discrete cosine transformations with linear interpolations in between. The extension to higher-order perturbations and the use of *DEGM* require that the linear interpolants used by Bayer and Luetticke (2020) to represent copulas and value functions are also replaced by spline interpolation. Table 4 evaluates the quality of the approximation introduced by the reduction step. We compare four variants. First, the no-reduction version ("None"), which defines both the distribution and the value functions on the full Cartesian product of the income and wealth grids.[15] Second, a version that reduces only the distribution function by writing it in terms of marginal distributions (defined on the full income and wealth grids) and a copula defined on a coarse wealth percentile grid. Third, we also use the DCT method of Bayer and Luetticke (2020) to reduce the number of controls representing the value function. Fourth, we apply the refinement developed in Bayer et al. (2024). This solves the model first at first order and finds a factor representation of the copula function based on its variance–covariance matrix obtained under the first-order representation.[16]

Panel A of Table 4 lists the number of states and controls, and hence the dimensionality, of the difference equation describing the macroeconomic model, see Appendix B. The maximum reduction removes 75% of all states and controls. This has no visible effect

[15] In practice, since we use an EGM to solve for optimal policies, we write the difference equation in terms of policy functions instead of value functions.
[16] This requires in practice a sufficiently rich set of shocks such that all prices that vary in the higher-order solution are also varying in the first-order solution. Here we achieve this by having TFP and also time-preference shocks that drive a wedge between the interest rate and the ratio of expected marginal utilities.

**Table 4**
Comparison of second-order solution with different state-space reductions.

| Reduction | Persistent None | Persistent Copula | Persistent Copula +DCT | Persistent Copula +DCT +Factor | Transitory None | Transitory Copula | Transitory Copula +DCT | Transitory Copula +DCT +Factor |
|---|---|---|---|---|---|---|---|---|
| **Panel A:** Dimensions | | | | | | | | |
| States | 403 | 214 | 214 | 109 | 403 | 214 | 214 | 112 |
| Controls | 412 | 412 | 98 | 98 | 412 | 412 | 95 | 95 |
| Total | 815 | 626 | 312 | 207 | 815 | 626 | 309 | 207 |
| **Panel B**: IRF statistic as in Bayer et al. (2024) | | | | | | | | |
| Capital stock | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wealth Gini | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| **Panel C:** Second-order moments [levels] | | | | | | | | |
| Capital stock | 25.54 | 25.54 | 25.54 | 25.54 | 25.29 | 25.29 | 25.29 | 25.29 |
| Wealth Gini | 0.61 | 0.61 | 0.61 | 0.61 | 0.43 | 0.43 | 0.43 | 0.43 |

*Notes*: "None" shows the results without reduction, "Copula" reduces the states for the distribution function, "Copula + DCT" additionally reduces the controls for the value function using the steady-state reduction approach of Bayer and Luetticke (2020), "Copula + DCT + Factor" derives a factor representation of the copula function from the first-order solution as in Bayer et al. (2024) and further reduces the dimensionality of the sparse set of basis functions to represent the copula, see Appendix B. The full state space refers to $n_k = 40$, $n_y = 10$ plus aggregate states and controls.
**Panel A** shows the number of states and controls that enter the difference equation representing the macroeconomic model and hence its dimensionality.
**Panel B** shows the $R^2$-like statistic from Bayer et al. (2024) for an impulse response following a 7.5 p.p. shock to $\delta$ (over 100 periods) with second-order (SO) perturbation solution. The $R^2$-like statistic is $1 - \sum_{h=1}^{H} \frac{(IRF_{Full}(h) - IRF_{Reduced}(h))^2}{\sum_{h=1}^{H} IRF_{Full}(h)^2}$.
**Panel C** shows ergodic moments in levels for the second-order solution.

**Table 5**
Ergodic moments under second- and third-order solution with investment risk.

| Variable | Persistent *LM* Mean | Persistent *LM* (SD) | Persistent *DEGM* Mean | Persistent *DEGM* (SD) | Transitory *LM* Mean | Transitory *LM* (SD) | Transitory *DEGM* Mean | Transitory *DEGM* (SD) |
|---|---|---|---|---|---|---|---|---|
| **Panel A**: Second-order solution [relative deviations in basis points] | | | | | | | | |
| Output | −2.7 | (60.5) | −2.7 | (62.6) | −3.1 | (62.8) | −2.9 | (63.5) |
| Capital stock | −8.4 | (189.0) | −8.4 | (195.6) | −9.7 | (196.2) | −9.2 | (198.3) |
| Wealth Gini | −2.6 | (19.0) | 0.6 | (15.3) | −18.2 | (76.4) | 1.1 | (55.0) |
| **Panel B**: Third-order solution [relative deviations in basis points] | | | | | | | | |
| Output | −3.6 | (-) | −1.5 | (-) | 4.8 | (-) | −3.7 | (-) |
| Capital stock | −11.2 | (-) | −4.7 | (-) | 15.0 | (-) | −11.4 | (-) |
| Wealth Gini | 2.0 | (-) | 10.7 | (-) | −149.7 | (-) | 2.1 | (-) |

