Higher-Order Perturbation in Sequence Space: the Certainty Correspondence

Adrien Auclert* Rodolfo Rigato† Matthew Rognlie‡ Ludwig Straub§

October 2026

Abstract

We show that higher-order perturbation solutions to rational expectations models can be obtained from the derivatives of the model's perfect foresight solution. This *certainty correspondence* result extends certainty equivalence beyond the first order. Our method delivers the same solution as state-space perturbation, but for models with large idiosyncratic state spaces it is significantly faster and more practical. We apply it to solve a HANK model and a menu cost model to third order in aggregate risk, finding significant time variation in the fiscal multiplier and the responsiveness of inflation to costs. For plausible shocks we consider, third-order perturbation performs well against nonlinear methods.

*Stanford University, CEPR and NBER. Email: aauclert@stanford.edu.
†European Central Bank. Email: rodolfo.dinis_rigato@ecb.europa.eu.
‡Northwestern University and NBER. Email: matthew.rognlie@northwestern.edu.
§Harvard University, CEPR and NBER. Email: ludwigstraub@g.harvard.edu.

We thank Marlon Azinovic-Yang, Regis Barnichon, Luigi Bocola, Timo Boppart, Benjamin Born, Jeppe Druedahl, Ben Hebert, Greg Kaplan, Per Krusell, Ralph Luetticke, Kurt Mitman, Ben Moll, Normann Rion, Thomas Sargent, Yucheng Yang, Christian Wolf, and Nathan Zorzi for helpful comments. This research is supported by the National Science Foundation grant number SES-234393 and the Harvard Chae Initiative. The views expressed are those of the authors and do not necessarily reflect those of the European Central Bank.

1 Introduction

Heterogeneous-agent models with aggregate risk are notoriously difficult to solve in general equilibrium. A recent literature (Boppart, Krusell and Mitman 2018, Auclert, Bardóczy, Rognlie and Straub 2021) has made progress by exploiting *certainty equivalence*: the fact that to first order in shock size, aggregate risk is neutral, and future variables can be replaced by their expectations. A consequence is that to first order, the impulse response to an "MIT shock"—a one-time, unexpected shock, with perfect foresight thereafter—is the same as in a stochastic model. Such impulse responses can be obtained with sequence-space methods, sidestepping the curse of dimensionality that often arises in the state space, and enabling faster and easier computation.

The restriction to first-order aggregate risk, however, is currently a major drawback of the sequence-space approach. In a first-order approximation, there is no role for size- or history-dependent effects, nor is there any precautionary response to aggregate risk. It is impossible to study how aggregate volatility, or policies responding to it, affects welfare.

In this paper, we overcome these limitations by extending sequence-space methods to higher order. We show that certainty equivalence can be generalized to what we call the *certainty correspondence*: a systematic relationship between MIT-shock impulse responses and the higher-order perturbation solution. We derive this explicitly for the second- and third-order cases, showing how each term can be obtained via the certainty correspondence, and apply these methods numerically to obtain the full third-order solution of a HANK model and a menu cost model. Our perturbation approach delivers the same solution as a state-space approach with pruning (Andreasen, Fernández-Villaverde and Rubio-Ramírez 2018), but comes with the computational benefits and speed of sequence-space methods. It also delivers high accuracy relative to a global solution for our partial-equilibrium model, where computing such a solution is straightforward.

As a result of the certainty correspondence, higher-order perturbation is straightforward to implement for anyone who can calculate nonlinear perfect-foresight impulse responses to shocks—extending the Boppart et al. (2018) insight beyond first order. We provide tools to speed up general equilibrium and avoid numerical differentiation, as well as formulas for generalized impulse response functions and moments of the stochastic solution. For a baseline heterogeneous-agent model, we obtain a full second-order solution and fifth-order accurate moments in less than a minute, and a complete third-order solution in less than 5 minutes on a standard laptop. A code repository illustrating our methods accompanies the paper.[1]

The foundation of our approach is the nonlinear sequence-space solution $X_t = \mathcal{X}(\sigma, \epsilon_t, \epsilon_{t-1}, \ldots)$, which expresses the path of an economy as a function of anticipated risk σ and the infinite history $\epsilon_t, \epsilon_{t-1}, \ldots$ of realized shocks.[2] Whenever a locally stable state-space solution $X_t = g(X_{t-1}, \epsilon_t, \sigma)$ exists, a sequence-space solution also exists, and we can use perturbation to approximate it around

[1] See `https://github.com/shade-econ/higher-order`.

[2] Lan and Meyer-Gohde (2013b) also study such a solution, which they call a "nonlinear moving average", while Lombardo and Uhlig (2018) and Bhandari, Bourany, Evans and Golosov (2023) study a similar function where the history is finite—starting at date 0—rather than infinite. We discuss the relation to these papers further below.

the deterministic steady state X. To first order, this is exactly the approach advocated by Boppart et al. (2018): since $\mathcal{X}_\sigma = 0$ (certainty equivalence), the first-order solution is determined by the derivatives $\mathcal{X}_j$ with respect to past shocks ϵ_{t-j}, which can be identified with the first-order impulse response to an MIT shock.

At second order, certainty equivalence no longer holds, since there is a constant risk term $\mathcal{X}_{\sigma\sigma}$. All other terms in the second-order expansion, however, are independent of risk σ, and can be obtained by differentiating impulse responses to MIT shocks: to get *size-dependence* effects $\mathcal{X}_{jj}$, we take the second derivative with respect to a new shock; while to get *history-dependence* effects $\mathcal{X}_{jk}$, we take the mixed second derivative with respect to two shocks that are $k - j$ periods apart. We further show that, remarkably, the risk term $\mathcal{X}_{\sigma\sigma}$ can *also* be obtained from the information in an MIT-shock solution: it is simply the effect from perturbing steady-state expectations by the impact size-dependence term $\mathcal{X}_{00}$. Intuitively, this is because in the risky steady state $X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2$ (the steady state when risk is anticipated, but no shocks are ever realized), the expectations of tomorrow's state are equal to the constant $X + \frac{1}{2}(\mathcal{X}_{\sigma\sigma} + \mathcal{X}_{00})\sigma^2$, so we can use our knowledge of $\mathcal{X}_{00}$ to solve for $\mathcal{X}_{\sigma\sigma}$. This is the second-order certainty correspondence.

Moving on to third order, most of the new terms in the expansion of $\mathcal{X}$ are still independent of risk: they are of the form $\mathcal{X}_{jkk'}$, and can be obtained from triple derivatives of responses to MIT shocks that are $k' - k$ and $k - j$ periods apart. The new risk-dependent terms have the form $\mathcal{X}_{\sigma\sigma j}$, and capture how risk alters the transmission from ϵ_{t-j} to X_t. We decompose these terms into two effects. First, there is a *risky steady state* effect: the second-order shift in the steady state changes first-order impulse responses. We can obtain this effect by comparing responses around the deterministic and risky steady states. Second, there is a *time-varying precautionary* effect, reflecting how realized shocks change the economy's response to anticipated future risk. We show that this effect can also be obtained from a perfect-foresight transition, with expectations perturbed by the known sequence $\mathcal{X}_{00j}$. This completes the third-order certainty correspondence.

We explain how to apply the second- or third-order sequence-space solutions to simulate the model, obtain its generalized impulse responses, and calculate its moments. For simulations, we evaluate the sequence space solution at order n on a sequence of shocks ϵ_{t-j}; this gives the same outcome as the pruned state-space simulations at order n when starting from the risky steady state (Lombardo and Uhlig 2018). For generalized impulse responses, we give explicit formulas that are almost identical to their MIT shock counterparts, but with corrections for expected size or size-history dependence as well as time-varying precautionary effects. Finally, for moments, we obtain parsimonious formulas using a formal perturbation of the moments of the nonlinear model. We show that evaluating covariances at any horizon using the first-order solution as in Auclert et al. (2021) is accurate to third order, and augmenting this with some terms from the second- and third-order solutions increases accuracy to fifth order.[3]

The only requirement of our approach is the ability to compute nonlinear responses to MIT

[3]These moment formulas differ from the more conventional approach of directly evaluating moments of the perturbation approximation, as in Andreasen et al. (2018).

shocks, allowing for deterministic perturbations to expectations, and to repeatedly differentiate these solutions around the steady state. However, since a naive implementation via numerical differentiation of a nonlinear general equilibrium routine would be prohibitively slow, we introduce additional results that allow for much more efficient computation. First, we introduce a "one-shot principle" that bypasses the general equilibrium nonlinear fixed point in sequence space: instead, we show how the kth-order solution can be obtained by evaluating equilibrium conditions on the $k-1$th-order solution, finding the kth-order error, and then applying the inverse Jacobian used to solve the first-order system. Second, we show how analytical differentiation that exploits the envelope theorem and handles the interaction between moving kinks in policy functions and the smooth distribution can further speed up the code by almost two orders of magnitude. Overall, these improvements bring down the time to obtain a complete third-order solution of a HANK model to under 5 minutes.

Our primary application is a standard though "smooth" heterogeneous-agent model, first in partial and then in general equilibrium.[4] In partial equilibrium, we study the response to income shocks—finding that, as expected from the concavity of the consumption function, the second-order term for contemporaneous consumption is negative. Interestingly, however, we find that the second-order terms for consumption in subsequent periods are *positive*—so that, as we scale up an income shock, the delayed effect on consumption increases more than proportionally. This result extends to a general equilibrium HANK setting where income fluctuates due to deficit-financed transfers. Prior work has argued that such delayed consumption effects were partly responsible for sustained high demand in the US after the 2020–21 fiscal interventions (e.g. Auclert, Rognlie and Straub 2023). Our finding that these delayed effects get even stronger for large shocks strengthens this argument, given the unprecedented size of these interventions.

We also study how the impact transfer multiplier varies with the level of current economic activity. We find that the multiplier can be twice as large in recessions, when the economy is depressed, balance sheets are weak, and MPCs are high. We show that this result hinges on demand shocks being a dominant driver of fluctuations, as found by our estimation procedure for the US over the 1966-2025 period. When instead only supply shocks are present, the effects are flipped, with multipliers being slightly larger in expansions. These findings may help explain why the empirical literature on the state dependence of fiscal multipliers has generated mixed results (e.g. Auerbach and Gorodnichenko 2012 vs. Ramey and Zubairy 2018).

Our second main application is a menu cost model with idiosyncratic shocks in the spirit of Golosov and Lucas (2007). For that model, we study how the responsiveness of inflation to nominal cost shocks is affected by the current level of inflation. In line with other recent studies that use similar models (e.g. Blanco, Boar, Jones and Midrigan 2024), we find a significantly higher responsiveness of inflation to costs in times of high inflation.

[4]The "smooth" model is a standard incomplete market (SIM) model with a persistent-transitory income process, but where the i.i.d. income shocks are drawn from a smooth density rather than a discrete distribution. As we discuss, this feature of the model is important to make the model differentiable to second and higher order, overcoming the difficulties with the non-smooth model highlighted by Bhandari et al. (2023).

We conduct extensive accuracy checks throughout the paper. First, we test our third-order sequence-space perturbation routines on small state spaces by showing that, in versions of a neoclassical growth model with N countries, they deliver exactly the same solution as third-order state-space perturbation, though the sequence-space approach becomes much faster and memory-efficient as N gets large. Second, we compare our partial equilibrium heterogeneous-agent model solution to that obtained from a global solution, which is attainable via a Bellman approach. We find that third order perturbation gets very close to the global solution. We also show that it is much more accurate than a perfect-foresight solution, since the former captures the effect of aggregate precautionary savings on MPCs that the latter misses. Finally, we show that the non-risk components of the third-order perturbation solution are very close to the nonlinear perfect-foresight solution for shocks of plausible magnitude, persistence, and distribution. This contrasts with the finding of significant deviations between the third-order and the nonlinear perfect-foresight solution in Bianchi and Kaplan (2026). As we show in appendix D, such strong nonlinearities tend to appear only when most or all of the following apply: very large shocks, lump-sum transfers, purely transitory shocks, and a focus on impact rather than cumulative responses.

Related literature. The two closest existing papers to ours are Lan and Meyer-Gohde (2013b) and Bhandari et al. (2023). Lan and Meyer-Gohde (2013b) studies a nonlinear sequence-space solution of the same form (a "nonlinear moving average") and its perturbation solution up to third order, also known as a Volterra expansion. Our paper differs in how we compute this solution: Lan and Meyer-Gohde (2013b) use techniques that parallel the state-space solution,[5] while we introduce the certainty correspondence, allowing us to obtain the solution using exclusively perfect-foresight methods.[6]

Bhandari et al. (2023) perform second-order analytic differentiation of heterogeneous-agent models, and also arrive at a second-order Volterra expansion. They further evaluate the terms in this expansion by applying directional derivatives in a way that economizes on state-space computations. We differ by deriving the certainty correspondence, yielding an alternative computational strategy that connects to nonlinear perfect-foresight solutions—which, thanks to our smooth model, we can implement with numerical differentiation. We also cover the third-order solution in detail, while Bhandari et al. (2023) only explicitly implement the second order.[7]

Another closely related paper is Lombardo and Uhlig (2018), which uses a nonlinear sequence-space solution to explain why the state-space perturbation solution should be pruned: namely, because this makes it equivalent to the perturbed sequence-space solution (see also Lan and Meyer-Gohde 2013a). We re-derive this important result in our setting. Unlike our paper and Lan and

[5] As they note in footnote 21: "Thus, our nonlinear moving average solution parallels nonlinear state space solutions in a manner analogous to the linear case, where the recursion is in the coefficients as opposed to the variables themselves."

[6] In addition, as we discuss in appendix B.2, the certainty correspondence as we define it does not hold in the Lan and Meyer-Gohde (2013b) framework, because of a difference in where expectations are taken.

[7] See also Rion (2026) for an implementation of the Bhandari et al. (2023) method in Dynare, and a reconciliation of this method with the sequence-space approach to first order.

Meyer-Gohde (2013a), however Lombardo and Uhlig (2018) and Bhandari et al. (2023) only track the finite history since time 0, which makes their sequence-space solution vary with t.

Our paper also connects to the broader state-space perturbation literature, including the seminal work of Schmitt-Grohé and Uribe (2004) and Kim, Kim, Schaumburg and Sims (2008), and the pruned state-space system of Andreasen et al. (2018). Several recent papers, including Bilal (2023), Reiter (2023), and Bayer, Luetticke, Weiss and Winkelmann (2026b), have extended state-space perturbation to the heterogeneous-agent context. For rich enough heterogeneity, however, this can require significant "model reduction"—as the size of a second-order state-space solution is cubic in the dimension of the state, and becomes prohibitively large and costly when this dimension is high, as we show in appendix E.7. To our knowledge, our work is the first to point out the certainty correspondence result, and also to obtain a third-order solution, without model reduction, in the presence of substantial heterogeneity.[8]

Finally, our paper is connected to a literature that seeks "global" solutions to heterogeneous-agent models to study the effect of aggregate risk beyond first order. This literature uses a variety of approaches, such as neural networks (Kase, Melosi and Rottner 2022), deep learning (Han, Yang and E 2021, Azinovic-Yang and Žemlička 2025, Druedahl, Huleux and Røpke 2026), structural reinforcement learning (Yang, Wang, Schaab and Moll 2025), and alternative methods based on simulating specific paths (Lin and Peruffo 2024, Lee 2025, Adjemian and Juillard 2026).[9] We view this literature as highly complementary. While our test against a global partial equilibrium solution performs well, global solutions may be useful for general equilibrium models with sufficiently large nonlinearities. Even then, our easy-to-obtain third-order perturbation solution can be used as a seed for global methods, as well as a benchmark to verify against. The extent to which global solutions differ from the third-order solution for typical general equilibrium heterogeneous-agent models remains an intriguing open question.

2 Framework

2.1 General setup

We consider a model driven by an exogenous shock process ϵ_t with i.i.d. innovations.[10] In the model, the endogenous variable $X_t \in \mathbb{R}^n$ is determined at t to solve the system of equations

$$F\left(X_{t-1}, X_t, \mathbb{E}_t\left[X_{t+1}\right], \epsilon_t\right) = 0, \tag{1}$$

[8] Auclert, Rognlie, Straub and Ťapák (2024b) use a second-order perturbation approach to solve for zero-th order portfolios and steady-state risk premia in a model with portfolio choice. Future work could combine this approach with the methods from this paper to solve for time-varying risk premia.

[9] Like our method, Adjemian and Juillard (2026)'s stochastic extended path (SEP) method also augments nonlinear perfect-foresight transitions to incorporate the effect of aggregate risk, but it does so by simulating and then integrating over shocks for a finite number of future periods, while our method is exact to the perturbation order considered for recurring aggregate risk over an infinite horizon, and does not require simulation.

[10] We assume a single exogenous shock for notational convenience. Results generalize in a straightforward way to multiple exogenous shocks; see appendix B.7 for details.

with F mapping $(\mathbb{R}^n)^3 \times \mathbb{R}$ to $\mathbb{R}^n$. Our results apply for any n, but they are especially useful when n is large—for instance, in discretized heterogeneous-agent models, where typically $n \geq 1000$.

The expectations in (1) are taken assuming that $\epsilon_t = \sigma\bar{\epsilon}_t$, where $\bar{\epsilon}_t$ are drawn i.i.d. from a distribution that is symmetric around zero, has unit variance, and has bounded support, and σ is a scalar that indexes the standard deviation of shocks. The actual, realized ϵ_t are given by $\epsilon_t = \sigma_R\bar{\epsilon}_t$. Our objective is to solve the rational expectations case with $\sigma = \sigma_R$, but our certainty correspondence result will connect this case to perfect-foresight or "MIT shock" impulse responses, where $\sigma = 0$ but $\sigma_R \neq 0$.

A *deterministic steady state* of the model is a constant vector $X \in \mathbb{R}^n$ that solves

$$F(X, X, X, 0) = 0.$$

A *nonlinear sequence-space solution* of the model around X is a continuous function $\mathcal{X}$, satisfying $\mathcal{X}(0, 0, 0, \ldots) = X$, such that

$$X_t = \mathcal{X}(\sigma, \epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots) \tag{2}$$

solves equation (1) for all σ and $\{\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots\}$ in some neighborhood of 0. This solution expresses the path of X_t as a function of anticipated risk and the infinite history of realized shocks.[11]

A *nonlinear state-space solution* of the model around X, by contrast, is a continuous function g, satisfying $g(X, 0, 0) = X$, such that

$$X_t = g(X_{t-1}, \epsilon_t, \sigma) \tag{3}$$

solves equation (1) for all X_{t-1}, ϵ_t, and σ in some neighborhood of $(X, 0, 0)$. In this more common solution concept, the history of past shocks is summarized by the n-dimensional lagged state X_{t-1}.[12] We say that g is *locally stable* if it is differentiable around $(X, 0, 0)$, and its derivative g_X there has all eigenvalues inside the unit circle.

We next make the following assumption, which parallels the usual conditions assumed in the linear rational expectations literature (e.g. Blanchard and Kahn 1980 or Klein 2000.) Here F_1, F_2, F_3 denote the partial derivatives of F at the steady state with respect to its first three arguments.

Assumption 1. *F is continuously differentiable in a neighborhood of the deterministic steady state, and:*

a) $\det(F_1 + \lambda F_2 + \lambda^2 F_3) = 0$ *has n roots $|\lambda| < 1$ inside the unit circle, counted including algebraic multiplicity, with no roots $|\lambda| = 1$ on the unit circle.*

b) The rank condition (A.1) is satisfied.

This assumption guarantees the existence and uniqueness of both a sequence-space solution

[11] Note that σ_R does not directly appear in (2): given σ, only the realized sequence of ϵ_t matters for X_t. This property is also satisfied by the state-space solution (3). Continuity is defined using the ℓ^∞ norm.

[12] The state-space solution has a nonlinear vector autoregressive (VAR) form while the sequence-space solution has a nonlinear vector moving average (VMA) form. It is a well-known result from time series (eg Hamilton 1994, chapter 10) that, under a stability condition, linear VAR and VMA are equivalent representations of the same process. Proposition 1 shows that this equivalence also holds for nonlinear solutions of the model in (1).

and a locally stable state-space solution, as the following proposition shows.[13]

Proposition 1. *Assume that assumption 1 holds. Then both a nonlinear sequence-space solution $\mathcal{X}$ and a locally stable nonlinear state-space solution g exist and are unique in a neighborhood of the steady state. Furthermore, they are equivalent to each other in the sense that*

$$\mathcal{X}\left(\sigma, \epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots\right) = g\left(\mathcal{X}\left(\sigma, \epsilon_{t-1}, \epsilon_{t-2}, \epsilon_{t-3}, \ldots\right), \epsilon_t, \sigma\right). \tag{4}$$

If F is $\mathcal{C}^r$, then g and $\mathcal{X}$ are also $\mathcal{C}^r$.

In the state space, "global" solution methods attempt to solve for the entire function g. Since this can be challenging, a popular alternative approach is to solve instead for a Taylor expansion of the function g in (3) around the deterministic steady state $(X, 0, 0)$: this is known as state-space perturbation (e.g., Schmitt-Grohé and Uribe 2004, Fernández-Villaverde, Rubio-Ramírez and Schorfheide 2016.) Our sequence-space perturbation approach parallels state-space perturbation: since solving for the function $\mathcal{X}$ directly is challenging,[14] we solve instead for a Taylor expansion of the $\mathcal{X}$ function in (2) around the deterministic steady state $(0, 0, 0, 0, \ldots)$.