*Notes*: $n_k = 40$, $n_y = 10$. Means and standard deviations (in brackets) across parameterization and methods are in basis point deviation from non-stochastic steady state. Moments are from closed-form solutions for pruned model dynamics (Andreasen et al., 2018-01-01). No standard deviations are reported for third-order perturbations because of memory requirements. Capital depreciation shock with $\sigma_\delta = 0.5\%$, $\tau_\delta = 1.2\%$.

on the second-order IRFs, Panel B. Also, the ergodic moments hardly change, Panel C. The quality of the reduction is independent of whether the income process is persistent or transitory. This is an important result because the number of derivatives to compute, and hence the size of the matrices to store, scales with the number of states and controls to the power of the order of the perturbation solution plus one (quadratic for first order, cubic for second order, etc.).

For this reason, a third-order solution to the full model is not feasible on a machine with less than 20 terabytes of memory.[17] However, the strong reduction allows us to solve the reduced system at third order (still requiring almost 2 terabytes of memory). We use the fact that the reduction is practically lossless at the second order as a heuristic to expect a reasonable quality of approximation also at the third order.

Higher-order solutions not only provide a better approximation of the dynamics, but also, importantly, capture the response of households to aggregate risk. Table 5 documents how the ergodic distribution with aggregate risk differs from the stationary equilibrium without aggregate risk. Aggregate risk here refers to investment risk. The top panel shows the change for a second-order perturbation, the bottom panel for a third-order perturbation. Appendix A discusses that the third-order solution captures the effect of the skewed distribution of aggregate shocks, while the second-order solution captures only the effect of the variance. As in Angeletos (2007), we find that aggregate investment risk reduces the aggregate capital stock because for most households the substitution effect is stronger than the income effect. For the second-order perturbation, we find small effects on aggregates, and

[17] We estimate this memory requirement by extrapolating the relationship between system size and memory requirements for smaller models. We predict that 20 terabytes of memory will be required for the third-order solution of the full model at the lowest resolution we consider, $n_k = 40, n_y = 5$.

the distributional nonlinearities that *LM* misses seem to be of little importance for the first two central moments of the ergodic distribution, which is in line with the similarities of the IRFs documented before in Fig. 3.

For the wealth distribution itself, in line with Bhandari et al. (2023), we find that the two methods are already for the second-order perturbation qualitatively different (even though differences are small). *LM* always predicts a decrease in wealth inequality while *DEGM*, which avoids the Bhandari et al. (2023) criticism, predicts that inequality might increase in response to aggregate risk compared to the stationary equilibrium. The underaccumulation and increase in wealth inequality are driven by lower savings of less wealthy households. For these households, the substitution effect dominates as they have little capital income. Moreover, a lower capital stock implies a lower wage rate and a higher rate of return on capital. For wealthy households, however, the income effect is key and they have strong precautionary saving motives given the aggregate investment risk. Thus, while a capital depreciation shock upon realization compresses the distribution of wealth,[18] as can be seen in Fig. 3, the risk of such a shock can increase wealth inequality on average.

At higher orders, the interaction of income and substitution effects starts to play a role: how is the savings response to a change in income altered by a simultaneous change in investment opportunities? This is relevant when considering the risk of capital destruction, which induces a negative correlation between today's capital income, and future returns on investment. However, this correlation does not only matter for the individual household's optimization problem, but also for the interaction of optimal policies in the *cross-section*. As we show in Appendix A.2, *LM* misses several of the higher-order terms when computing IRFs and the ergodic distribution. The reason is that the higher-order effect of a change in the savings policy on the wealth distribution in $t+1$ depends on the curvature of the wealth distribution in $t$ (in Appendix A.3 we derive this analytically). Since *LM* models a discrete distribution, it only accounts for the *level* of the "density" at a specific point, namely the size of the mass point. In contrast, *DEGM* models the continuous limit of the wealth distribution. Thereby, it accounts for *marginal* transition flows across wealth levels. The marginal flows are characterized by the slope and the curvature of the density.[19]

## 5. Conclusion

We propose a novel endogenous gridpoint method for distributional dynamics (*DEGM*). Our method retains the tractability and speed of the lottery/histogram methods commonly used in the literature, while requiring significantly fewer gridpoints and capturing all nonlinear effects of distributional dynamics. By preserving the nonlinearities critical to heterogeneous agent models, *DEGM* provides an improved framework for studying models with household heterogeneity and aggregate risk. It allows for a straightforward implementation in the established sequence and state-space approaches for solving heterogeneous agent models with aggregate shocks (Auclert et al., 2021; Bayer et al., 2024). We provide an example of a state-space solution with a third-order perturbation. In particular, we propose a Krusell and Smith (1998) model with aggregate investment risk. In this model, we show that aggregate investment risk affects inequality.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmoneco.2026.103895.