Proposition 1 implies that state-space and sequence-space perturbation solutions are closely connected. We review this connection, which has previously been established by Lan and Meyer-Gohde (2013b) and Lombardo and Uhlig (2018), later in the paper, and use it to validate our results. Our main theoretical contribution in this paper is instead to connect the sequence-space perturbation solution to the perfect-foresight solution, $\mathcal{X}(0, \epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$.

We finally note that since $\bar{\epsilon}_t$ is distributed symmetrically around zero, the distributions of $\sigma\bar{\epsilon}_t$ and $-\sigma\bar{\epsilon}_t$ are identical. This implies the following lemma, which we will use frequently.

Lemma 1 (Symmetry in σ)**.** Both g and $\mathcal{X}$ are even functions of σ. Hence every derivative containing an odd number of differentiations with respect to σ, provided it exists, is zero at $\sigma = 0$.

2.2 Examples

We use three simple model examples to illustrate our theoretical results. In section 5, we introduce more complex models for advanced applications.

Example 1: The Neoclassical Growth Model. Our first example is a neoclassical growth model with shocks, as in e.g. Schmitt-Grohé and Uribe (2004). We will use this classic model to verify that our methods deliver the same results as standard state-space perturbation. In this model, a planner maximizes the utility $\mathbb{E}\sum\beta^t u(C_t)$ of a representative agent, with CRRA period utility $u(C) = C^{1-\gamma}/(1-\gamma)$, subject to the feasibility constraint $C_t + K_t - (1-\delta)K_{t-1} = e^{a_t}K_{t-1}^{\alpha}$, and an

[13] We show in appendix A.2 that assuming continuous differentiability of F, the root and rank conditions are equivalent to the invertibility of the sequence-space Jacobian system formed by the derivatives of F, which can be tested following Auclert, Majic, Rognlie and Straub (2025).

[14] See Azinovic-Yang and Žemlička (2025) for a deep learning approach to solving $\mathcal{X}$ globally.

AR(1) shock process for log TFP $a_t = \rho_a a_{t-1} + \epsilon_t$. The optimum in this model is characterized by an Euler equation $C_t^{-\gamma} = \beta \mathbb{E}_t \left[\left(1 - \delta + \alpha e^{a_{t+1}} K_t^{\alpha - 1} \right) C_{t+1}^{-\gamma} \right]$.

To put this in the notation of (1), we define $\lambda_t \equiv \left(1 - \delta + \alpha e^{a_t} K_{t-1}^{\alpha-1} \right) C_t^{-\gamma}$ to be the shadow value of capital at t, and define $X_t \equiv (a_t, \lambda_t, K_t, C_t)$. Then we can stack

$$F\left(X_{t-1}, X_t, \mathbb{E}_t\left[X_{t+1}\right], \epsilon_t\right) \equiv \begin{bmatrix} \lambda_t - \left(1 - \delta + \alpha e^{a_t} K_{t-1}^{\alpha-1} \right) C_t^{-\gamma} \\ C_t^{-\gamma} - \beta \mathbb{E}_t\left[\lambda_{t+1}\right] \\ C_t + K_t - (1-\delta) K_{t-1} - e^{a_t} K_{t-1}^{\alpha} \\ a_t - \rho_a a_{t-1} - \epsilon_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{5}$$

Observe that, by appropriate choice of λ_t, we have cast the model in such a way that expectations at time t show up only through $\mathbb{E}_t[X_{t+1}]$, fitting the framework in (1).

Example 2: The Smooth Standard Incomplete Markets (SIM) Model. We next consider a partial equilibrium heterogeneous-agent model in which consumers face idiosyncratic income risk and a borrowing constraint. This model is convenient because it can be solved globally in the state space, and so we can verify the accuracy of our perturbation against this global solution.[15] The Bellman equation for a consumer with assets a, persistent income state e, and discount factor β is

$$\begin{aligned} V_t(a, e, \beta) &= \mathbb{E}_\xi \left[\max_{c, a'} u(c) + \beta \mathbb{E}_t[V_{t+1}(a', e', \beta') | e, \beta] \right] \\ \text{s.t.} \qquad & a' + c = (1 + r)a + Z_t \xi e + Tr; \qquad a' \geq \underline{a} \end{aligned} \tag{6}$$

where $u(\cdot)$ continues to be CRRA utility, r is the interest rate, and Tr is a constant transfer. Labor income $Z_t \xi e$ is made up of three parts: an aggregate component Z_t, which follows an AR(1) process, $Z_t - Z_{ss} = \rho_Z (Z_{t-1} - Z_{ss}) + \epsilon_t$ with mean Z_{ss},[16] a persistent idiosyncratic component e, which follows a finite Markov chain with a fixed transition matrix Π^e, and an independent transitory idiosyncratic component ξ, which is i.i.d. smoothly distributed with density $f(\cdot)$ realized at the beginning of the period, with $\mathbb{E}[\xi] = \mathbb{E}[e] = 1$. Patience β evolves independently of income according to a finite Markov chain with a fixed transition matrix Π^β, with the overall transition matrix between discrete states (e, β) denoted by $\Pi = \Pi^e \otimes \Pi^\beta$.

Appendix A.5 shows that this model can be put into the form in (1) by stacking the marginal value function $V_{a,t}$ at each discretization point into a vector $\boldsymbol{v}_t$, the distribution Φ_t at each discretization point into a vector $\boldsymbol{\phi}_t$, and then forming the large vector $X_t \equiv (\boldsymbol{v}_t, \boldsymbol{\phi}_t, C_t, Z_t)$, where C_t is aggregate consumption (the integral of the consumption policy across the distribution), such

[15] We thank Ben Moll for this suggestion.

[16] We consider values of $\sigma_R < \frac{(1-\rho_Z) Z_{ss}}{\sup |\bar{\epsilon}|}$, so that Z_t remains strictly positive for any possible realization of $\bar{\epsilon}_t$.

that

$$F\left(X_{t-1}, X_t, \mathbb{E}_t\left[X_{t+1}\right], \epsilon_t\right) \equiv \begin{bmatrix} v_t - \mathcal{V}(Z_t, \mathbb{E}_t v_{t+1}) \\ \boldsymbol{\phi}_t - \Lambda(Z_t, \mathbb{E}_t v_{t+1}, \boldsymbol{\phi}_{t-1}) \\ C_t - \mathcal{C}(Z_t, \mathbb{E}_t v_{t+1}, \boldsymbol{\phi}_{t-1}) \\ Z_t - Z_{ss} - \rho_Z(Z_{t-1} - Z_{ss}) - \epsilon_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{7}$$

for some functions $\mathcal{V}$, Λ, and $\mathcal{C}$ representing respectively the Bellman equation, the law of motion of the distribution, and the aggregation equation. Observe that here also, expectations at time t enter as $\mathbb{E}_t[X_{t+1}]$ because we have made the marginal value function v_t an element of X_t.

Appendix C provides more details for both the underlying smooth model and its computation. There, we emphasize the importance of the smooth ξ shock in making the overall model smooth: the way we numerically implement them, $V_{a,t}$ is a differentiable function of a, Φ_t is a differentiable function of a above the constraint $\underline{a}$, and C_t is a differentiable function of Z_t, each to order at least three. By contrast, the model with a discretized ξ shock, or with no ξ shock at all, is not differentiable beyond first order. As emphasized by Bhandari et al. (2023), this non-smoothness significantly complicates the implementation of higher-order perturbation. We also show that our introduction of a smooth distribution does not change the economics of the model—which, conceptually, takes in a continuous income process anyway—or any of its first-order implications, but merely adds the convenience of higher-order differentiability.

Our calibration, which is at a quarterly frequency, mostly follows Auclert, Rognlie and Straub (2024a), with the smooth i.i.d. income process calibrated to the magnitude of transitory income shocks inferred from Ganong, Noel, Patterson, Vavra and Weinberg (2025). We describe this calibration in appendix A.5, and summarize our calibration parameters in table 1.

Example 3: The Smooth Intertemporal Keynesian Cross (IKC) Model. Finally, we develop a New Keynesian general equilibrium version of the smooth SIM model, along the lines of the "intertemporal Keynesian cross" model in Auclert et al. (2024a). This model constitutes a natural use case for our method: since it features high n and general equilibrium, a global solution and a state-space perturbation solution without model reduction are both similarly out of reach.

The underlying partial equilibrium problem is the same as in Example 2: households still solve the Bellman equation (6), except that now $Z_t = (1 - \tau_t)Y_t$ is equal to production net of taxes. Production is from labor, $Y_t = N_t$, nominal prices are flexible, nominal wages are sticky, and monetary policy maintains the ex-ante real interest rate equal to r at all times.[17] The government's budget constraint is

$$B_t + \tau_t Y_t = (1 + r)B_{t-1} + Tr + G$$

The level of bonds B_t is stochastic, and follows an AR(1) process $B_t - B = \rho_B(B_{t-1} - B) + \epsilon_t$,[18] with the government adjusting the tax rate τ_t each period to meet this bond target. In equilibrium,

[17] Nominal wages are set by unions on behalf of households, with labor N_t entering additively separably into the utility function. Given the real interest rate rule and the fact that bonds are real assets, the resulting Phillips curve is not relevant for the dynamics of real variables. The quantitative model in section 5.1 relaxes this assumption.

[18] Here again we consider σ_R low enough that B_t stays strictly positive for any possible realization of $\bar{\epsilon}_t$.

Table 1: Baseline calibration of heterogeneous-agent models

Parameter	Description	Value
Standard incomplete-markets model (partial equilibrium)		
r	Real interest rate	0.005
γ	Risk aversion	1
ρ_e	Persistence of persistent component	0.9767
σ_e	Innovation s.d. of persistent component	0.1964
σ_ξ	Innovation s.d. of transitory component	0.0964
β_L	Low-type discount factor	0.95996
β_H	High-type discount factor	0.99049
Z	Aggregate non-transfer income	0.67
Tr	Transfer income	0.14
A	Aggregate assets	4
$\partial C_0/\partial Z_0$	Aggregate MPC	0.2
ρ_Z	Aggregate income persistence	0
σ	Aggregate income innovation s.d.	0.05
IKC model (general equilibrium)		
Y	GDP	1
B	Government bonds	4
G	Government spending	0.17
τ	Marginal tax rate	0.33
ρ_B	Bond persistence	0.95
σ	Bond innovation s.d.	0.05

the goods market must clear, so $C_t + G_t = Y_t$, and the asset market must clear, so $A_t = B_t$.

Appendix A.5 describes the calibration of this model—which is summarized in table 1—and shows that it can be put in the form (1), described by equation (A.20). The important difference relative to (7) is the addition of an aggregate market clearing condition. This reflects a key difficulty of general equilibrium: due to the endogeneity (here) of aggregate post-tax income Z_t, the model can no longer be easily solved globally in the state space.

3 Certainty correspondence

In this section, we characterize the nonlinear sequence-space solution $\mathcal{X}(\sigma, \epsilon_t, \epsilon_{t-1}, \ldots)$ to third order in ϵ and σ. Our main contribution is to do so using only two kinds of objects: first, the responses of the economy to up to three consecutive unanticipated shocks (or "MIT shocks"), and second, the response of the economy to certain deterministic perturbations of expectations. Remarkably, this "perfect-foresight information" is sufficient to obtain the higher-order solution with aggregate risk.

3.1 First-order perturbation and certainty equivalence

We begin with the first order, which is already well understood. We evaluate the first-order Taylor approximation of the sequence-space solution $\mathcal{X}$ in (2) around the deterministic steady state

$(\sigma, \epsilon_t, \epsilon_{t-1} \ldots) = (0, 0, 0, \ldots)$ (or $\mathbf{0}$ for short). We write $\mathcal{X}_\sigma \equiv \frac{\partial \mathcal{X}}{\partial \sigma}(\mathbf{0})$ and $\mathcal{X}_j \equiv \frac{\partial \mathcal{X}}{\partial \epsilon_{-j}}(\mathbf{0})$. Assuming differentiability, and noting that $\mathcal{X}_\sigma = 0$ by lemma 1, this Taylor approximation is

$$X_t \approx X + \sum_{j=0}^{\infty} \mathcal{X}_j \epsilon_{t-j} \equiv X_t^{(1)} \tag{8}$$

Hence, to first order, $X_t - X$ is $MA(\infty)$ in ϵ_t. An important feature of (8) is that σ does not appear in $X_t^{(1)}$. Anticipated risk does not affect the first-order solution.

To compute the $MA(\infty)$ coefficients $\mathcal{X}_j$ that appear in (8), we define

$$X_j^{MIT}(\nu) \equiv \mathcal{X}(0, 0, \ldots, 0, \underbrace{\nu}_{j}, 0 \ldots) \tag{9}$$

as the response to a single MIT shock of size ν: a one-time, unanticipated shock at date 0, starting from the deterministic steady state, and with perfect foresight in every period j after. Since $\mathcal{X}_j$ is evaluated around this steady state and with $\sigma = 0$, we have:

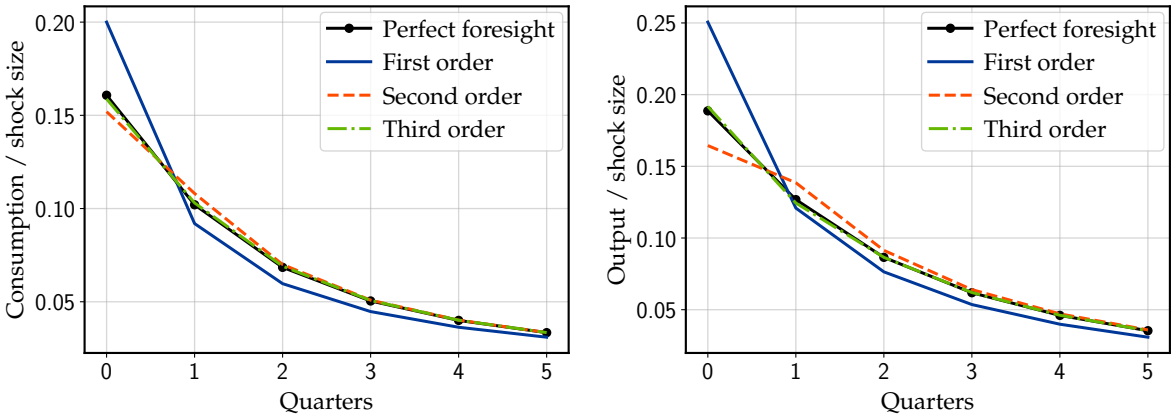
$$\mathcal{X}_j = (X_j^{MIT})'(0) \tag{10}$$

Hence, the $MA(\infty)$ coefficients in (8) can be obtained as the local impulse response to an MIT shock (see, e.g., Fernández-Villaverde et al. 2016 and Boppart et al. 2018). This is an especially helpful result for heterogeneous-agent models, since there are well-known procedures to calculate this MIT shock impulse response in these models, both in partial equilibrium where a simple backward-forward procedure suffices, or in general equilibrium where one can use the sequence-space Jacobian method (Auclert et al. 2021). We summarize this result as follows:

Proposition 2 (Certainty equivalence)**.** *The first-order perturbation of X_t in the sequence space is given by $X_t^{(1)}$ in (8). The constant term is the deterministic steady state X, and the coefficients on past innovations in the linear MA(∞) process are given by (10), so they can be obtained from linearized perfect-foresight impulse responses.*

The blue lines in figure 1 visualize the first-order impulse response of consumption $\mathcal{X}_j$ to a one-time innovation to ϵ in our partial equilibrium model (left) and our general equilibrium model (right). In the PE model with an i.i.d. process for aggregate labor income Z_t, a unit shock to ϵ increases Z_t by 1 on impact, and has no effect on Z_t in subsequent periods. As a result, the left panel traces out the intertemporal marginal propensities to consume $\partial C_j / \partial Z_0$ (iMPCs), with the impact effect calibrated to 0.2, and spending remaining elevated in the following quarters. In the GE model, a unit shock to ϵ lowers taxes T_t by 1 on impact and raises them slightly going forward as bonds return to the steady state at rate ρ_B. This triggers an intertemporal Keynesian cross mechanism (Auclert et al. 2024a), with equilibrium spending and income following the same pattern as iMPCs, but with a larger impact effect and slightly more persistence.

The black lines show the perfect foresight impulse responses to an innovation of size $\nu = 0.10$ (10% of quarterly GDP, equal to two standard deviations in our calibration), normalized by the size



(a) PE (intertemporal MPCs)

(b) GE impulse response

Figure 1: Impulse response of consumption over time, PE and GE

Note: The lines show the dynamic responses of consumption to a shock of size of 10% of deterministic steady state GDP divided by the shock size. We set $\sigma = 0$ and assume the economy is at the deterministic steady state when shocked.

of the shock, $(C_j^{MIT}(\nu) - C)/\nu$. The normalized impact effect is lower, but the effect is also more persistent in both partial and general equilibrium. This is a general feature of nonlinear iMPCs that we will discuss in the next section. For a shock of this magnitude, the difference between the first-order and the nonlinear perfect-foresight solution is significant, as pointed out by Bianchi and Kaplan (2026).

3.2 Second-order perturbation and certainty correspondence

We continue with a second-order Taylor approximation of $\mathcal{X}$ around the deterministic steady state. We now write $\mathcal{X}_{jk} \equiv \frac{\partial^2 \mathcal{X}}{\partial \epsilon_{-j} \partial \epsilon_{-k}}(\mathbf{0})$, and similarly for derivatives with respect to σ. Assuming second-order differentiability, a second-order Taylor approximation of X_t can be written as

$$X_t \approx X + \sum_{j=0}^{\infty} \mathcal{X}_j \epsilon_{t-j} + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2 + \frac{1}{2}\sum_{j,k\geq 0} \mathcal{X}_{jk}\epsilon_{t-j}\epsilon_{t-k} \equiv X_t^{(2)} \tag{11}$$

This expression is also known as the second-order Volterra expansion of X_t. It implies that, up to constants, X_t is a quadratic $MA(\infty)$ process in ϵ_t. By lemma 1, all mixed derivatives $\mathcal{X}_{\sigma j}$ are zero.

Using symmetry of mixed partials, the new second-order terms in (11) can be collected into three distinct groups:

$$X_t^{(2)} - X_t^{(1)} = \underbrace{\frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2}_{\text{Risk}} + \underbrace{\frac{1}{2}\sum_{j=0}^{\infty} \mathcal{X}_{jj}\epsilon_{t-j}^2}_{\text{Size}} + \underbrace{\sum_{j=0}^{\infty}\sum_{k>j}^{\infty} \mathcal{X}_{jk}\epsilon_{t-j}\epsilon_{t-k}}_{\text{History}} \tag{12}$$

Different from the first-order expansion, there is now an explicit term that captures the anticipation of future risk, $\frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2$. In particular, note that in the *risky steady state* in which agents continually

anticipate aggregate risk, even though this risk never materializes (Coeurdacier, Rey and Winant 2011), we have $\epsilon_t = 0$ at all dates, so the system is perpetually at $X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2$. Building on this observation, we now show that all the new terms in (12) can still be computed with perfect-foresight information from MIT shock responses and perturbed-expectation steady states.

For the size and history terms, since the derivatives $\mathcal{X}_{jk}$ are evaluated at the deterministic steady state, we use the same connection to MIT shock responses as above. In particular, the size effect in the quadratic MA expression, $\mathcal{X}_{jj}$, directly corresponds to the size effect in the MIT shock response,

$$\mathcal{X}_{jj} = (X_j^{MIT})''(0) \tag{13}$$

Similarly, the history effect $\mathcal{X}_{jk}$ corresponds to the response to two subsequent MIT shocks. To see this, define, for $l > 0$,

$$X_j^{MIT,l}(\nu, \epsilon) \equiv \mathcal{X}(0, 0, \ldots 0, \underbrace{\nu}_{j}, 0, \ldots, 0, \underbrace{\epsilon}_{j+l}, 0 \ldots) \tag{14}$$

as the response to the following sequence of perfect-foresight ($\sigma = 0$) experiments: first, starting from the deterministic steady state, an unexpected shock of size ϵ occurs, with no further shock expected. Then, after l periods, a second unexpected shock of size ν occurs, again with no further shock expected, and we trace out the effect of this combination of shocks after j periods.[19] Since $\mathcal{X}_{jk}$ is also evaluated around $\sigma = 0$, we have, for $k > j$ that

$$\mathcal{X}_{jk} = \frac{\partial^2 X_j^{MIT,k-j}}{\partial \nu \partial \epsilon}(0,0) \tag{15}$$

To characterize the risk term, we introduce the idea of a *perturbed-expectations steady state*. Consider the steady state that obtains if the expectations of X are always shifted by a small amount $\Delta \in \mathbb{R}^n$, solving

$$F\left(X^{pss}, X^{pss}, X^{pss} + \Delta, 0\right) = 0 \tag{16}$$

Let $X^{pss}(\Delta)$ denote the function implicitly defined by (16). Then, we have:

$$\mathcal{X}_{\sigma\sigma} = \left.\frac{\partial X^{pss}(\mathcal{X}_{00}\delta)}{\partial \delta}\right|_{\delta=0} \tag{17}$$

The intuition for this remarkable result is as follows. In the risky steady state with $X_t = X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2 \equiv X^*$, (11) shows that expectations of the future are equal to the constant $\mathbb{E}_t[X_{t+1}] = X^* + \frac{1}{2}\mathcal{X}_{00}\sigma^2$. Hence, if we solve for the perturbed-expectations steady state (16) that obtains when we set $\Delta = \frac{1}{2}\mathcal{X}_{00}\sigma^2$, we recover X^*. Letting $\delta = \frac{1}{2}\sigma^2 \to 0$ delivers (17).

Taking these results together, we obtain our second-order *certainty correspondence* result:

[19] Section 4.2 gives further detail about how to implement this sequence practically given an algorithm to solve nonlinear perfect-foresight impulse responses from steady state.

Proposition 3 (Second-order certainty correspondence)**.** *The second-order perturbation of X_t in the sequence space is given by the quadratic MA(∞) in (11). The new coefficients $\mathcal{X}_{jj}$, $\mathcal{X}_{jk}$ and $\mathcal{X}_{\sigma\sigma}$ are respectively given by equations (13), (15), and (17), so they can be entirely obtained from MIT shock responses and a perturbed-expectation steady state.*

A complete proof is given in appendix B.2. Again, what is striking is that only perfect foresight information is needed for to construct all second-order MA coefficients. The effect of anticipated risk only requires a perturbed-expectations steady state, with expectations perturbed by the anticipation of one-period-ahead size dependence $\mathcal{X}_{00}$.

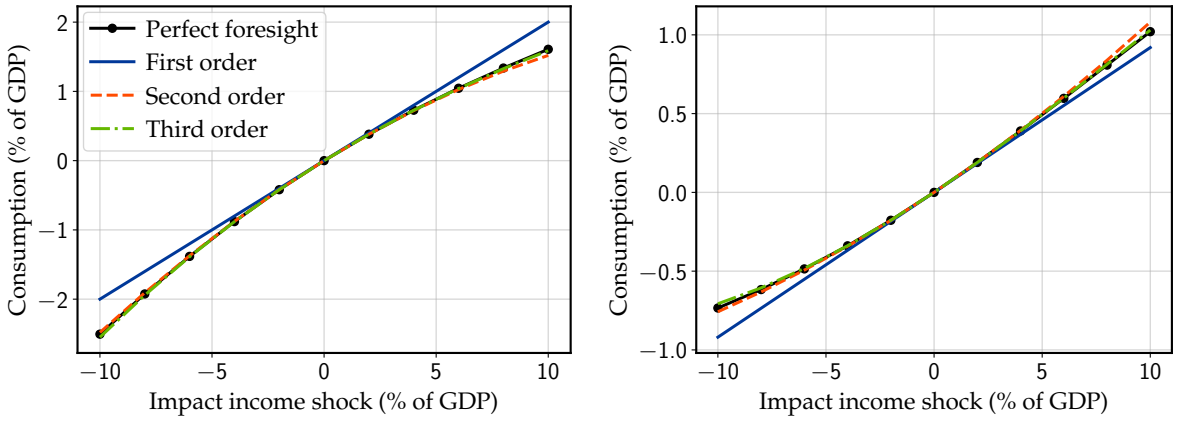
Notions of steady states. Using the second-order solution, we have made a distinction between the deterministic steady state X, with $\sigma = \sigma_R = 0$, and the risky steady state $X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2$, in which $\sigma > 0$ but $\sigma_R = 0$. There is also a further distinction with the *ergodic mean*, which is the average value of the system when risk realizes at its expected value $\sigma_R = \sigma > 0$. This mean can be immediately derived from (11) as

$$\mathbb{E}[X_t] = X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2 + \frac{1}{2}\sum_{j=0}^{\infty}\mathcal{X}_{jj}\sigma^2. \tag{18}$$

While these notions all collapse to the deterministic steady state X in the first-order solution (8), they generally differ from it in the second-order solution (see also Lan and Meyer-Gohde 2013b and Bhandari et al. 2023 for similar expressions.)

Equivalence to pruned state-space. As shown by Lombardo and Uhlig (2018), the second-order sequence space perturbation in (11) corresponds exactly to the so-called second-order pruned state space solution (Andreasen et al. 2018). Indeed, Lombardo and Uhlig (2018) derive a theory of pruning to any order based on this equivalence. We re-derive this result in appendix B.5, where we show how to map the coefficients in a state space law of motion (derivatives of g) to those in a Volterra expansion (derivatives of $\mathcal{X}$). Then, in appendix B.6, we use this result to verify the accuracy of our routines: we check that, in the the neoclassical growth model, the pruned state-space solution obtained using Kim et al. (2008)'s method, as currently implemented in Dynare, numerically coincides with our sequence-space solution.

Prior work on sequence-space perturbation. Lan and Meyer-Gohde (2013b) also pursue perturbation in the sequence space, and introduce the Volterra expansion in (11). Their paper obtains expressions for the coefficients in this expansion in terms of the derivatives of the F function, paralleling the usual approach to perturbation in the state space. By contrast, our result in proposition 3 shows that we can solve for these coefficients using perfect-foresight impulse responses, bypassing the need to explicitly obtain and manipulate derivatives of F. This is a significant improvement because obtaining these derivatives can be very complicated, especially in the case of



(a) Response at $t = 0$ (b) Response at $t = 1$

Figure 2: Size dependence in the response of consumption to income shocks (PE model)

Note: The lines show the responses of aggregate consumption at $t = 0$ and $t = 1$ for various shocks around the deterministic steady state with $\sigma = 0$. All units are in % of deterministic steady state GDP.

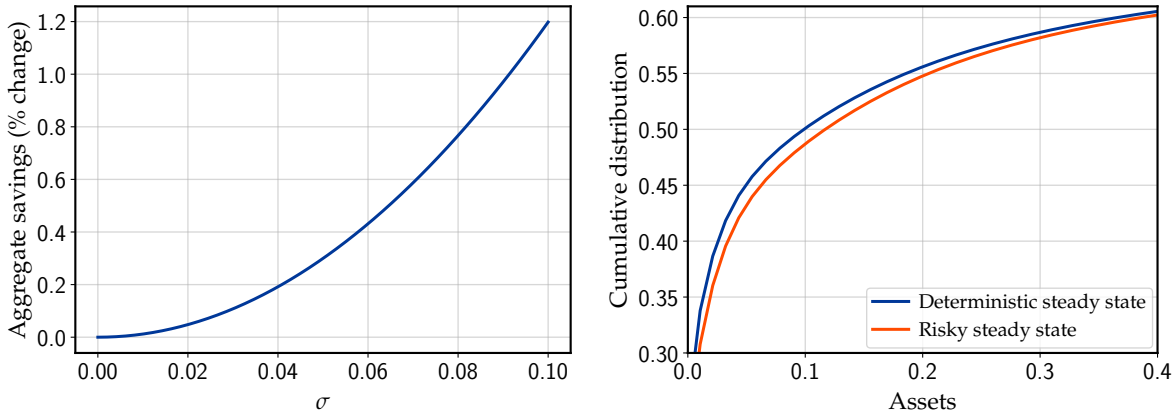
heterogeneous-agent models.[20]

Bhandari et al. (2023) also obtain a Volterra expansion closely related to (11), although their risk term for them is a time varying $\mathcal{X}_{\sigma\sigma t}$, reflecting a transition between the deterministic and the risky steady state. They offer a construction for the coefficients $\mathcal{X}_{jk}$ and $\mathcal{X}_{\sigma\sigma t}$ based on recursions in the sequence space using directional derivatives, bypassing the need for a full state-space law of motion, which is costly in heterogeneous-agent applications. The forcing term in their $\mathcal{X}_{\sigma\sigma t}$ recursion is also $\mathcal{X}_{00}$, and in the limit of $t \to \infty$, their $\mathcal{X}_{\sigma\sigma t}$ converges to our $\mathcal{X}_{\sigma\sigma}$. By contrast, we show how to explicitly construct the terms $\mathcal{X}_{jk}$ and $\mathcal{X}_{\sigma\sigma}$ by solving perfect-foresight transitions and a perturbed-expectations steady state. Our papers also differ because we cover the third order case, and because we offer to compute derivatives numerically on a smooth model as a benchmark, whereas they treat non-smooth model differentiability via generalized derivatives.

Application to heterogeneous-agent models. The orange dashed lines in figure 1 display the second-order normalized impulse response (the time path of $\mathcal{X}_j + \frac{1}{2}\mathcal{X}_{jj}\nu$ for aggregate consumption) for a shock of size $\nu = 2\sigma = 10\%$ of quarterly GDP, in partial and general equilibrium.

This second-order impulse captures a key feature of the nonlinear solution: in both models, the impact effect is lower than the first order impact due to the concavity of the consumption function ($\mathcal{X}_{00} < 0$), but the effect in later periods is higher, reflecting convexity in later periods ($\mathcal{X}_{jj} > 0$ for $j > 0$). This is a manifestation of the intertemporal budget constraint of agents, which holds irrespective of shock size: any additional savings from period 0 must later be spent down as agents return to their buffer stock of assets.

[20] Lan and Meyer-Gohde (2013b) also take a different starting point than (1) and write their model with the expectation operator on the outside of the F function, i.e. $\mathbb{E}_t\left[F\left(X_{t-1}, X_t, X_{t+1}, \epsilon_t\right)\right] = 0$. Our convention is useful because it enables us to obtain the $\mathcal{X}_{\sigma\sigma}$ term solely via (17), rather than needing to take into account the interaction of risk in X_{t+1} with the nonlinearity of F (see appendix B.2). As our examples illustrate, (1) usually obtains without loss of generality.



(a) Aggregate assets

(b) Wealth distribution

Figure 3: The risky steady state: aggregate savings and aggregate income uncertainty (PE model)

Note: The left panel shows the percent increase in savings in the risky steady state relative to the deterministic one as a function of σ. The right panel shows the cumulative wealth distribution in the deterministic steady state, and in risky steady state for $\sigma = 0.05$.

Figure 2 visualizes these results in a different way, by considering how the level of the consumption response net of the risky steady state varies with shock size on impact ($j = 0$, left panels) and at date 1 ($j = 1$, right panels) in the PE model (the results in the GE model are very similar and displayed in appendix figure B.4). Here again, the blue line represents the linear approximation, and the dashed orange line the second order approximation: this has a concave shape at date 0, and a convex shape at date 1, for the reason just discussed. Compared to the nonlinear perfect foresight solution $C_j^{MIT}(\nu) - C$ in the black line, for the range of shocks considered here (aggregate income shocks of up to 10% of quarterly GDP), the second order approximation does an excellent job, although the fit slightly deteriorates for very large shocks, as visible in figure 1.

We next turn to the "risk" effect $\mathcal{X}_{\sigma\sigma}$. While aggregate uncertainty affects all the elements of X, we focus on the effect of σ on aggregate asset accumulation in the risky steady state in the partial equilibrium model,[21] with corresponding results for the general equilibrium model in appendix figure B.5. Certainty correspondence allows us to obtain this in a straightforward way: we obtain the impact size dependence $\mathcal{X}_{00}$ in the marginal value function v_t, since this is the only term entering with expectations in (7), and then solve for the risky steady state $\mathcal{X}_{\sigma\sigma}$ using (17).

Figure 3 visualizes the outcome of this procedure. The left panel plots the effect of aggregate risk on risky steady-state aggregate assets as a function of σ, in other words the term $\frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2$ corresponding to total assets. The right panel plots the cumulative distribution function of assets, both at the deterministic steady state and at the risky steady state with $\sigma = 0.05$. Perceived aggregate risk makes agents save more at every point in the distribution, resulting in a rightward shift of the cumulative distribution function and a significant increase in aggregate precautionary savings. This shift lowers the mass of agents at the borrowing constraint by a significant amount.

[21] Note that this is closely related to the effect on aggregate consumption, since the consolidated budget constraint of the economy is $C = Y + rA$.

As a result, MPCs are significantly lower in the risky steady state, compared to the deterministic steady state. This has important implications for the third order perturbation below.

The left panel of Figure 4 concludes this section by visualizing the three new types of terms from the second-order perturbation as impulse responses to consumption. The dashed line shows the constant positive effect on risky-steady state consumption from higher asset accumulation. The blue line shows the contemporaneous size effect $\frac{1}{2}\mathcal{X}_{jj}$, which is negative and positive, reflecting the concave-convex pattern from figure 2. The orange line shows the lagged size effect $\frac{1}{2}\mathcal{X}_{j+1,j+1}$, which is positive everywhere with a decaying pattern, because large past shocks lead to additional asset accumulation, which is then spent down. Finally, the green line shows the history effect $\mathcal{X}_{j,j+1}$, which has the same pattern as the contemporaneous size effect: a past shock strengthens balance sheets and mitigates the impact consumption effect of a new shock, leading to more saving and therefore higher spending later on. One benefit of our perturbation approach over a global solution is that all these terms can be calculated and analyzed separately.

3.3 Third-order perturbation and certainty correspondence

We proceed with a third-order Taylor approximation of $\mathcal{X}$ around the deterministic steady state. We continue to write $\mathcal{X}_{jkk'} \equiv \frac{\partial^3 \mathcal{X}}{\partial \epsilon_{-j} \partial \epsilon_{-k} \partial \epsilon_{-k'}}(\mathbf{0})$, and similarly for derivatives with respect to σ. Assuming third-order differentiability, and using lemma 1 again to observe that all mixed derivatives $\mathcal{X}_{\sigma jk}$ are zero, a third-order Taylor approximation (Volterra expansion) of X_t can now be written as

$$X_t \approx X_t^{(2)} + \frac{1}{2}\sum_{j=0}^{\infty} \mathcal{X}_{\sigma\sigma j}\epsilon_{t-j}\sigma^2 + \frac{1}{6}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{k'=0}^{\infty} \mathcal{X}_{jkk'}\epsilon_{t-j}\epsilon_{t-k}\epsilon_{t-k'} \equiv X_t^{(3)} \tag{19}$$

where $X_t^{(2)}$ is the second-order Volterra expansion in equation (11). Hence, to third order, up to constants, X_t is a cubic $MA(\infty)$ process in ϵ_t.

The new term on the right in (19) corresponds to all triple interactions of shocks, and using the symmetry of mixed partials again, can further be rewritten as

$$\frac{1}{6}\sum_{j=0}^{\infty} \mathcal{X}_{jjj}\epsilon_{t-j}^3 + \frac{1}{2}\sum_{j=0}^{\infty}\sum_{k>j}^{\infty} \mathcal{X}_{jjk}\epsilon_{t-j}^2\epsilon_{t-k} + \frac{1}{2}\sum_{j=0}^{\infty}\sum_{k>j}^{\infty} \mathcal{X}_{jkk}\epsilon_{t-j}\epsilon_{t-k}^2 + \sum_{j=0}^{\infty}\sum_{k>j}^{\infty}\sum_{k'>k}^{\infty} \mathcal{X}_{jkk'}\epsilon_{t-j}\epsilon_{t-k}\epsilon_{t-k'}$$

These correspond, respectively, to cubic size dependence, the interaction of quadratic size dependence with history, the quadratic size history effect, and the history effect from the interaction of two past shocks. Just like with second-order certainty correspondence, these terms are all straightforwardly related to corresponding MIT shock responses. Indeed, for $j \geq 0$ we have

$$\mathcal{X}_{jjj} = (X_j^{MIT})'''(0) \tag{20}$$

while for $k > j$ we have

$$\mathcal{X}_{jjk} = \frac{\partial^3 X_j^{MIT,k-j}}{\partial^2 \nu \partial \epsilon}(0,0) \qquad \text{and} \quad \mathcal{X}_{jkk} = \frac{\partial^3 X_j^{MIT,k-j}}{\partial \nu \partial^2 \epsilon}(0,0) \tag{21}$$

Finally, we introduce for $j \geq 0$ and $l' > l > 0$,

$$X_j^{MIT,l,l'}(\nu, \epsilon, \epsilon') \equiv \mathcal{X}(0,0,\ldots 0, \underbrace{\nu}_{j}, 0, \ldots 0, \underbrace{\epsilon}_{j+l}, 0, \ldots, 0, \underbrace{\epsilon'}_{j+l'}, 0 \ldots) \tag{22}$$

as the response to three consecutive MIT shocks, with the first two separated by $l' - l$ and the second and third separated by l, again starting from the deterministic steady state and with no anticipation of a further shock at any point. Then, for $k' > k > j$ we have

$$\mathcal{X}_{jkk'} = \frac{\partial^3 X_j^{MIT,k-j,k'-j}}{\partial \nu \partial \epsilon \partial \epsilon'}(0,0,0). \tag{23}$$

This completes our characterization of the third-order size and history terms in (19).

The remaining term in (19) is that emerging from the $\epsilon_{t-j}\sigma^2$ interactions. This term quantifies both how the presence of aggregate risk affects the impulse response to shocks, and how the occurrence of past shocks affects the precautionary motive. Appendix B.3 shows that the overall effect is in fact the sum

$$\mathcal{X}_{\sigma\sigma j} = \left(\mathcal{X}_{\sigma\sigma j}\right)^{rss} + \left(\mathcal{X}_{\sigma\sigma j}\right)^{tvp}. \tag{24}$$

where *rss* stands for "risky steady state", and *tvp* for "time-varying precautionary".

The *rss* term in (24) corresponds to the perfect-foresight response to a small shock starting from the risky steady state.[22] The *tvp* term, on the other hand, captures the effect that a shock that hit j periods ago has on agents' behavior by changing their perception of risk $\mathcal{X}_{00j}\frac{\sigma^2}{2}$.

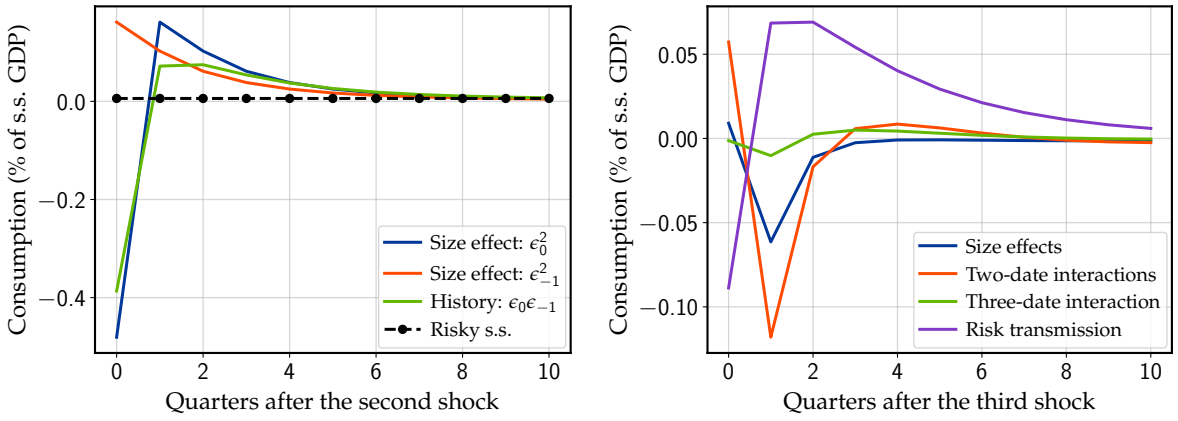
To define these terms formally, we introduce the notion of a *perturbed-expectations transition*, which we define as the deterministic transition emerging from a sequence of perturbed expectations $\Delta_j \in \mathbb{R}^n$, one for each $j \geq 0$, and a shock of size ν at $j = 0$, solving

$$F\left(X_{j-1}^{pt}, X_j^{pt}, X_{j+1}^{pt} + \Delta_j, \nu \mathbf{1}_{j=0}\right) = 0 \tag{25}$$

with boundary conditions $X_{-1}^{pt} = \lim_{j\to\infty} X_j^{pt} = X^{pss}(\Delta_\infty)$. Then, denoting by $X_j^{pt}(\{\Delta_i\}, \nu)$ the function implicitly defined in (25), the *rss* term corresponds to the effect of a small shock ν given a constant perturbation of expectations $\mathcal{X}_{00}\delta$,

$$\left(\mathcal{X}_{\sigma\sigma j}\right)^{rss} = \left.\frac{\partial^2 X_j^{pt}\left(\{\mathcal{X}_{00}\delta, \mathcal{X}_{00}\delta, \ldots\}, \nu\right)}{\partial \delta \partial \nu}\right|_{\delta=\nu=0} \tag{26}$$

[22] The *rss* term is the one solved by the literature that performs "perturbation around the risky steady-state" (e.g. Andreasen, Caggiano, Castelnuovo and Pellegrino 2024). However, while our approach leads us to solve for the *rss* and the *tvp* term separately, a consistent stochastic perturbation delivers the sum of the two, i.e. (24).



(a) Second order

(b) Third order

Figure 4: Decomposing second and third order impulse on consumption (PE model)

Note: Panel (a) shows the contribution of second-order terms to consumption following two consecutive shocks at $t = -1$ and $t = 0$. Panel (b) shows the contribution of different sets of third-order terms following three consecutive shocks at $t \in \{-2, -1, 0\}$. All shocks have size 10% of deterministic steady state GDP and $\sigma = 0.05$.

while the *tvp* term is the effect of a time-varying sequence of expectations $\mathcal{X}_{00,j+1}\delta$ when $\nu = 0$,

$$\left(\mathcal{X}_{\sigma\sigma j}\right)^{tvp} = \left.\frac{\partial X_j^{pt}\left(\{\mathcal{X}_{001}\delta, \mathcal{X}_{002}\delta, \ldots\}, 0\right)}{\partial \delta}\right|_{\delta=0} \tag{27}$$

This completes our argument that we can obtain all the terms in the third-order perturbation using perfect-foresight information.

Proposition 4 (Third-order certainty correspondence)**.** *The third-order perturbation of X_t in the sequence space is given by the cubic MA(∞) in equation (19). The coefficients on the new terms are given by equations (20), (21), (23), (24), (26) and (27), so they can be obtained with perfect-foresight information.*

We prove the proposition in appendix B.3. To our knowledge, this characterization of the third-order solution is completely new: although Lan and Meyer-Gohde (2013b) gets closest by deriving the Volterra expansion for the third-order sequence-space solution, they do not connect the terms in their expansion to the perfect-foresight solution, or interpret the $\mathcal{X}_{\sigma\sigma j}$ term in terms of the two effects in (24). Again, our solution coincides theoretically with a pruned third-order state space perturbation, and in appendix B.6 we verify that it does in practice against the Dynare implementation of the Andreasen et al. (2018) formulas on the neoclassical growth model.