## Data availability

No data was used for the research described in the article.

## References

Aiyagari, S.R., 1994. Uninsured idiosyncratic risk and aggregate saving. Q. J. Econ. 109 (3), 659–684.
Andreasen, M.M., Fernández-Villaverde, J., Rubio-Ramírez, J.F., 2018-01-01. The pruned state-space system for non-linear DSGE models: Theory and empirical applications. Rev. Econ. Stud. 85 (1), 1–49.
Angeletos, G.-M., 2007. Uninsured idiosyncratic investment risk and aggregate saving. Rev. Econ. Dyn. 10 (1), 1–30.
Auclert, A., Bardóczy, B., Rognlie, M., Straub, L., 2021. Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models. Econometrica 89 (5), 2375–2408.
Barro, R.J., 2006. Rare disasters and asset markets in the twentieth century*. Q. J. Econ. 121 (3), 823–866.
Bayer, C., Born, B., Luetticke, R., 2024. Shocks, frictions, and inequality in US business cycles. Am. Econ. Rev. 114 (5), 1211–1247. http://dx.doi.org/10.1257/aer.20201875.
Bayer, C., Luetticke, R., 2020. Solving discrete time heterogeneous agent models with aggregate risk and many idiosyncratic states by perturbation. Quant. Econ. 11 (4), 1253–1288.
Benhabib, J., Cui, W., Miao, J., 2024. Capital income jumps and wealth distribution. Mimeo, UCL.
Bhandari, A., Bourany, T., Evans, D., Golosov, M., 2023. A perturbational approach for approximating heterogeneous agent models. NBER Working Paper No 31744.

[18] Luetticke et al. (2025) find that wars are best characterized by capital destruction shocks and do compress inequality.

[19] For example, at a wealth level $a$ where the density $f(a)$ is falling and convex, the marginal outflow when transitioning to other points on the distribution is dominated by households with wealth slightly below $a$. This means that the substitution effect is *marginally* more important for the change in distribution induced by the outflows from point $a$, if the substitution effect is more important for the poorer households, while conversely the income effect dominates for richer households.

Carroll, C., 2006. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. Econom. Lett. 91 (3), 312–320.

Challe, E., Ragot, X., 2016. Precautionary saving over the business cycle. Econ. J. 126 (590), 135–164.

Den Haan, W.J., Judd, K.L., Juillard, M., 2010. Computational suite of models with heterogeneous agents: Incomplete markets and aggregate uncertainty. J. Econom. Dynam. Control 34 (1), 1–3.

Fagereng, A., Holm, M., Moll, B., Natvik, G., 2020. Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains. CEPR Discussion Paper, (14355).

Fernández-Villaverde, J., Levintal, O., 2024. The distributional effects of asset returns. Mimeo University of Pennsylvania and Reichman University..

Gomez, M., 2024. Wealth inequality and asset prices. Mimeo Columbia University.

Heer, B., Maussner, A., 2009. Dynamic General Equilibrium Modeling. Springer.

Krusell, P., Smith, A.A., 1998. Income and wealth heterogeneity in the macroeconomy. J. Political Econ. 106 (5), 867–896.

Levintal, O., 2017-07. Fifth-order perturbation solution to DSGE models. J. Econom. Dynam. Control 80, 1–16.

Luetticke, R., Meyer, T., Müller, G., Schularick, M., 2025. The Great Leveler According to HANK. CEPR Discussion Papers DP20943.

Reiter, M., 2009. Solving heterogeneous-agent models by projection and perturbation. J. Econom. Dynam. Control 33 (3), 649–665.

Ríos-Rull, J.-V., 1997. Computation of Equilibria in Heterogeneous Agent Models. Tech. Rep., Federal Reserve Bank of Minneapolis.

Rudebusch, G.D., Swanson, E.T., 2012. The bond premium in a DSGE model with long-run real and nominal risks. Am. Econ. J.: Macroecon. 4 (1), 105–143.

Tauchen, G., Hussey, R., 1991. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. Econ.: J. Econ. Soc. 371–396.

Weidner, J., Kaplan, G., Violante, G., 2014. The wealthy hand-to-mouth. Brookings Pap. Econ. Act. 1, 77–153.

Young, E.R., 2010. Solving the incomplete markets model with aggregate uncertainty using the Krusell Smith algorithm and non-stochastic simulations. J. Econom. Dynam. Control 34 (1), 36–41, Computational Suite of Models with Heterogeneous Agents: Incomplete Markets and Aggregate Uncertainty.