Application to heterogeneous-agent models. We continue to illustrate these results using our smooth SIM model in PE and GE. Figure 1 shows that the fit is now almost exact compared to nonlinear perfect foresight: the dash-dot green line (which shows the terms $\mathcal{X}_j + \frac{1}{2}\mathcal{X}_{jj}\nu + \frac{1}{6}\mathcal{X}_{jjj}\nu^2$ for $\nu = 10\%$ of quarterly GDP), is nearly identical to the solid black line in either model. Figure 2 shows that this continues to be the case for shocks of up to 10% of quarterly GDP in the PE model, and Figure B.4 shows the same for the GE model. Hence, compared to the nonlinear

perfect foresight solution, a third order perturbation works extremely well in our context: it not only captures the basic concavity/convexity patterns of nonlinear iMPCs, but also corrects for the changes in this curvature in a way that extrapolates well for fiscal shocks of a plausible size.

The right panel of figure 4 visualizes the new terms that come from the third order perturbation. First, the purple line shows the $\mathcal{X}_{\sigma\sigma t}$ term, reflecting the overall interaction of risk with the impulse response. This term is negative on impact and positive later on, notably because of precautionary savings at the risky steady state, which lowers MPCs on impact and raise them later on. Next, the blue line shows the overall size effects from a combination of three shocks, and the orange line shows the mixed size-history effects from sequences of large and small shocks. These terms have the opposite shape as the risk term. Finally, the green line visualizes the interaction term from a sequence of three shocks, which can be interpreted as the additional effect on the impulse response of consumption from the presence of two shocks at periods $t = -2$ and $t = -1$, relative to the effect of either in isolation. This effect is quite small (compare the magnitudes on the right and the left panel of figure 4), reflecting fast decay in history terms that is suggestive that a fourth order approximation will likely not add much additional precision.

3.4 Generalized impulse response functions and model moments

Two objects of wide interest in applied research are the model's generalized impulse responses and its ergodic moments. This section shows how to solve for both using the second and third order sequence-space solution.

Generalized impulse response functions. A common way to summarize the behavior of the model is in terms of its impulse responses. We define the generalized impulse response (GIRF) of X as the revision of the forecast of X, j periods ahead due to an innovation to ϵ at date t, $IRF_j^X \equiv \mathbb{E}_t[X_{t+j}] - \mathbb{E}_{t-1}[X_{t+j}]$. This revision generally depends on the size of perceived shocks σ, on the size of the date-t innovation ν, and on the state entering period t, though it does not further depend on the calendar time t beyond these states. If we consider a state-space solution (3), then the state entering date t is X_{t-1},

$$IRF_j\left(\nu, X_{t-1}, \sigma\right) \equiv \mathbb{E}\left[X_{t+j}|\epsilon_t = \nu, X_{t-1}, \sigma\right] - \mathbb{E}\left[X_{t+j}|X_{t-1}, \sigma\right] \quad j \geq 0 \tag{28}$$

while if we consider a sequence-space solution (2), then the state is $\{\epsilon_{t-k}\}_{k=1}^{\infty}$,

$$IRF_j\left(\nu, \{\epsilon_{t-k}\}_{k=1}^{\infty}, \sigma\right) \equiv \mathbb{E}\left[X_{t+j}|\epsilon_t = \nu, \{\epsilon_{t-k}\}_{k=1}^{\infty}\right] - \mathbb{E}\left[X_{t+j}|\left\{\epsilon_{t-k}\right\}_{k=1}^{\infty}\right] \quad j \geq 0 \tag{29}$$

Proposition 5 (GIRFs)**.** *Given history summarized by past shocks* $\{\epsilon_{t-l}\}_{l=1}^{\infty}$*, the second and third order*

GIRFs are given, respectively, by

$$IRF_j^{(2)}(\nu,\sigma) = \mathcal{X}_j\nu + \frac{1}{2}\mathcal{X}_{jj}\left(\nu^2-\sigma^2\right) + \sum_{k>j}^{\infty}\mathcal{X}_{jk}\nu\epsilon_{t-(k-j)} \tag{30}$$

$$\begin{aligned} IRF_j^{(3)}(\nu,\sigma) &= IRF_j^{(2)}(\nu,\sigma) + \frac{1}{2}\left(\mathcal{X}_{\sigma\sigma j} + \sum_{0\le i<j}\mathcal{X}_{iij}\right)\sigma^2\nu + \frac{1}{6}\mathcal{X}_{jjj}\nu^3 + \frac{1}{2}\left(\sum_{k>j}^{\infty}\mathcal{X}_{jjk}\epsilon_{t-(k-j)}\right)\left(\nu^2-\sigma^2\right) \\ &+ \frac{1}{2}\sum_{k>j}^{\infty}\mathcal{X}_{jkk}\epsilon_{t-(k-j)}^2\nu + \sum_{k>j}^{\infty}\sum_{k'>k}^{\infty}\mathcal{X}_{jkk'}\epsilon_{t-(k-j)}\epsilon_{t-(k'-j)}\nu. \end{aligned} \tag{31}$$

Unlike the first order impulse response, the second-order GIRF is both size- and history-dependent. Size dependence comes from the second term in (30) and gives rise to concavity in ν if $\mathcal{X}_{jj} < 0$ and convexity otherwise. The correction $-\sigma^2$ emerges because agents anticipate a non-zero value of X due to future size dependence (see also Andreasen et al. 2018). History dependence comes from the last term in (30) and captures how a past innovation $k-j$ periods ago changes the impulse response to a new shock by an amount proportional to $\mathcal{X}_{jk}$.

The third-order GIRF (31) similarly contains additional terms from the third order perturbation. For instance, the impulse response around 0 is modified by an amount that scales in risk σ^2 due to both time-variation in precautionary savings and mixed size-history dependence.

Moments and moment-based estimation. The ergodic moments of the model are an object of broad interest, in particular for applications to estimation. A typical approach in the literature is to use the moments from the perturbation approximation—for instance, second moments calculated from the $MA(\infty)$ implied by the first-order approximation are routinely used to estimate models via method of moments or likelihood methods (e.g. Auclert et al. 2021). Here, we ask how these calculations relate to the perturbation of the primitive ergodic model moments as a function of σ, and show that we can calculate fifth-order accurate moments using the third order approximation derived above.

In particular, for vector-valued X_t, we are interested in the autocovariance matrix

$$\Sigma_h(\sigma) = \text{Cov}[\mathcal{X}(\sigma,\sigma\bar{\epsilon}_t,\sigma\bar{\epsilon}_{t-1},\ldots),\mathcal{X}(\sigma,\sigma\bar{\epsilon}_{t-h},\sigma\bar{\epsilon}_{t-h-1},\ldots)] \tag{32}$$

Assuming that $\bar{\epsilon}$ has the same moments, up to at least fourth order, as a standard normal distribution,[23] the following proposition characterizes $\Sigma_h(\sigma)$.

Proposition 6 (Ergodic moments)**.** *A fifth-order accurate solution to the autocovariance matrix (32) is*

$$\Sigma_h(\sigma) = \sigma^2\Sigma_{h,2} + \sigma^4\Sigma_{h,4} + O(\sigma^6)$$

[23] This additional assumption is only required for the fourth-order term $\sigma^4\Sigma_{h,4}$. Our original assumption of symmetric $\bar{\epsilon}$ is sufficient for the $\sigma^2\Sigma_{h,2}$ to be accurate to third order.

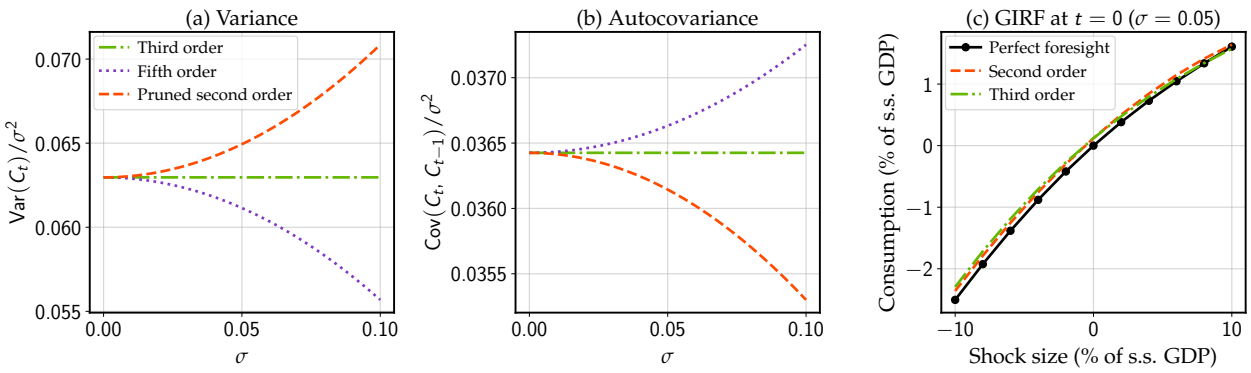


Figure 5: Ergodic moments and GIRF (PE model)

Note: The figures show the variance and first autocovariance of consumption normalized by σ^2, computed using Proposition 6. The third order approximation is computed by setting $\Sigma_{h,4} = 0$, and the pruned second order approximation is obtained by setting third-order terms to zero when computing $\Sigma_{h,4}$. GIRFs are computed as in (31) assuming past shocks $\{\epsilon_{-t}\}_{t=1}^{\infty}$ are zero.

where $\Sigma_{h,2} = \sum_{i=0}^{\infty} \mathcal{X}_{h+i}\mathcal{X}_i^\top$ *is the autocovariance matrix of the first-order perturbation solution, and*

$$\Sigma_{h,4} = \frac{1}{2}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \mathcal{X}_{h+i,h+j}\mathcal{X}_{ij}^\top + \frac{1}{2}\sum_{i=0}^{\infty}\left(\mathcal{X}_{h+i}\mathcal{X}_{\sigma\sigma i}^\top + \mathcal{X}_{\sigma\sigma,h+i}\mathcal{X}_i^\top\right) + \frac{1}{2}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\left(\mathcal{X}_{h+i}\mathcal{X}_{ijj}^\top + \mathcal{X}_{h+i,jj}\mathcal{X}_i^\top\right).$$

Proposition 6 shows, first, that the second-moments calculated from the first-order solution are accurate up to third order in σ. This provides a justification for the typical procedure in the literature. The proposition further offers a correction to the moments that is accurate up to fifth order. This correction involves the second and third derivatives of $\mathcal{X}$, which we have shown how to obtain in sections 3.2 and 3.3 respectively. Notably, however, it does not require *all terms* from the second and third-order solutions: for instance, the triple interactions $\mathcal{X}_{ijk}$ with three distinct indices, which are the most numerous and therefore computationally intensive part of the third-order solution, do not appear.[24]

Figure 5 illustrates these results by visualizing the ergodic variance of consumption $\mathrm{Var}(C_t)$ and the first order autocovariance of consumption $\mathrm{Cov}(C_t, C_{t-1})$, normalized by income shock variance σ^2, for different values of σ. The horizontal blue line shows the third-order accurate approximation to variance—the usual term obtained in the literature, including from typical Dynare output—which is constant once normalized by σ^2. The orange line shows how this term is modified in a fifth-order accurate approximation of the moments. For large aggregate shocks ($\sigma \approx 0.10$), the scaled ergodic variance is around 12% lower, and the ergodic autocovariance around 2% higher. The figure also shows, in dash-dot green, the moments that one would obtain if one naively computed moments from the pruned second-order solution without properly expanding $\Sigma_h(\sigma)$ in σ. Mathematically, this ignores the second and third terms in the expression for $\Sigma_{h,4}$. In

[24] This is in contrast to the more common approach in, for instance, Andreasen et al. (2018), where the entire second- or third-order pruned perturbation solution is used when approximating autocovariances. Our derivation shows that other components of the solution are not needed for fifth-order accuracy—and while they might be needed to higher order, an even higher-order approximation to moments would involve terms from the fourth- and fifth-order solutions.

this example, these pruned second-order moments correct the variance and autocovariance in the wrong direction relative to fifth-order accurate moments.

To conclude, using fifth order accurate moments can make a significant difference in certain applications, and these moments are obtainable by computing only certain parts of the third-order sequence-space solution, making it especially computationally efficient to do so. Armed with these moments, we can estimate model parameters using a generalized method of moments approach: we provide an example of this procedure in section 5.1.

Other applications of the Volterra expansion. The Volterra expansions in (11) and (19) can be useful for other purposes if we are willing to make assumptions on the distribution of the shocks ϵ_t. For instance, in section 5, we assume a normal distribution of shocks and obtain the full ergodic distribution of the second-order solution. Alternative applications may also be possible, though are potentially computationally challenging: those include likelihood-based estimation, extending the method for the first-order solution in Auclert et al. (2021), or filtering unobserved shocks, extending the method for the first-order solution in Rigato (2026).

3.5 Comparison to global solution

A key benefit of our partial equilibrium model is that it is possible to solve it globally using a Bellman equation approach. Indeed, the model in (7) can also be written as:

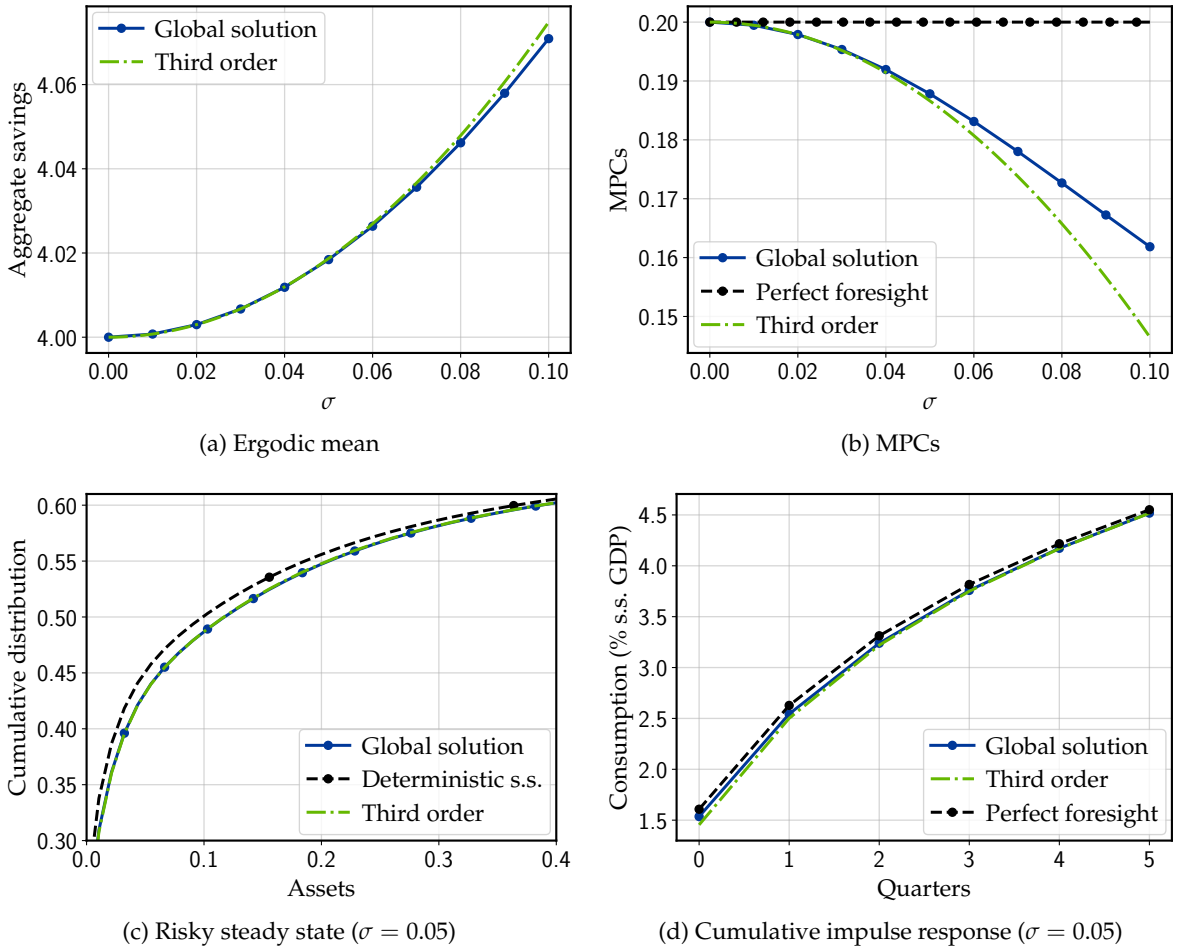
$$\begin{aligned} V(a,e,\beta,Z) &= \mathbb{E}_\xi\left[\max_{c,a'} u(c)+\beta\mathbb{E}[V(a',e',\beta',Z')|e,\beta\,Z]\right] \\ \text{s.t.} \quad & a'+c=(1+r)a+Z\xi e+Tr; \qquad a'\geq \underline{a} \\ & Z'-Z_{ss}=\rho_Z(Z-Z_{ss})+\epsilon \end{aligned} \tag{33}$$

Solving this Bellman equation directly delivers the state space solution $X_t = g(X_{t-1}, \epsilon_t)$. By contrast, our general equilibrium model cannot be cast in this form: as Yang et al. (2025) point out, the market-clearing level of output Y_t (and therefore of post-tax income Z_t) does not have a Markovian structure, so a similar strategy is not available.

We implement the global solution as follows. Since Z_t is i.i.d. in our main calibration, $\rho_Z = 0$, it does not even need to enter as a state variable in the Bellman equation (33). We assume a normal distribution for ϵ with standard deviation σ and compute expectations with respect to ϵ using Gauss-Hermite quadrature with 9 nodes. We then compare our resulting global solution to those from the third order perturbation.

Figure 6 provides this comparison. We first consider how aggregate savings vary with σ in the ergodic mean of the model (panel a), and the distribution of assets in the risky steady state (panel c).[25] Figure 6 shows that our third-order solution is extremely close to the global solution. Panel

[25] We obtain the risky steady state in the global solution by considering the idiosyncratic steady state that results when all agents follow the policy $c(a, e, \beta, Z = Z_{ss})$.



(a) Ergodic mean

(b) MPCs

(c) Risky steady state ($\sigma = 0.05$)

(d) Cumulative impulse response ($\sigma = 0.05$)

Figure 6: Testing global solution vs 3rd order sequence space

Note: Panel (a) shows the ergodic mean of aggregate savings in levels. Panel (b) shows MPCs, defined as $\partial C_0/\partial Z_0$, which for the third-order solution correspond to $\mathcal{X}_0 + \frac{1}{2}\mathcal{X}_{\sigma\sigma 0}\sigma^2$. Panel (c) shows the cumulative wealth distribution for $\sigma = 0.05$. Panel (d) shows the cumulative response of consumption to a one-time income shock of 10% of deterministic s.s. GDP. Perfect foresight starts from the deterministic s.s.; global and third order responses start from their respective risky s.s.

b of figure 6 compares how MPCs fall in σ and finds that up to $\sigma \approx 0.05$, the third-order solution provides an excellent fit. For larger σ, MPCs fall by less in the global solution.

Finally, we consider a notion of impulse response, where in the global model we start from the risky steady state and then consider the effect of the sequence of shocks $(\epsilon_0, \epsilon_1, \epsilon_2, \ldots) = (0.1, 0, 0, \ldots)$.[26] Panel d shows that this impulse response is well approximated by the third order solution, though it undershoots slightly on impact. By contrast, the perfect foresight solution overshoots the global solution on impact as it misses that aggregate risk reduces MPCs.

In summary, the third order perturbation and the global solution get extremely close up to the third order in our PE model, which is comforting. An important question remains to determine how much the global solution adds to the third order perturbation in more complicated, general

[26] This differs from the usual impulse response concept because it does not trace out the expected outcome on average, but is more easily obtained than a generalized impulse response in the global solution.

equilibrium models, for which no standard global method exists.

3.6 How large are nonlinearities?

We have shown that perturbation in the sequence space works well: it approximates the perfect-foresight solution accurately to third (and sometime also to second or first) order, and captures key features of the global solution for a model in which we can compute that solution. This provides a useful path to solving heterogeneous-agent models in which aggregate risk matters more broadly.

In a related context, however, Bianchi and Kaplan (2026) have instead argued that perturbation works poorly, including up to third order, in comparison to the perfect foresight solution. In appendix C we highlight that it is crucial to use a smooth model for higher-order perturbation. A regular model with a discrete income process is not smooth: its distribution is a set of mass points, rather than a continuous and differentiable density as in ours.[27]

In fact, for the experiments studied in Bianchi and Kaplan (2026), we find in appendix D that the third-order perturbation in our smooth model provides an excellent approximation for fiscal shocks of up to 10% of quarterly GDP. Significant errors in the third order approximation relative to the nonlinear perfect foresight solution only appear if most or all of the following are true: 1) transfers are lump-sum instead of proportional, 2) transfers are one-time instead of persistent, 3) one solely focuses on the impact effect instead of a cumulative one over multiple quarters, and 4) transfers are larger than a plausible range for business-cycle fluctuations.

4 Implementation in the sequence space

Our certainty correspondence result showed that a complete third-order solution can be obtained via derivatives of the responses of the economy to up to three consecutive MIT shocks, $\mathcal{X}_j, \mathcal{X}_{jk}, \mathcal{X}_{jkl}$, as well as those of perturbed-expectations steady state and transitions, allowing us to obtain $\mathcal{X}_{\sigma\sigma}$ and $\mathcal{X}_{\sigma\sigma j}$. In this section, we discuss efficient ways to compute these terms. We first describe a "one-shot principle" that allows for such computations in general equilibrium while circumventing nonlinear fixed points. We then explain how to compute the $\mathcal{X}_j, \mathcal{X}_{jk}, \mathcal{X}_{jkl}, \mathcal{X}_{\sigma\sigma}$ and $\mathcal{X}_{\sigma\sigma j}$ terms practically using either numerical or analytical differentiation, discussing the speed involved with each. Finally, we conclude with a comparison to state-space methods. A public code repository illustrates our implementation.[28]

4.1 General equilibrium and the one-shot principle

In many macroeconomic models, including in particular heterogeneous-agent models, the equations governing the perfect foresight solution $\mathcal{X}(0, \epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$ have a special structure, which

[27] Bayer, Briglia, Luetticke, Weiss and Winkelmann (2026a) make a similar point, and propose an alternative solution based on the method in Bayer et al. (2026b).

[28] See `https://github.com/shade-econ/higher-order`.

can be described as follows. First, given $(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$, equilibrium can be reduced to the problem of finding the time path of a small set of "unknown" variables $\mathbf{U} = (U_t, U_{t+1}, U_{t+2}, \ldots)$, where for all $s \geq 0$, $U_{t+s} \in \mathbb{R}^m$ with $m \ll n$, as a solution to a system of equations

$$\mathbf{H}\left(\mathbf{U}, \epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots\right) = 0, \tag{34}$$

where $\mathbf{H} = (H_t, H_{t+1}, H_{t+2}, \ldots)$, and for all $s \geq 0$, $H_{t+s} \in \mathbb{R}^m$. Second, there exists a simple function $\widetilde{\mathcal{X}}(\mathbf{U}, \epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$ such that, given the function $\mathbf{U}\left(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots\right)$ implicitly defined by (34), the complete perfect foresight solution is given by

$$\mathcal{X}\left(0, \epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots\right) = \tilde{\mathcal{X}}\left(\mathbf{U}\left(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots\right), \epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots\right).$$

Both our heterogeneous-agent examples from section 2, for example, can be cast in this form with $m = 1$. In the partial equilibrium model, $U_t = Z_t$ is exogenous after-tax income, and equation (34) takes the simple form $Z_t = Z_{ss} + \sum_{k \geq 0} \rho_Z^k \epsilon_{t-k}$.[29] In the general equilibrium model, $U_t = Z_t$ is endogenous after-tax income, and equation (34) can be written as $\mathcal{A}_t\left(\{Z_s\}\right) - B_t = 0$ at all t, which requires solving a nonlinear fixed point for the time path of $\{Z_t\}$, for any initial condition $(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$ determining both the path of future bonds B_{t+s} and the initial distribution Φ_{t-1}. This type of structure, where blocks of the economy interact via a limited set of m aggregates so that general equilibrium can be reduced to finding the time path of these aggregates, is the characteristic of the sequence space as described, for instance, in Auclert et al. (2021).

Nonlinear solutions in the sequence space, however, are computationally cumbersome: an exact solution to the perfect-foresight system (34) requires a nonlinear fixed point for any choice of $(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$. Fortunately, a perturbation of order k only requires accuracy of up to order k in the vector of ϵ's, and this can be obtained without a nonlinear fixed point, as we now show.

Let $\boldsymbol{\epsilon} = (\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$. We say that a function $\mathbf{U}^{[k]}(\boldsymbol{\epsilon})$ approximates $\mathbf{U}(\boldsymbol{\epsilon})$ to order k if it is C^k and satisfies

$$\mathbf{U}^{[k]}(\boldsymbol{\epsilon}) = \mathbf{U}(\boldsymbol{\epsilon}) + o(||\boldsymbol{\epsilon}||^k).$$

The following result shows how to use such an order-k approximation to the unknowns to construct, in a single step, an approximation of the whole sequence space solution of order $k+1$.

Proposition 7 (One-shot principle). *Let $\mathbf{H}_U$ denote the sequence-space Jacobian of the $\mathbf{H}$ function at the steady state $(\mathbf{U}_{ss}, 0, 0, \ldots)$. Assume that $\mathbf{H}$ and $\widetilde{\mathcal{X}}$ are sufficiently smooth, that $\mathbf{H}_U$ has a bounded inverse, and that $\mathbf{U}^{[k]}(\boldsymbol{\epsilon})$ approximates $\mathbf{U}\left(\boldsymbol{\epsilon}\right)$ to order k for some $k \geq 0$. Then, defining $\mathbf{U}^{[k+1]}(\boldsymbol{\epsilon})$ and $X_t^{[k+1]}(\boldsymbol{\epsilon})$ as*

$$\begin{aligned} \mathbf{U}^{[k+1]}(\boldsymbol{\epsilon}) &= \mathbf{U}^{[k]}(\boldsymbol{\epsilon}) - \mathbf{H}_U^{-1}\mathbf{H}\left(\mathbf{U}^{[k]}(\boldsymbol{\epsilon}), \boldsymbol{\epsilon}\right), \\ X_t^{[k+1]}(\boldsymbol{\epsilon}) &= \tilde{\mathcal{X}}\left(\mathbf{U}^{[k+1]}(\boldsymbol{\epsilon}), \boldsymbol{\epsilon}\right), \end{aligned} \tag{35}$$

we have that $\mathbf{U}^{[k+1]}(\boldsymbol{\epsilon})$ approximates $\mathbf{U}(\boldsymbol{\epsilon})$ to order $k+1$, and $X_t^{[k+1]}\left(\boldsymbol{\epsilon}\right)$ approximates the nonlinear

[29] In the iid shock case with $\rho_Z = 0$, this step isn't even necessary, and we can set $m = 0$.

perfect-foresight sequence space solution to order $k+1$, in the sense that

$$X_t^{[k+1]}(\boldsymbol{\epsilon}) = \mathcal{X}(0, \boldsymbol{\epsilon}) + o(||\boldsymbol{\epsilon}||^{k+1})$$

Concretely, for given $\boldsymbol{\epsilon}$, we can construct a sequence of order-k approximations $\mathbf{U}^{[k]}(\boldsymbol{\epsilon})$ by forming $\mathbf{U}^{[0]}(\boldsymbol{\epsilon}) = \mathbf{U}_{ss}$ and applying (35) recursively k times. Alternatively, we can construct $\mathbf{U}^{[k]}(\boldsymbol{\epsilon})$ as a Taylor polynomial of order k in $\boldsymbol{\epsilon}$, and use equation (35) to get the order-$k+1$ derivative.

The idea behind Proposition 7 is that (35) is a quasi-Newton iteration with the Jacobian fixed at the steady state, which improves the accuracy of any error by one order in each step—hence the name "one-shot principle". A concrete implication of this result is that we can replace, in the calculation of kth order derivatives of $\mathcal{X}$, all nonlinear fixed points involving an unknown and possibly large number of evaluations of $\mathbf{H}$ by only k nonlinear evaluations of $\mathbf{H}$, each re-equilibrated by the application of a pre-constructed inverse Jacobian $\mathbf{H}_U^{-1}$.[30] This will turn out to be critical in implementations of our method that rely on numerical differentiation, as described further below.

A similar principle applies to steady state calculations. The equations for the perturbed expectations steady state (16), specialized to finding the unknowns U, are given by $H(U, \Delta) = 0$ where $H, U \in \mathbb{R}^m$, and Δ is the expectational shifter with $\Delta = 0$ corresponding to the deterministic steady state U_{ss}. Then, given the Jacobian H_U evaluated at that steady state, for any Δ we can obtain steady-state unknowns accurately to order k via the recursions $U^{[k]} = U^{[k-1]} - H_U^{-1} H(U^{[k-1]}, \Delta)$ starting from $U = U_{ss}$. This is relevant for practical implementation of routines to obtain the risky steady state $\mathcal{X}_{\sigma\sigma}$ and the time-varying risk term $\mathcal{X}_{\sigma\sigma j}$, since those involve perturbations Δ to the steady state routines.

4.2 Practical implementation

In practical implementations of algorithms, the function $\widetilde{\mathcal{X}}$ described in the previous section takes the form of a function $\tilde{X}_j^{MIT}(\nu, X_-, \mathbf{U})$ that, given a shock size ν, an inherited state X_-, and a set of relevant aggregate variables $\mathbf{U}$, gives us the date-j outcome of the economy after a shock at date 0. Using these functions, we can compute the nonlinear responses to one, two, or three shocks, $j > 0$ periods after the latest shock, with a lag of $l > 0$ to the second shock and a lag of $l' > l$ to the oldest shock, as:

$$X_j^{MIT}(\nu) = \tilde{X}_j^{MIT}(\nu, X, \mathbf{U}(\nu, X)) \tag{36}$$

$$X_j^{MIT,l}(\nu, \epsilon) = \tilde{X}_j^{MIT}\left(\nu, X_{l-1}^{MIT}(\epsilon), \mathbf{U}\left(\nu, X_{l-1}^{MIT}(\epsilon)\right)\right) \tag{37}$$

$$X_j^{MIT,l,l'}(\nu, \epsilon, \epsilon') = \tilde{X}_j^{MIT}\left(\nu, X_{l-1}^{MIT,l'-l}(\epsilon, \epsilon'), \mathbf{U}\left(\nu, X_{l-1}^{MIT,l'-l}(\epsilon, \epsilon')\right)\right) \tag{38}$$

[30] This idea is closely related to that, in state-space perturbation, of using only linear operations to solve for higher-order terms once the first-order law of motion is known (e.g. Schmitt-Grohé and Uribe 2004, Kim et al. 2008).

To get the effect on the economy from a single shock, we start from the steady state X, calculate the unknowns $\mathbf{U}(\nu, X)$ implied by the system (34), and then use the $\tilde{X}_j^{MIT}$ function to trace out the perfect-foresight evolution out to period j. To evaluate two consecutive MIT shocks separated by l periods, we first solve the MIT response to the earlier shock ϵ and use the resulting date-l state $X_{l-1}^{MIT}(\epsilon)$ as the inherited state when the later shock ν arrives, as expressed by equation (37). To get the effect of three consecutive surprise shocks, with the first two separated by $l' - l$ and the next two by l periods, the logic is similar: we start from the inherited state $X_{l-1}^{MIT,l'-l}(\epsilon, \epsilon')$ after l periods of two shocks ϵ, ϵ' that were $l' - l$ periods apart, and trace out the effect of a completely new shock ν from that point onwards, as expressed by equation (38).

This approach allows us to numerically evaluate the derivatives $\mathcal{X}_j, \mathcal{X}_{jk}, \mathcal{X}_{jkl}$ required for the certainty correspondence up to third order. Critically, we can use Proposition 7 to replace, in the calculation of derivatives of order k, the nonlinear $\mathbf{U}$ by its k-th order approximation $\mathbf{U}^{[k]}$ in (36)–(38). For instance, to calculate the size-dependence term $\mathcal{X}_{jj}$ for any $j \geq 0$, we can use:

$$\mathcal{X}_{jj} = \frac{d^2}{d\nu^2} \tilde{X}_j^{MIT} \left(\nu, X, \mathbf{U}^{[2]}(\nu, X) \right) \Big|_{\nu=0} \tag{39}$$

since this is a second derivative and therefore requires second-order correct unknowns.

Likewise, to get $\mathcal{X}_{\sigma\sigma}$, we can follow a similar procedure, with (17) combined with the steady-state version of the one-shot principle implying that we can obtain the unknowns in the risky steady state as

$$U_{\sigma\sigma} = -H_U^{-1} \frac{\partial H \left(U_{ss}, \mathcal{X}_{00}\delta \right)}{\partial \delta} \tag{40}$$

from which the $\mathcal{X}_{\sigma\sigma}$ follows. Similar procedures, described in appendix E.2, apply to obtain the remaining second and third order derivatives of $\mathcal{X}$.

We conclude by noting that we typically do not need to compute or store every element of $\mathcal{X}$. For example, our heterogeneous-agent applications, we may be interested in keeping track of only a few aggregates such as post-tax income Z_t, GDP Y_t, consumption C_t, taxes T_t and bonds B_t (i.e. the elements of endogenous unknowns U_t plus a few more), rather than the full evolution of the marginal value function $\boldsymbol{v}_t$ and the distribution $\boldsymbol{\phi}_t$. This lowers the number of derivatives that need to be computed and kept track of to the number of aggregates we care about. To obtain the risk terms for these objects, we need to compute the larger $\mathcal{X}_{00}$ and $\mathcal{X}_{00j}$, but not the full $\mathcal{X}_{jk}$ or $\mathcal{X}_{jkk'}$ terms, and only for the part of the X_t that enters the F function via expectations. In our examples, this means computing impact size-dependence and mixed history-impact size for the marginal value $\boldsymbol{v}_t$, but not for the distribution $\boldsymbol{\phi}_t$.

4.3 Choice of truncation

In practice, we need to truncate our solution in two ways. First, for solving any individual MIT shock impulse response, we calculate up to some horizon T following the most recent shock. The choice of T is discussed at length in Auclert et al. (2021), and we do not dwell on it here: it must

be high enough that there is convergence back to the steady state to a tight numerical tolerance. For our applications here, we use $T = 300$, which is sufficient.[31]

Second, we must similarly truncate history dependence. In particular, when calculating (37)–(38), we choose some terminal lag L, where $0 < l < l' < L$. From these MIT shocks, we can calculate second-order coefficients of the form $\mathcal{X}_{j,j+l}$, for $0 \leq j < T$ and $0 \leq l < L$, and third-order coefficients of the form $\mathcal{X}_{j,j+l,j+l'}$, for $0 \leq j < T$ and $0 \leq l, l' < L$. These give us truncated approximations to the second- and third-order expansions in (11) and (19).

To avoid truncation error, we must choose L high enough that the truncated approximation is a good approximation to the overall expansion. There are several ways in which this can be measured, and we will focus on two. First, we ask: what is the average error in a *simulated* nth order approximation $X_t^{(n)}$, for random ϵ_{t-j}, from truncating the series? Second, we ask: what is the error in calculating *moments* from using a truncated series?

Figure 7(a) shows both notions of error for the second-order approximation to the IKC model as a function of L, using a very long history at $L = 300$ as the approximate no-truncation benchmark. The orange line shows the root mean square error for simulated $X_t^{(2)}$, while the green line shows the truncation error for the analytical standard deviation using proposition 6, with the third-order terms dropped.[32] We also plot the error for the analytical $h = 5$ autocovariance (purple), and the largest coefficient omitted by truncation (blue) for reference. We assume $\sigma = 0.05$, as in figure 6, and normalize errors by the standard deviation of the first-order solution, so that the errors in 7(a) are relevant to typical movement in the series.[33]

All errors decay steadily on a log scale as we increase L, with a temporary dip between $L = 100$ and $L = 150$, corresponding to a sign change in the GE impulse response. Interestingly, the error in calculating moments decays much more quickly than the average error in the simulated series; this is because the second-order terms in proposition 6 enter quadratically, so that smaller terms corresponding to high l quickly become irrelevant.

Of course, the $L = 300$ benchmark in panel (a) is often not available when choosing L in practice. Panel (b) therefore constructs an internal estimate of the remaining second-order error, using only coefficients calculated through the current L. We compare the errors contributed by the two most recent blocks of 20 lags, extrapolate the rate of decay between them, and sum the implied remaining tail. (See appendix E.3 for details.) At $L = 100$, this procedure produces an estimate close to, though somewhat larger than, the benchmark error in panel (a)—suggesting that the extrapolation is, if anything, slightly conservative.

Panel (c) applies the same idea to truncation of the third-order solution, holding the second-order solution fixed at $L = 100$. Here, again, there is a steady decline in the error in log terms as L increases, and errors in moments are smaller than average errors in the simulated series—though

[31] T must be quite large because increasing T affects the solution at every date before T, via the inverse $T \times T$ Jacobian $\mathbf{H}_U^{-1}$. By contrast, increasing L does not change the solution for lags before L, making for a simpler accuracy check.

[32] For a precise estimate of root mean square error, we calculate it analytically rather than from a Monte Carlo simulation. We assume that ϵ_t have the same moments as a standard normal; see appendix E.3 for details of the calculation.

[33] To keep units consistent, the autocovariance is normalized by the variance of the first-order solution.

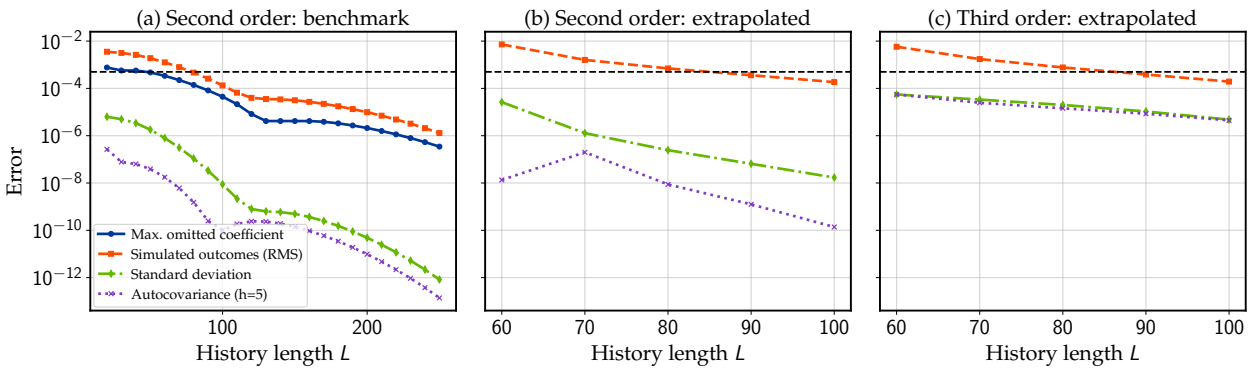


Figure 7: Truncation error in IKC (GE) solution by history length L

now they decay at a similar rate, because third-order terms enter linearly in proposition 6.

The dashed horizontal line gives our error target of 5×10^{-4} for each order, which allows the combined error from both orders to remain below 10^{-3}. All diagnostics are comfortably below this target at $L = 100$, motivating our use of $L = 100$ throughout the paper.

4.4 Numerical differentiation

The most straightforward approach to computing derivatives such as (39) and (40) is numerical differentiation. For instance, we use numerical differentiation and the fake news algorithm of Auclert et al. (2021) to form the Jacobian $\mathbf{H}_U$. This gives us the first-order solution for unknowns $\mathbf{U}^{[1]}(h)$, and we can then obtain the $\mathcal{X}_j$'s via first-order symmetric difference:

$$\mathcal{X}_j \simeq \frac{X_j^{MIT}(h) - X_j^{MIT}(-h)}{2h} \tag{41}$$

Moreover, to obtain (39), we compute the central symmetric difference:

$$\mathcal{X}_{jj} \simeq \frac{X_j^{MIT}(h) + X_j^{MIT}(-h) - 2X_{ss}}{h^2} \tag{42}$$

for an appropriate choice of h, and using the one-shot principle for the underlying calculation of $\mathbf{U}$.[34] For mixed derivatives such as $\mathcal{X}_{jk}$ for $k > j$, a similar approach is also feasible. However, given that we are typically computing $\mathcal{X}_{jj}$ and $\mathcal{X}_{kk}$ anyway, a faster approach is to use the so-called *polarization identity*:

$$\mathcal{X}_{jk} \simeq \frac{X_j^{MIT,k-j}(h,h) + X_j^{MIT,k-j}(-h,-h) - 2X_{ss}}{2h^2} - \frac{\mathcal{X}_{jj} + \mathcal{X}_{kk}}{2} \tag{43}$$

[34] In other words, we implement $\mathcal{X}_{jj} \simeq \frac{\tilde{X}_j^{MIT}(h, X_{ss}, \mathbf{U}^{[2]}(h, X_{ss})) + \tilde{X}_j^{MIT}(-h, X_{ss}, \mathbf{U}^{[2]}(-h, X_{ss})) - 2X_{ss}}{h^2}$. In practice, we pick $\mathbf{U}^{[2]}(h, X_{ss}) = \mathbf{U}_{ss} + \mathbf{U}_v h + \mathbf{U}_{vv}\frac{h^2}{2}$.

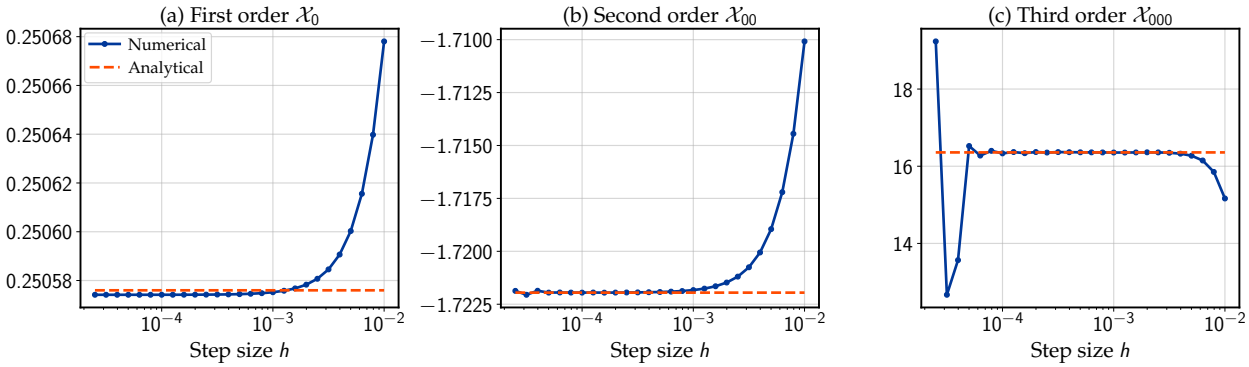


Figure 8: Testing step size h in numerical differentiation for impact output in IKC model.

This requires half as many evaluations of the $X_j^{MIT,l}$ function, compared with a direct formula that would use all four corners $(\pm h, \pm h')$, and is therefore about twice as fast. We use similar principles for the risk terms and the third numerical derivatives $\mathcal{X}_{jkl}$, as described in appendix E.4.

Choice of h and accuracy. The optimal choice of step size h in numerical differentiation, such as in (41) and (43), balances error from higher-order terms that contaminate the result when h is too high against error from roundoff and other noise that becomes inflated when h is too low. To choose h, we recommend recomputing some derivatives for a range of h, and finding the h at which the results appear most stable. Figure 8 shows an example of this procedure, where we recalculate size effects at first, second, and third order for various h. For this application, we find that an appropriate h is 10^{-4} for first order, 2×10^{-4} for second order, and 10^{-3} for third order.

Local stability with respect to step size h is also an important test of whether numerical differentiation is working properly. In our case, this works because the underlying model is smooth—but in nearly identical models without the same smoothness, the second and third numerical derivatives can be highly unstable, as we show in figure C.7.

Computational speed. The first column of table 2 displays computation speed on an ordinary laptop for implementing a second and third order solution this way. The table first reports the time spent in pre-computation, to obtain the model steady state as well as other setup such as calculating the steady-state Jacobian H_U and the sequence-space Jacobian $\mathbf{H}_U$, which we do using the fake news algorithm in Auclert et al. (2021). The table then reports times to compute the second and the third-order solution for the two heterogeneous-agent models in section 2. In addition to times used to obtain the full second and third order solution, which require respectively all terms $(\mathcal{X}_{jk}, \mathcal{X}_{\sigma\sigma})$ and $(\mathcal{X}_{jkl}, \mathcal{X}_{\sigma\sigma j})$, we also report the time required to compute the ergodic moments implied by these solutions. As shown in proposition 6, those only involve a subset of the $\mathcal{X}$ derivatives.

A few observations stand out. First, solving the general equilibrium takes almost exactly the same time as solving the partial equilibrium. This is due to the one-shot principle, which calcu-

Model	Object to compute	Type of differentiation	
		Numerical	Analytical
Pre-computation *(needed for PE & GE)*	Steady state	4.0 s	4.0 s
	Other setup	0.0 s	1.2 s
	First-order solution	10.2 s	0.3 s
SIM model (PE)	RSS and ergodic mean	16.0 s	0.5 s
	Second-order solution	4.7 m	10.6 s
	Moments to fifth order	12.7 m	14.7 s
	Third-order solution	3.7 h	4.5 m
IKC model (GE)	RSS and ergodic mean	15.3 s	0.4 s
	Second-order solution	4.7 m	10.7 s
	Moments to fifth order	15.4 m	16.8 s
	Third-order solution	3.9 h	4.4 m

Note: Times are cumulative within each block (pre-computation, SIM model and IKC model).

Table 2: Computation speed for sequence space solution

lates the endogenous unknowns using only linear operations in addition to the work done for partial equilibrium. Second, the time taken depends mostly on how many $\mathcal{X}$ derivatives are calculated. For instance, calculating the RSS and ergodic mean only requires computing a single size-dependent impulse response of $\mathcal{X}_{jj}$ and the resulting $\mathcal{X}_{\sigma\sigma}$, while calculating the full second-order solution involves a much larger number of history terms $\mathcal{X}_{jk}$. Similarly, using proposition 6 to calculate fifth-order accurate moments requires only mixed size-history terms of the form $\mathcal{X}_{jkk}$ and the resulting $\mathcal{X}_{j\sigma\sigma}$, rather than the much larger number of triple interactions $\mathcal{X}_{jkl}$.

Most of the times reported here are quite practical, but the full third-order solutions are costly, at roughly 4 hours. These solutions are only necessary, however, if we want to explicitly simulate the third-order solution in for a history of shocks of length at least L.[35] If we are only interested in fifth-order accurate moments that summarize the solution, however, then a much shorter calculation suffices.

4.5 Analytical differentiation

While numerical differentiation is always available, we can achieve significant gains in speed, and also slightly higher accuracy, by using analytical differentiation. The general idea is to systematically apply the chain rule to our nonlinear routines to obtain derivatives, as in automatic differentiation (Griewank and Walther 2008).

Rather than using a canned automatic differentiation package for these computations, however, we manually differentiate each step of our routines and write code to accumulate the derivatives. This is for several reasons. First, the economics of the model simplify some of the deriva-

[35] If we only wanted to simulate the third-order solution in response to a shorter history of shocks of length $L' < L$, then we could truncate the triple interaction terms at L', and reduce computational time by approximately $(L'/L)^2$.

tives: for instance, the envelope theorem typically governs backward propagation of the value function. Second, for higher derivatives, singularities in the model can contribute additional terms: for instance, in our smooth SIM model, the second-order solution includes an interaction between the moving kink in the policy function and the incoming smooth density of households (see also Bhandari et al. 2023 and Bianchi and Kaplan 2026). Finally, we can speed up computation by precomputing steady-state objects that are used repeatedly. We provide more details in the context of our smooth SIM and IKC models in appendix E.5, illustrating each of these three ideas.

The second column of table 2 shows the resulting improvements in computational speed. The steady state is identical, after which we precompute objects that are used in the rest of the procedure in about a second. Times for our first-, second-, and third-order calculations are then a factor of 30 to 60 below the corresponding times with numerical differentiation. This makes even the full third-order solution much more practical, at less than 5 minutes, and all other parts of the solution finish in seconds.

These gains are large because analytical differentiation replaces repeated nonlinear calculations, as needed for (41)–(43), with purely linear calculations with mostly precomputed coefficients. For instance, the steady-state marginal propensity to save governs the first-order transmission from tomorrow's marginal value function to today's.[36]

The expressions for analytical derivatives become progressively more complicated with higher order, as we illustrate in appendix E.5, and in isolation it is easy to make mistakes. Crucially, however, our smooth model makes it possible to use numerical differentiation as a benchmark for analytical derivative—both for the ultimate results and, when debugging, for intermediate derivatives. This straightforward numerical validation provides a check against human error, and is also particularly valuable when asking AI agents to write code for analytical derivatives. One example of this validation is in figure 8, where analytical derivatives agree almost perfectly with numerical derivatives when the latter are in the stable region.

Overall, for the second order, our analytical differentiation procedure is similar to that proposed by Bhandari et al. (2023): we precompute steady state objects, obtain the higher-order solution by using the first-order system using directional derivatives that involve the lower-order solution as forcing variables (due to what we call the one-shot principle), and take care of movements of kinks in policies and the mass of agents at the constraint using Leibniz's rule. Our contribution here is to extend these ideas to the third order and to offer numerical differentiation on a smooth model as a benchmark.

Further speedup by approximating with exponential mixtures. In applications where we need to repeatedly calculate a full third-order solution, or where there is a larger model with multiple shocks as in section 5.1, it may be useful to further improve on the times in table 2. In appendix E.6, we propose a method for this, which relies on the fact that the first-order impulse response of unknowns $\mathbf{U}$ will often be very well-approximated by a mixture of exponentials. For instance,

[36] See, e.g., Farhi, Olivi and Werning (2022), Bilal and Goyal (2025), or Auclert et al. (2024a).

Sequence-space			State-space	
Perturbation	Key object	Size	Key object	Size
First order	First-order impulse response	T	First-order LoM	n^2
Second-order	Precautionary effect	1	Second-order LoM	n^3
	Size dependence impulse	T		
	History dependence impulse	TL		
	Total	TL	*Total*	n^3
Third-order	Precautionary impulse	T	Third-order LoM	n^4
	Mixed size-history impulses	TL		
	History interaction impulses	TL^2		
	Total for 5th-order moments	TL		
	Total for full solution	TL^2	*Total*	n^4

Table 3: Dimensionality of sequence-space vs. state-space perturbation

in the IKC model, we can represent the first-order path of dZ_t to six digits of accuracy with a mixture of seven exponentials. Rewriting the first-order response in terms of this mixture allows us to derive relatively low-dimensional recursions for the second- and third-order solutions. As we show in appendix E.6, this reduces the time relative to the analytical solution in table 2 by a further factor of about 35, with a negligible loss in accuracy that is closely tied to the accuracy of the first-order exponential mixture approximation.

4.6 Dimensionality and sequence space vs. state space

We conclude this section by discussing the dimension of the sequence-space solution: that is, the number of terms that need to be computed and stored in memory. This serves as a complement to our speed calculations, and is useful in part because it does not depend on a particular computer or implementation. It also allows for a comparison with the state-space perturbation approach.

The left side of table 3 shows how many coefficients are needed at each order to describe the sequence-space solution for a single outcome variable and a single aggregate shock. This involves the truncation horizon for impulse responses T, as well as the truncation horizon for history terms L. Recall that for our heterogeneous-agent model we have found that $T = 300$ and $L = 100$ provide us with the accuracy we need. Given these values, a second-order perturbation requires calculating approximately 30,000 terms, and a third-order perturbation approximately 3 million.

As already discussed, and validated on the neoclassical growth model in appendix B.6, a state-space perturbation solution gives the same responses as our sequence-space solution at the same order. However, the two approaches have very different storage requirements, as illustrated on the right side of table 3, which reports the number of coefficients in the state-space law of motion for n variables.

In our model, the household distribution has approximately $n = 2800$ entries: 200 asset grid

points, 7 productivity states, and 2 patience types. A full, dense state-space law of motion would therefore require roughly 8 million coefficients at first order, 22 billion at second order, and 61 trillion at third order. Simply storing the third-order solution would be prohibitive. By comparison, our third-order sequence-space solution requires about 3 million coefficients for a single aggregate outcome, regardless of the dimension of the household distribution—much more practical in this case.

Speed. Finally, we return to the question of speed, and ask how these differences between sequence and state space extend to computation time, in models where both approaches are feasible. In appendix E.7, we compare the two approaches in an N-country version of our neoclassical growth model, varying N to change the size of the model. At third order, the state-space solution calculated by Dynare is faster for small N, but its computation time grows more rapidly with N. Sequence space becomes faster at around 60 countries, corresponding to roughly 200 variables.

5 Applications

In this section, we apply the methods introduced above to solve more complex models of the type that arise in practical quantitative work. The first is a richer version of the general equilibrium HANK model analyzed earlier in the paper. The second is a state-of-the-art menu cost model of price setting.

5.1 Quantitative HANK model

We expand upon the HANK model from section 2.2 by incorporating additional model ingredients and shocks. Specifically, we allow for a Taylor rule for the nominal interest rate, and a Phillips curve relating inflation to economic activity, which brings about new sources of nonlinearities. We also add government spending and wage markup shocks.

In this updated model, the central bank follows a Taylor rule for the nominal interest rate:

$$1 + i_t = (1 + r)\,(1 + \pi_t)^{\phi} \tag{44}$$

We maintain our assumption that households are only able to invest in real bonds. However, the model also features risk-neutral arbitrageurs, who have the the possibility of investing in either nominal or real bonds. In equilibrium these arbitrageurs make no profits, and their portfolio choice optimality condition implies that the real interest rate promised at t on real bonds paying at $t+1$ is

$$1 + r_t = \mathbb{E}_t\left[\frac{1 + i_t}{1 + \pi_{t+1}}\right] \tag{45}$$

We allow for shocks to government spending G_t in addition to bonds B_t. Given the time-

varying real interest rate on bonds, the government budget constraint is now

$$B_t + \tau_t Y_t = (1 + r_{t-1}) B_{t-1} + Tr + G_t \tag{46}$$

The Taylor rule implies that inflation is now relevant for the determination of aggregate demand. We assume Rotemberg wage setting, as in Auclert et al. (2024a), and further assume that unions evaluate welfare using average consumption, using the average discount factor across households $\bar{\beta} \equiv (\beta_L + \beta_H)/2$, and that they are subject to exogenous markup shocks μ_t. Motivated by the widespread empirical observation that prices are more flexible than wages, we continue to assume flexible prices, implying that the real wage is constant at 1. This combination of assumptions leads to the nonlinear wage Phillips curve:

$$\pi_t (1 + \pi_t) = \kappa N_t \left(v'(N_t) - \frac{\epsilon - 1}{\epsilon} \frac{1 - \tau_t}{\mu_t} u'(C_t) \right) + \bar{\beta} \mathbb{E}_t \left[\pi_{t+1} (1 + \pi_{t+1}) \right] \tag{47}$$

Compared to our simple IKC model, this model contains new sources of nonlinearities due to the presence of inflation in the denominator in (45) and entering as a square in (47). To take this into account, we define the model by keeping track of $\frac{1}{1+\pi_t}$ and $\pi_t(1 + \pi_t)$ as separate variables in X_t, so that its equations can take the form in in (1). We consider a standard calibration with $\phi = 1.5$, $\kappa = 0.01$, an isoelastic disutility of labor with Frisch elasticity of 0.5 ($v'(N) = \zeta N^2$), and a steady state inverse wage markdown $\frac{\epsilon}{\epsilon - 1} = 1.1$ and wedge $\mu = 1$.

We assume that the model is hit by exogenous shocks to government spending G_t, bonds B_t and inverse markups μ_t^{-1}.[37] We model these as AR(1) processes in levels, with persistences (ρ_G, ρ_B, ρ_p) and innovations standard deviations $(\sigma_G, \sigma_B, \sigma_p)$ respectively. The government spending shock G_t has the interpretation of a general aggregate demand shock emanating outside the household sector, while the markup shock μ_t has the interpretation of a general supply shock, pushing up inflation at a given level of activity.

Estimation. We first estimate the model's shock processes using a generalized method of moments. We assume no fiscal shocks, $\sigma_B = 0$, as those tend to be small in the data, and set $\rho_B = 0.95$, as in our IKC example from section 2. We then estimate the four other shock parameters, $\rho_G, \sigma_G, \rho_p, \sigma_p$ by GMM with five moments: the first two autocovariances of output and inflation, as well as the contemporaneous covariance between output and inflation.[38] We construct the model moments using the fifth order correct moments from section 3.4.

This procedure delivers the parameter estimates listed in table 4. We find demand shocks to be quite persistent ($\rho_G \simeq 0.90$), with supply shocks a little less persistent ($\rho_p \simeq 0.75$). Both shocks are important, but demand shocks account for approximately 76% of the variance of total output;[39]

[37] We model shocks to inverse markups rather than markups to avoid the nonlinearity from the shock μ_t itself entering in the denominator of (47)

[38] Appendix F.1.8 describes the GMM formulas and the data used in the estimation.

[39] This is defined as the fifth-order approximation of $\text{Var}(Y_t)$ setting $\sigma_p = 0$ relative to what is obtained with the estimated σ_p. The contribution of the supply shock to total output variance is 22%, with the remaining 2% corresponding

Moment	Data	Model	Parameter	Estimate (s.e.)
$\text{Var}(Y_t)$	4.99631	4.82543	ρ_G	0.90338 (0.03041)
$\text{Cov}(Y_t, Y_{t-1})$	4.56912	4.24239	σ_G	0.00890 (0.00089)
$\text{Var}(\pi_t)$	0.41064	0.43022	ρ_p	0.74933 (0.06833)
$\text{Cov}(\pi_t, \pi_{t-1})$	0.33420	0.35290	σ_p	0.14089 (0.04811)
$\text{Cov}(Y_t, \pi_t)$	0.12010	0.13740		

Table 4: Moments and estimated parameters

Note: Y_t in the data is constructed as real GDP divided by potential GDP. π_t in the data is PCE inflation. Units of σ_G are in percentage points of GDP, σ_p is in percentage points of annualized inflation, and all moments are expressed in squares of percentage points.

this is necessary to match the observed positive covariance between inflation and output.

Risky steady state. Figure 9 visualizes the risky steady state in our model as a function of the overall standard deviation of output σ_Y, keeping constant the estimated ratio σ_G/σ_p. Contrary to the partial equilibrium model where higher σ raised aggregate asset demand in the risky steady state (see figure 3), aggregate assets do not adjust here as bond supply is fixed in the risky steady state. Instead, greater precautionary saving is associated with lower interest rates, which help the government budget and imply that disposable income and GDP rise. In turn, inflation is lower, reflecting a positive supply-side effect of lower taxes in the wage Phillips curve. For our estimated parameter values, these effects are all relatively mild.

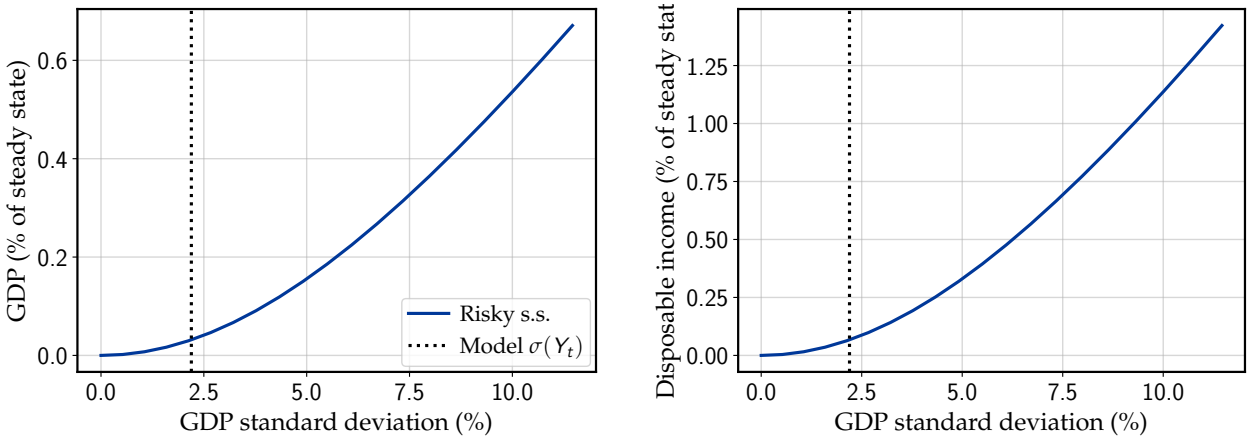
Size dependence. Figure 10 presents the impact ($t = 0$) and delayed ($t = 1$) effects of a deficit-financed bond shock on output in our quantitative HANK model as a function of the size of the shock ν, presented as in figure 2 so that it is directly comparable to the perfect foresight solution.[40] The concave and convex patterns for the two effects (panels a and b) are similar to those we observed for the simple PE and GE models from section 2.2. These terms of the third-order solution continue to be an excellent approximation of the perfect-foresight solution.

Panel c presents the effects on inflation in period $t = 0$. Here, we find modest nonlinearities to second and third order, despite multiple sources of nonlinearities in the model. This is because inflation is a forward-looking average of future aggregate demand, so that the concavity in aggregate demand in $t = 0$ is partially offset by the convexity in later periods (see appendix D for a further discussion of this point). This effect goes away in $t = 1$ resulting in convexity becoming dominant, as shown in panel d.

These results are dominated by the nonlinearities of the household block; the ones due to the presence of nonlinear terms in the Fisher equation (45) and the Phillips curve (47) do not play a

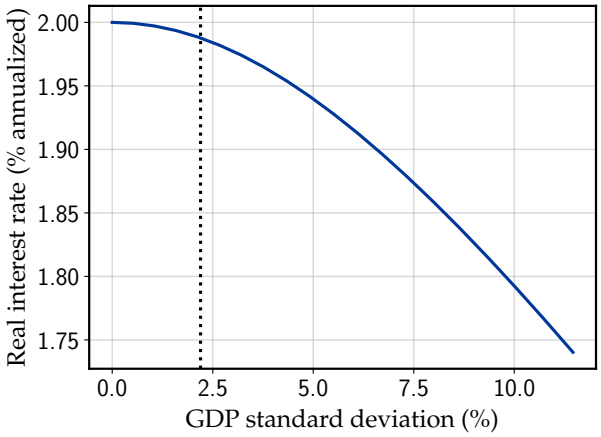
to their interaction.

[40] As proposition 5 shows, the generalized impulse responses would look slightly different due to the expected size dependence, which we do not include in figure 10. Quantitatively, expected size dependence would be relatively small for the range of shock size shown in figure 10.

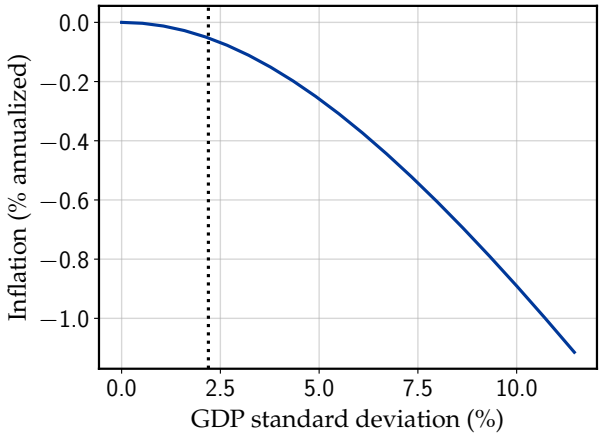


(a) GDP

(b) Disposable income



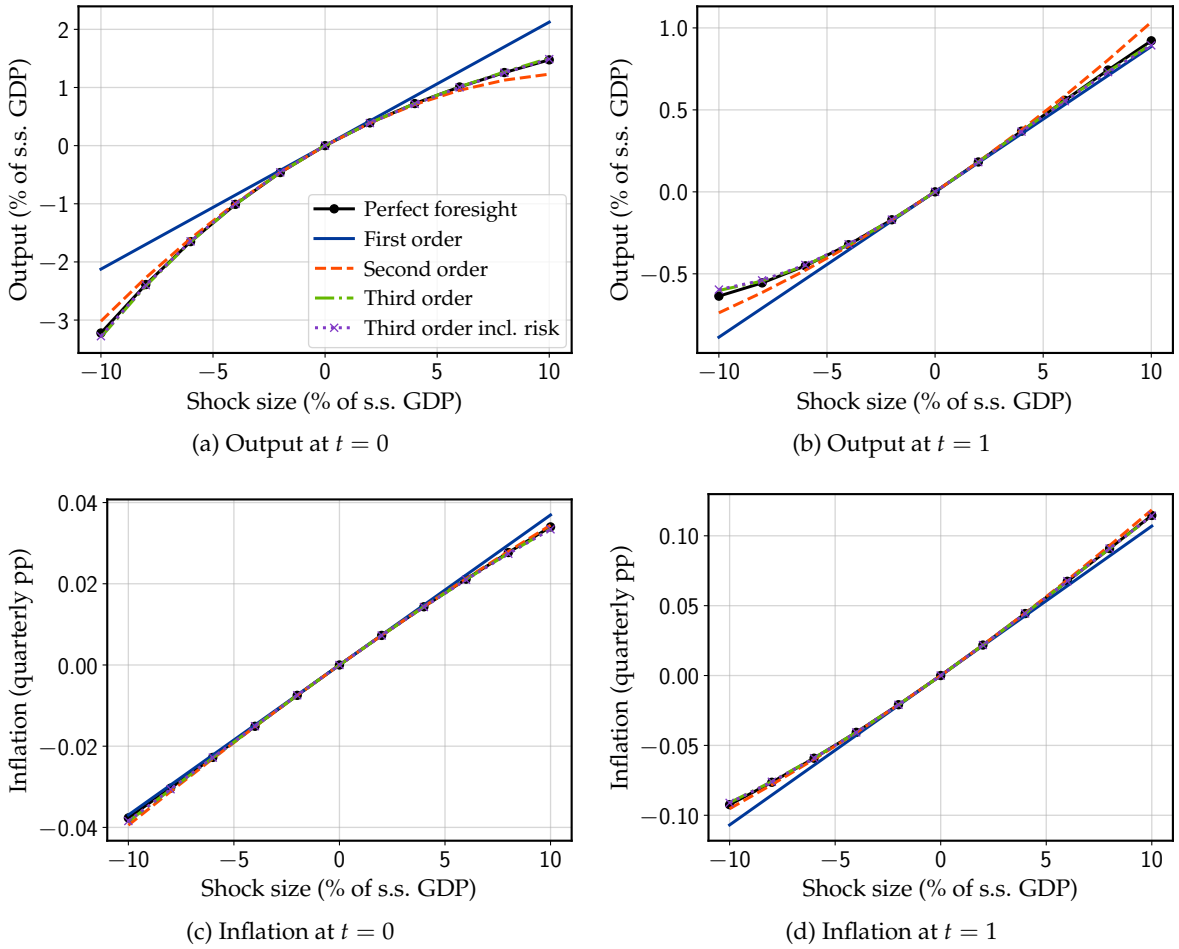
(c) Real interest rate



(d) Inflation

Figure 9: Risky steady state in the quantitative HANK model

Note: Shock standard deviations σ_G and σ_p are scaled proportional to their estimated values. The horizontal axis shows GDP standard deviation computed with a fifth order approximation.



(a) Output at $t = 0$

(b) Output at $t = 1$

(c) Inflation at $t = 0$

(d) Inflation at $t = 1$

Figure 10: Size dependence in the quantitative HANK model

Note: All responses assume $\sigma_G = \sigma_p = 0$, except for the purple line, which computes responses under estimated shock variances. Responses are computed around the deterministic steady state.

major role.

Estimated model and nonlinear multipliers. To conclude, we study history dependence of the transfer multiplier, dY_t/dT_t. Using our estimated shocks, we construct the full third-order Volterra expansion of the HANK model. We do so allowing for the two estimated government spending and cost-push shocks as well as tax cut shocks with standard deviation σ and persistence ρ as before. We then simulate data from our model for 2,000 quarters. After each history of shocks, we compute the transfer multiplier from a small tax cut shock.[41]

Figure 11(b) plots the distribution of impact multipliers of the transfer shock. There is a wide range of multipliers, ranging roughly from 0.15 to 0.30. Panel a shows that the multipliers are are countercyclical over the business cycle: When output is high, multipliers tend to be low, and

[41] Strictly speaking, for this exercise a full third-order Volterra expansion with three shocks is not necessary. Only terms in which the tax cut shock ϵ_t does not enter (but the other two shocks do, to up to third order) or enters linearly (with the other two shocks entering to up to second order) need to be considered.

vice versa when output is low. This stems entirely from government spending shocks: Panel c considers the same simulation but assumes no cost-push shocks. Then, transfer multipliers are strongly countercyclical, as positive government spending shocks as well as tax cut shocks act via after-tax household income, and when income is already high due to past positive government spending shocks, an additional tax cut is significantly less impactful. When government spending shocks instead of cost-push shocks are switched off, the multiplier becomes mildly procyclical.

These findings show that the notion of "state dependence" can be subtle to separate empirically: here, the level of output does not summarize the state, and the nature of shocks that pushed it there matters. This may help explain why the empirical literature on the state dependence of fiscal multipliers has generated mixed results, with some papers finding that fiscal multipliers are much larger in recessions, while others find a much weaker relationship (see e.g. Auerbach and Gorodnichenko 2012 vs. Ramey and Zubairy 2018).[42]

5.2 Nonlinear Phillips curve in a menu cost model

Our second application is a menu cost model with idiosyncratic shocks in the spirit of Golosov and Lucas (2007), and Alvarez, Le Bihan and Lippi (2016). The model features a random menu cost, trend inflation, and strategic complementarity, as in Auclert, Rigato, Rognlie and Straub (2024c) and Alvarez, Lippi and Souganidis (2023). Firms have a quadratic objective and try to keep their price gap x as close as possible to the aggregate log nominal marginal cost MC_t. They can reset their price either by paying a fixed menu cost ξ or by waiting for a free opportunity to adjust their price, which arrives with Calvo probability λ. If the price is not reset, the price gap evolves as an idiosyncratic random walk, $x' = x + \epsilon$, with ϵ drawn from a distribution with smooth density $f(\epsilon)$. This ensures that the only idiosyncratic state of the model is the price gap x.

The Bellman equation of a price setter with price gap x at date t is given by:

$$V_t(x) = \frac{1}{2}(x - MC_t)^2 + \beta \mathbb{E}_t \left[(1-\lambda) \mathbb{E}_\epsilon [\min\{V_{t+1}(x+\epsilon), \xi + \min_{x^*} V_{t+1}(x^*)\}] + \lambda \min_{x^*} V_{t+1}(x^*) \right]$$

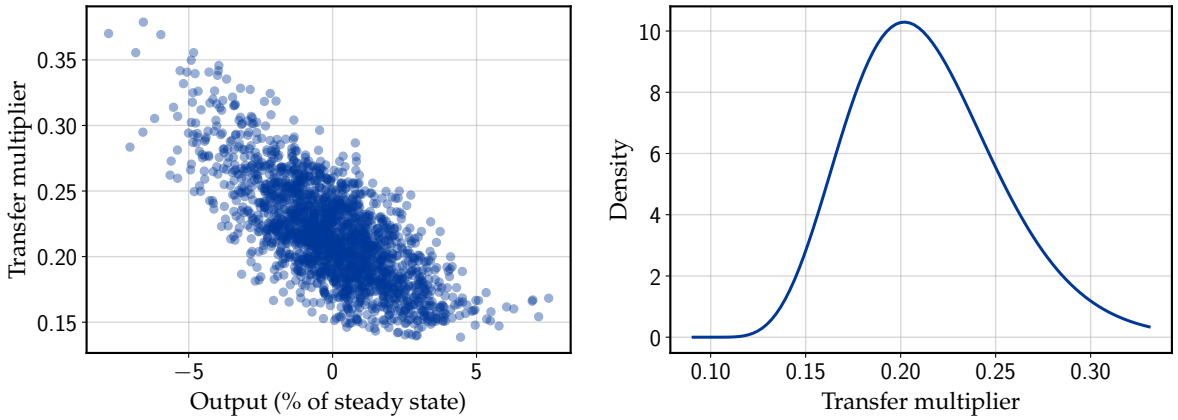
As is common in the literature (e.g. Alvarez et al. 2016), we assume linear aggregation, so that the log price index is

$$P_t = m^* + \int x dD_t(x) \tag{48}$$

where m^* is the desired (flexible-price) markup over marginal cost, and D_t is the distribution of gaps after adjustment. Finally, we assume that strategic complementarity in price setting arises from roundabout production, with firms producing from an aggregate of labor and an input bundle made up of production of other firms; this implies that log nominal marginal cost is

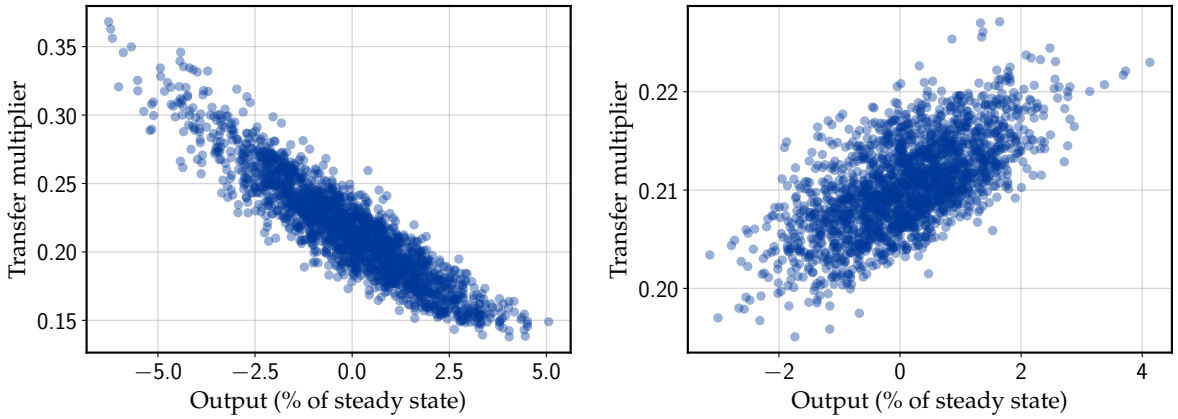
$$MC_t = \chi W_t + (1-\chi) P_t \tag{49}$$

[42] See Albertini, Auray, Bouakez and Eyquem (2021), Hagedorn, Manovskii and Mitman (2019) and Auclert et al. (2024a) for other papers studying the state dependence of fiscal multipliers in a HANK environment.



(a) Simulated values with both shocks

(b) Ergodic distribution with both shocks



(c) Simulated values with ϵ^G shock only

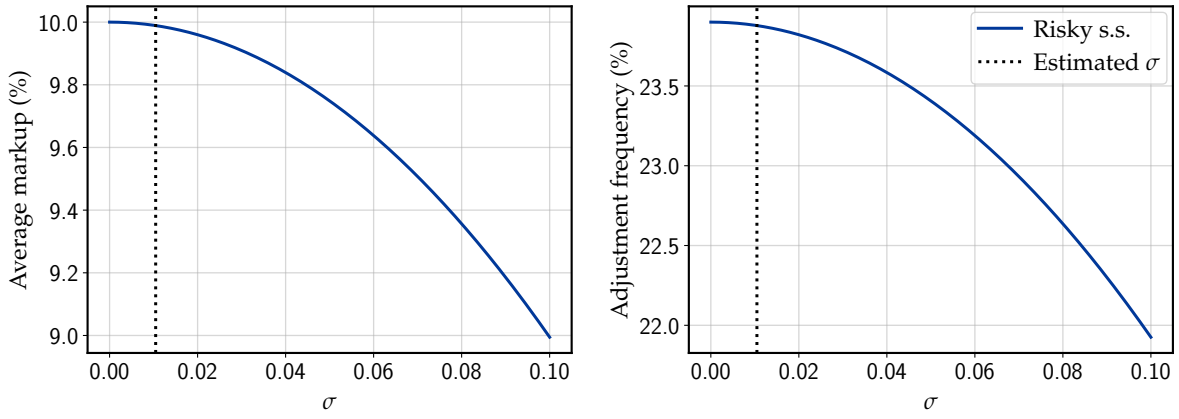
(d) Simulated values with ϵ^p shock only

Figure 11: Time-varying transfer multipliers in the quantitative HANK model

Note: On the horizontal axis, output is reported in percent deviations from the deterministic steady state. Panel (a) shows a simulation of 2000 periods using estimated parameter values. Panel (b) shows the ergodic distribution of the transfer multiplier $\partial Y_t / \partial \epsilon_t^B$. Panels (c) and (d) set realized markup and government spending innovations, respectively, to zero, while keeping anticipated shock variances at their estimated values.

where χ is the fraction of value added in production. The log nominal wage trends at rate μ, $W_t = \mu t + \tilde{W}_t$, where $\tilde{W}_t = \rho \tilde{W}_{t-1} + \sigma \bar{\epsilon}_t$ follows an AR(1) process and is the source of aggregate risk in the model.

We calibrate the model at a quarterly frequency, with $\beta = 0.99$. We assume that the idiosyncratic shock distribution f is normal with mean 0 and variance σ_ϵ^2, and jointly choose ξ, λ, and σ_ϵ to match a price adjustment frequency in the deterministic steady state of freq $= 0.239$, a median absolute size of price changes of 0.085, and a share of free adjustments in total adjustments of 0.75. These numbers follow Nakamura and Steinsson (2010), converted to quarterly data (see Auclert et al. 2024c). We further set the share of value added in production to $\chi = 0.7$ and trend inflation to $\mu = 0.8512\%$ quarterly (about 3.4% p.a.), which is average PCE inflation in the U.S. from 1966Q1 to 2025Q4. Finally, we set m^* so that the average log price-wage markup is equal to 10% in the deterministic steady state.



(a) Average price-wage markup

(b) Price adjustment frequency

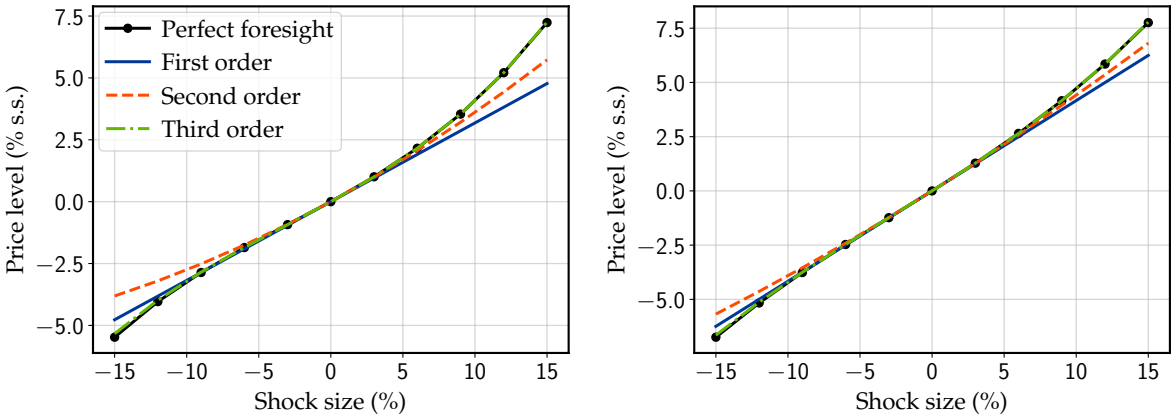
Figure 12: Risky steady state

Note: Risky steady state as a function of the standard deviation of the nominal wage shock σ.

Like in the IKC model or the quantitative HANK model, this model features a general equilibrium fixed point problem: the price level P_t depends on nominal marginal cost through firms' price-setting decisions, while marginal cost itself depends on the price level through strategic complementarity (see e.g. Alvarez et al. 2023). As further described in appendix F.2, we solve this model in the steady state by detrending the model by wage inflation μ, and ensuring that the model is smooth. We formulate the model as in (1) by keeping track of $W_t(x) = \mathbb{E}_\epsilon[\min\{V_{t+1}(x+\epsilon), \xi + \min_{x^*} V_{t+1}(x^*)\}]$ as an element of the X_t vector. We then solve for the higher-order solutions of the model using a reduced system with P_t as the unknown variable and applying the one-shot principle.

Estimation of σ and ρ. We estimate the standard deviation and persistence of wage shocks using GMM, similar to our estimation in the previous subsection. We match the fifth-order correct variance and first autocovariance of the stationary log price level in the model to the corresponding moments of the HP-filtered log PCE price index (1966Q1–2025Q4), using an HP smoothing parameter of 1600. We find $\sigma = 1.047$ (s.e. 0.344, both in pp.) and $\rho = 0.74991$ (s.e. 0.09553).

Risky steady state. Figure 12 plots the average log price-wage markup (panel a) and the adjustment frequency (panel b) as function of σ in the risky steady state of the estimated model. Recall that we have calibrated the price-wage markup to 10% in the deterministic steady state. The figure shows that additional cost uncertainty σ widens the inaction region, reducing the frequency of adjustment (panel b), but it also induces firms to reset prices more conservatively and place comparatively less weight on trend inflation. This *precautionary pricing* mechanism therefore brings down markups in the risky steady state (panel a).



(a) Response at $t = 0$

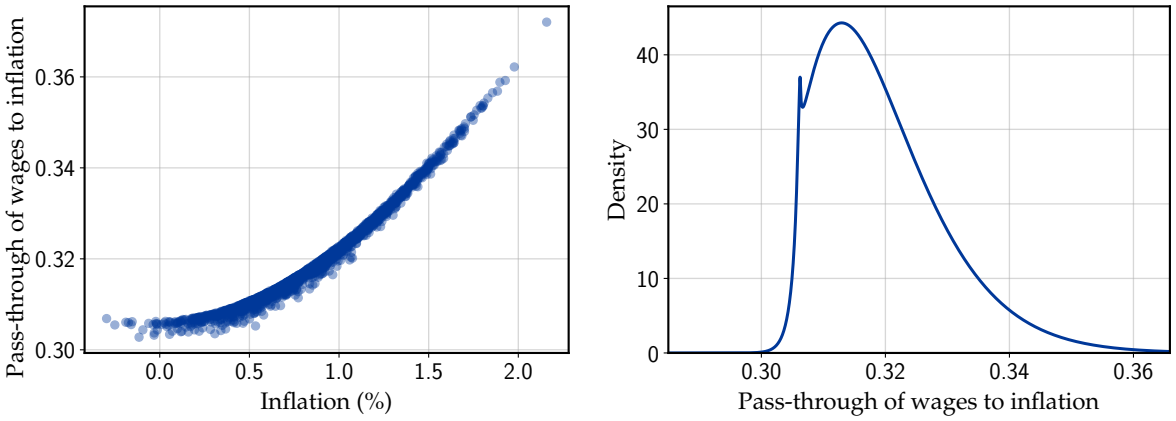
(b) Response at $t = 1$

Figure 13: Response of the price level to nominal wage shocks

Note: Log price level responses to a nominal wage shock starting at the deterministic steady state under perfect foresight. The first-, second- and third-order curves show the corresponding approximations to the perfect-foresight response.

Impulse responses. Next, we turn to the impulse response of the log price level to wage shocks of different sizes. Menu cost models are well known to have nonlinear effects of costs on prices, as large shocks of either sign trigger more widespread price adjustment. Figure 13 cleanly shows this pattern of sign and size effects in the model here: large positive as well as large negative shocks show a larger pass through of wages to prices, both at $t = 0$ (panel a) and at $t = 1$ (panel b).

History-dependent cost pass-through. We use our methods to construct the third-order Volterra expansion and simulate 2,000 quarters of data. After each history, we use the Volterra expansion to extract the contemporaneous pass-through of the wage shock ϵ_t to inflation $d\pi_t / d\epsilon_t$, which one can regard as a measure of the local slope of the Phillips curve. Figure 14 shows that there is nontrivial variation over time in the pass-through of wages to inflation, with a coefficient fluctuating between about 0.30 to about 0.35 (panel b), and faster pass-through typically occurring in high-inflation environments (panel a). This effect is particularly pronounced in the third-order solution, which produces right skewness in panel b and the convex pattern visible in panel a. Our finding that higher inflation coincides with faster pass-through of costs to inflation is emphasized by a recent literature on inflation in the 2020s (e.g. Cavallo, Lippi and Miyahara 2024, Blanco et al. 2024, Blanco, Boar, Jones and Midrigan 2026, Karádi, Nakov, Nuño, Pastén and Thaler 2026, Gagliardone, Gertler, Lenzu and Tielens 2026). While this literature often looks for approximations to menu cost models, here we are able to solve the complete menu cost model to third order, bringing about not just nonlinearities but also state dependence and precautionary pricing.



(a) Simulated values

(b) Ergodic distribution

Figure 14: Time-varying pass-through of nominal wages to inflation

Note: Panel (a) shows 2000 periods simulated from the menu cost model with estimated shock process parameters. Panel (b) shows the ergodic distribution of $\partial \pi_t / \partial \epsilon_t$.

6 Conclusion

This paper shows how to obtain higher-order perturbation solutions to heterogeneous-agent models using information derived from well-known routines to solve the model under perfect foresight, which we augment to solve rapidly for higher-order general equilibrium. This allows us to solve a HANK model and a menu cost model to third order at a much lower computational burden than a state-space solution would require. Our checks for accuracy against both nonlinear perfect foresight and a global solution perform well. An open question, however, is the extent to which our third order perturbation solutions differ from those obtained by global solution methods. This is an active area of research.

References

Adjemian, Stéphane and Michel Juillard, "Stochastic Extended Path," *Journal of Economic Dynamics and Control*, 2026, *182*, 105227.

Albertini, Julien, Stéphane Auray, Hafedh Bouakez, and Aurélien Eyquem, "Taking Off into the Wind: Unemployment Risk and State-Dependent Government Spending Multipliers," *Journal of Monetary Economics*, 2021, *117*, 990–1007.

Alvarez, Fernando, Francesco Lippi, and Panagiotis Souganidis, "Price Setting with Strategic Complementarities as a Mean Field Game," *Econometrica*, 2023, *91* (6), 2005–2039.

_ , Hervé Le Bihan, and Francesco Lippi, "The Real Effects of Monetary Shocks in Sticky Price Models: A Sufficient Statistic Approach," *American Economic Review*, 2016, *106* (10), 2817–2851.

Andreasen, Martin M., Giovanni Caggiano, Efrem Castelnuovo, and Giovanni Pellegrino,

"Does Risk Matter More in Recessions Than in Expansions? Implications for Monetary Policy," *Journal of Monetary Economics*, 2024, *143*, 103533.

_ , **Jesús Fernández-Villaverde, and Juan F. Rubio-Ramírez**, "The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications," *Review of Economic Studies*, 2018, *85* (1), 1–49.

Auclert, Adrien and Matthew Rognlie, "Inequality and Aggregate Demand," 2018. Manuscript.

_ , **Bence Bardóczy, Matthew Rognlie, and Ludwig Straub**, "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," *Econometrica*, 2021, *89* (5), 2375–2408.

_ , **Evan Majic, Matthew Rognlie, and Ludwig Straub**, "Determinacy and Large-Scale Solutions in the Sequence Space," *Journal of Political Economy Macroeconomics*, 2025. Forthcoming.

_ , **Matthew Rognlie, and Ludwig Straub**, "The Trickling Up of Excess Savings," *AEA Papers and Proceedings*, 2023, *113*, 70–75.

_ , _ , **and** _ , "The Intertemporal Keynesian Cross," *Journal of Political Economy*, 2024, *132* (12), 4068–4121.

_ , _ , _ , **and Tomáš Ťapák**, "When Do Endogenous Portfolios Matter for HANK?," 2024. Manuscript.

_ , **Rodolfo Rigato, Matthew Rognlie, and Ludwig Straub**, "New Pricing Models, Same Old Phillips Curves?," *Quarterly Journal of Economics*, 2024, *139* (1), 121–186.

Auerbach, Alan J. and Yuriy Gorodnichenko, "Measuring the Output Responses to Fiscal Policy," *American Economic Journal: Economic Policy*, 2012, *4* (2), 1–27.

Azinovic-Yang, Marlon and Jan Žemlička, "Deep Learning in the Sequence Space," arXiv Preprint arXiv:2509.13623 2025.

Bayer, Christian, Luigi-Maria Briglia, Ralph Luetticke, Maximilian Weiss, and Yannik Winkelmann, "Local Methods for Large Transfers," 2026. Manuscript.

_ , **Ralph Luetticke, Maximilian Weiss, and Yannik Winkelmann**, "An Endogenous Gridpoint Method for Distributional Dynamics," *Journal of Monetary Economics*, 2026, *158*, 103895.

Bhandari, Anmol, Thomas Bourany, David Evans, and Mikhail Golosov, "A Perturbational Approach for Approximating Heterogeneous-Agent Models," 2023. Manuscript.

Bianchi, Javier and Greg Kaplan, "How Small Is Small? Non-Linearities in Heterogeneous Agent Models," 2026. Manuscript.

Bilal, Adrien, "Solving Heterogeneous Agent Models with the Master Equation," 2023. Manuscript.

_ **and Shlok Goyal**, "Some Pleasant Sequence-Space Arithmetic in Continuous Time," *Journal of Political Economy Macroeconomics*, 2025. Forthcoming.

Blanchard, Olivier J. and Charles M. Kahn, "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 1980, *48* (5), 1305–1311.

Blanco, Andrés, Corina Boar, Callum J. Jones, and Virgiliu Midrigan, "Non-Linear Inflation Dynamics in Menu Cost Economies," 2024. Manuscript.

_ , _ , _ , and _ , "The Inflation Accelerator," *American Economic Review*, 2026. Forthcoming.

Boppart, Timo, Per Krusell, and Kurt Mitman, "Exploiting MIT Shocks in Heterogeneous-Agent Economies: The Impulse Response as a Numerical Derivative," *Journal of Economic Dynamics and Control*, 2018, *89*, 68–92.

Carroll, Christopher D., "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, 2006, *91* (3), 312–320.

Cavallo, Alberto, Francesco Lippi, and Ken Miyahara, "Large Shocks Travel Fast," *American Economic Review: Insights*, 2024, *6* (4), 558–574.

Christiano, Lawrence J., "Linear-Quadratic Approximation and Value-Function Iteration: A Comparison," *Journal of Business & Economic Statistics*, 1990, *8* (1), 99–113.

Coeurdacier, Nicolas, Hélène Rey, and Pablo Winant, "The Risky Steady State," *American Economic Review*, 2011, *101* (3), 398–401.

de Groot, Oliver, "Computing the Risky Steady State of DSGE Models," *Economics Letters*, 2013, *120* (3), 566–569.

Druedahl, Jeppe, Raphaël Huleux, and Jacob Røpke, "Fiscal Multipliers in a Globally Solved HANK Model with Aggregate Risk," 2026. Manuscript.

Farhi, Emmanuel, Alan Olivi, and Iván Werning, "Price Theory for Incomplete Markets," 2022. Manuscript.

Fernández-Villaverde, Jesús, Juan F. Rubio-Ramírez, and Frank Schorfheide, "Solution and Estimation Methods for DSGE Models," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2, Elsevier, 2016, chapter 9, pp. 527–724.

Gagliardone, Luca, Mark Gertler, Simone Lenzu, and Joris Tielens, "Micro and Macro Cost-Price Dynamics in Normal Times and during Inflation Surges," 2026. Manuscript.

Ganong, Peter, Pascal Noel, Christina Patterson, Joseph Vavra, and Alexander Weinberg, "Earnings Instability," 2025. Manuscript.

Golosov, Mikhail and Robert E. Lucas Jr., "Menu Costs and Phillips Curves," *Journal of Political Economy*, 2007, *115* (2), 171–199.

Griewank, Andreas and Andrea Walther, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, SIAM, 2008.

Hagedorn, Marcus, Iourii Manovskii, and Kurt Mitman, "The Fiscal Multiplier," 2019.

Manuscript.

Hamilton, James D., *Time Series Analysis*, Princeton University Press, 1994.

Han, Jiequn, Yucheng Yang, and Weinan E, "DeepHAM: A Global Solution Method for Heterogeneous Agent Models with Aggregate Shocks," arXiv Preprint arXiv:2112.14377 2021.

Hansen, Lars P., "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1982, *50* (4), 1029–1054.

Imhof, Jean-Pierre, "Computing the Distribution of Quadratic Forms in Normal Variables," *Biometrika*, 1961, *48* (3/4), 419–426.

Iskhakov, Fedor, Thomas H. Jørgensen, John Rust, and Bertel Schjerning, "The Endogenous Grid Method for Discrete-Continuous Dynamic Choice Models with (or without) Taste Shocks," *Quantitative Economics*, 2017, *8* (2), 317–365.

Kameník, Ondra, "Solution of Specialized Sylvester Equation," 2003. Manuscript.

Karádi, Péter, Anton Nakov, Galo Nuño, Ernesto Pastén, and Dominik Thaler, "Strike While the Iron Is Hot: Optimal Monetary Policy under State-Dependent Pricing," *American Economic Review*, 2026. Forthcoming.

Kase, Hanno, Leonardo Melosi, and Matthias Rottner, "Estimating Nonlinear Heterogeneous Agents Models with Neural Networks," 2022. Manuscript.

Kim, Jinill, Sunghyun Kim, Ernst Schaumburg, and Christopher A. Sims, "Calculating and Using Second-Order Accurate Solutions of Discrete Time Dynamic Equilibrium Models," *Journal of Economic Dynamics and Control*, 2008, *32* (11), 3397–3414.

King, Robert G. and Sergio T. Rebelo, "Resuscitating Real Business Cycles," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1, Part B, Elsevier, 1999, chapter 14, pp. 927–1007.

Klein, Paul, "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control*, 2000, *24* (10), 1405–1423.

Kopecky, Karen A. and Richard M. H. Suen, "Finite State Markov-Chain Approximations to Highly Persistent Processes," *Review of Economic Dynamics*, 2010, *13* (3), 701–714.

Lan, Hong and Alexander Meyer-Gohde, "Pruning in Perturbation DSGE Models: Guidance from Nonlinear Moving Average Approximations," 2013. Manuscript.

_ **and** _ , "Solving DSGE Models with a Nonlinear Moving Average," *Journal of Economic Dynamics and Control*, 2013, *37* (12), 2643–2667.

Le Grand, François and Xavier Ragot, "Refining the Truncation Method to Solve Heterogeneous-Agent Models," *Annals of Economics and Statistics*, 2022, (146), 65–92.

Lee, Hanbaek, "Global Nonlinear Solutions in Sequence Space and the Generalized Transition

Function," 2025. Manuscript.

Lin, Alessandro and Marcel Peruffo, "Aggregate Uncertainty, HANK, and the ZLB," 2024. Manuscript.

Lombardo, Giovanni and Harald Uhlig, "A Theory of Pruning," *International Economic Review*, 2018, *59* (4), 1825–1836.

Nakamura, Emi and Jón Steinsson, "Monetary Non-Neutrality in a Multisector Menu Cost Model," *Quarterly Journal of Economics*, 2010, *125* (3), 961–1013.

Newey, Whitney K. and Kenneth D. West, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 1987, *55* (3), 703–708.

Ramey, Valerie A. and Sarah Zubairy, "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data," *Journal of Political Economy*, 2018, *126* (2), 850–901.

Reiter, Michael, "State Reduction and Second-Order Perturbations of Heterogeneous Agent Models," 2023. Manuscript.

Rigato, Rodolfo, "A Least-Squares Filter for Sequence-Space Models," *Economics Letters*, 2026, *265*, 112988.

Rion, Normann, "Rich Heterogeneity in Dynare 7: A Practical Description," 2026. Manuscript.

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

_ **and** _ , "Heterogeneous Downward Nominal Wage Rigidity: Foundations of a Nonlinear Phillips Curve," 2022. Manuscript.

Taylor, John B. and Harald Uhlig, "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods," *Journal of Business & Economic Statistics*, 1990, *8* (1), 1–17.

Yang, Yucheng, Chiyuan Wang, Andreas Schaab, and Benjamin Moll, "Structural Reinforcement Learning for Heterogeneous Agent Macroeconomics," arXiv Preprint arXiv:2512.18892 2025.

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

Appendix for "Higher-Order Perturbation in the Sequence Space: the Certainty Correspondence"

A Appendix to section 2

A.1 Rank condition for assumption 1

Together, the root and rank conditions in assumption 1 are the standard conditions ensuring that, for every initial state x_{-1}, there is a unique solution to the linearized perfect-foresight system $F_1 x_{t-1} + F_2 x_t + F_3 x_{t+1} = 0$ for $t \geq 0$ that converges back to the steady state.

Concretely, stack the system as

$$N \begin{pmatrix} x_{t+1} \\ x_t \end{pmatrix} = M \begin{pmatrix} x_t \\ x_{t-1} \end{pmatrix},$$

where

$$M \equiv \begin{pmatrix} -F_2 & -F_1 \\ I & 0 \end{pmatrix}, \qquad N \equiv \begin{pmatrix} F_3 & 0 \\ 0 & I \end{pmatrix}.$$

These matrices satisfy $\det(M - \lambda N) = \det(F_1 + \lambda F_2 + \lambda^2 F_3)$, so the roots λ in assumption 1 are the generalized eigenvalues of (M, N), meaning that $Mz = \lambda Nz$ for some nonzero z. For a real root, the stacked path $\lambda^t z$ then solves the system and converges to zero when $|\lambda| < 1$. More generally, the "stable subspace" associated with these roots gives exactly the space of possible pairs (x_t, x_{t-1}) that can be continued with paths that converge to zero. The root condition in assumption 1 says that this subspace is n-dimensional.

What remains is to ensure that, for every initial x_{-1}, the stable subspace contains exactly one pair (x_0, x_{-1}). To state this formally, let $Z_s \in \mathbb{R}^{2n \times n}$ be any real basis matrix for the stable subspace, and let $Z_{-1} \in \mathbb{R}^{n \times n}$ be its last n rows.[A-1] Every pair in the stable subspace can be written as $Z_s a$ for some $a \in \mathbb{R}^n$, and its last n entries are $Z_{-1} a$. Thus each x_{-1} corresponds to exactly one such pair if and only if $x_{-1} = Z_{-1} a$ has exactly one solution a; otherwise, some x_{-1} will have no stable solution, and others will have more than one.

The corresponding rank condition we require in assumption 1 is therefore

$$\operatorname{rank}(Z_{-1}) = n, \qquad \text{or equivalently} \qquad \det Z_{-1} \neq 0. \tag{A.1}$$

Condition (A.1) is exactly the invertibility condition in Assumption 4.5 of Klein (2000), although the notation there is different. See also section 5.3.1 in that paper for discussion of why root count alone is not sufficient.

A similar rank or invertibility condition appears in other canonical linear rational expectations

[A-1] This can be obtained, for instance, with an ordered real QZ decomposition of (M, N), as we do in the next section. (A.1) is unaffected by the choice of basis matrix.

solution methods. See, for instance, the rank assumption immediately preceding proposition 1 in Blanchard and Kahn (1980). It is worth noting that rank conditions like (A.1) hold generically, in the sense that failure requires the stable subspace to line up exactly so that x_0 can change without any change in x_{-1}. This is an alignment that small perturbations would ordinarily break, unless it is enforced by the structure of a model.

A.2 Sequence-space form of assumption 1

Assuming F is continuously differentiable, we show that the root and rank conditions in assumption 1 are equivalent to the invertibility of a block sequence-space Jacobian. Since this Jacobian is block Toeplitz, its invertibility can be tested directly using the methods in Auclert, Majic, Rognlie and Straub (2025), allowing us to verify assumption 1 while remaining in the sequence space.[A-2]

Concretely, let ℓ^2 be the space of square-summable paths $x = (x_0, x_1, \ldots)$ in $\mathbb{R}^n$. Define the bounded linear operator $J : \ell^2 \to \ell^2$ as

$$(Jx)_0 = F_2 x_0 + F_3 x_1, \qquad (Jx)_t = F_1 x_{t-1} + F_2 x_t + F_3 x_{t+1} \quad (t \geq 1). \tag{A.2}$$

Here Jx is the path of linearized equilibrium residuals produced by x, with $x_{-1} = 0$. Membership in ℓ^2 requires $x_t \to 0$, and invertibility of J means that every residual path $r \in \ell^2$ corresponds to exactly one square-summable path. The proposition identifies this property with the root and rank conditions.

Proposition 1. The root and rank conditions in assumption 1 hold if and only if J is invertible on ℓ^2.

Proof. Suppose first that the root and rank conditions hold. To construct J^{-1}, we first derive from the stable subspace a matrix P satisfying $F_1 + F_2 P + F_3 P^2 = 0$, and then use it to solve $Jx = r$ recursively.

Partition the basis matrix for the stable subspace from appendix A.1 as $Z_s = \binom{Z_0}{Z_{-1}}$. Since Z_{-1} is invertible by the rank condition (A.1), for each $a \in \mathbb{R}^n$ the stable subspace contains exactly one vector with lower block a, namely (Pa, a), where $P \equiv Z_0 Z_{-1}^{-1}$. This vector gives (x_0, x_{-1}) for a solution satisfying $F_1 x_{t-1} + F_2 x_t + F_3 x_{t+1} = 0$ and $x_t \to 0$. The next vector (x_1, x_0) also lies in the stable subspace. Since $x_0 = Pa$, it must be $(P^2 a, Pa)$. Therefore $F_1 a + F_2 Pa + F_3 P^2 a = 0$ for every a, which gives

$$F_1 + F_2 P + F_3 P^2 = 0.$$

Moreover, the eigenvalues of P are precisely the roots inside the unit circle, so $\rho(P) < 1$.[A-3]

[A-2] See proposition 11 in that paper for the main winding number test, which—as the paper discusses—corresponds to the root condition. See also the discussion in section 3.5 for how to use nonexistence to identify the "non-generic" case where the winding number test is not enough, which corresponds to failure of the rank condition.

[A-3] For a concrete construction, take the n stable columns Z_s from an ordered real QZ decomposition of (M, N). The QZ identities imply $MZ_s = NZ_s A_s$, where the eigenvalues of A_s are the stable roots. The lower block gives $Z_0 = Z_{-1} A_s$, so $P = Z_{-1} A_s Z_{-1}^{-1}$. The upper block then gives the matrix quadratic above.

We now return to $Jx = r$. For a general path x, define $u_t \equiv x_t - Px_{t-1}$. Substituting $x_t = Px_{t-1} + u_t$ and $x_{t+1} = Px_t + u_{t+1}$ into $(Jx)_t = r_t$ gives, after cancellation,

$$r_t = Ku_t + F_3 u_{t+1}, \qquad K \equiv F_2 + F_3 P. \tag{A.3}$$

To solve (A.3) backward—that is, to determine u_t from $r_t, r_{t+1}, \ldots$—we need K to be invertible and the powers of $R \equiv -K^{-1}F_3$ to decay. We now establish both facts from the root condition.

Substituting $F_1 = -F_2 P - F_3 P^2$ gives

$$F_1 + \lambda F_2 + \lambda^2 F_3 = (K + \lambda F_3)(\lambda I - P). \tag{A.4}$$

Because $\det(\lambda I - P) = 0$ at the n eigenvalues of P, the second factor accounts for all n roots inside the unit circle. Thus K must be invertible: otherwise the first factor would also be singular at $\lambda = 0$, adding another root inside the unit circle. The remaining finite roots therefore lie outside the unit circle. We may now define $R \equiv -K^{-1}F_3$. Since $K + \lambda F_3 = K(I - \lambda R)$, the nonzero eigenvalues of R are the reciprocals of these remaining roots; a root at infinity corresponds to a zero eigenvalue of R. Thus $\rho(R) < 1$, as needed.

Equation (A.3) can now be written as

$$u_t = K^{-1} r_t + R u_{t+1}.$$

For any solution in ℓ^2, $u \in \ell^2$. Iterating this equation toward future dates and using square summability to eliminate the terminal term $R^{L+1}u_{t+L+1}$ gives the first expression below. Iterating $x_t = Px_{t-1} + u_t$ forward from $x_{-1} = 0$ gives the second:

$$u_t = \sum_{j=0}^{\infty} R^j K^{-1} r_{t+j}, \qquad x_t = \sum_{k=0}^{t} P^k u_{t-k}. \tag{A.5}$$

The first expression discounts future residuals back to date t; the second propagates their effects forward through the stable law of motion P.[A-4] Since the powers of P and R decay geometrically, these formulas give a unique $x \in \ell^2$ for every $r \in \ell^2$, with $\|x\|_2 \le C\|r\|_2$. Hence J is invertible.

Conversely, suppose that J is invertible. We first rule out roots on the unit circle; this also shows that the determinant polynomial is not identically zero. We then identify the pairs (x_0, x_{-1}) that admit an ℓ^2 continuation and use invertibility of J to establish the root count and rank condition.

Suppose, to the contrary, that there is a unit root: $(F_1 + \lambda F_2 + \lambda^2 F_3)v = 0$ for some $|\lambda| = 1$ and $v \neq 0$.[A-5] We use this root to construct paths with fixed ℓ^2 norm and arbitrarily small $\|Jx\|_2$, contradicting boundedness of J^{-1}. For arbitrary $L \ge 2$, set $x_t^{(L)} = L^{-1/2}\lambda^{t+1}v$ for $0 \le t < L$ and zero thereafter; we note that $\|x^{(L)}\|_2 = |v|$. The root equation then implies $(Jx^{(L)})_t = 0$ for

[A-4] A closely related construction of the sequence-space Jacobian that solves a linearized rational expectations system appears in appendices B.7–B.8 of Auclert et al. (2025).

[A-5] If λ or v is complex, apply the argument to complex-valued paths. The real operator J^{-1} extends to them by acting separately on their real and imaginary parts.

$1 \le t \le L-2$; only $(Jx^{(L)})_0$, $(Jx^{(L)})_{L-1}$, and $(Jx^{(L)})_L$ can be nonzero. At these three dates, the root equation gives

$$(Jx^{(L)})_0 = -L^{-1/2}F_1 v, \qquad (Jx^{(L)})_{L-1} = -L^{-1/2}\lambda^{L+1}F_3 v, \qquad (Jx^{(L)})_L = L^{-1/2}\lambda^L F_1 v.$$

Thus $\|Jx^{(L)}\|_2 = L^{-1/2}(2|F_1 v|^2 + |F_3 v|^2)^{1/2} \to 0$, contradicting boundedness of J^{-1}. Thus there are no roots on the unit circle, and $\det(F_1 + \lambda F_2 + \lambda^2 F_3)$ cannot be identically zero.

It remains to count the roots inside the unit circle and establish the rank condition. With roots on the unit circle ruled out, the stable subspace described in appendix A.1 consists exactly of the pairs (x_0, x_{-1}) for which there is a path $x \in \ell^2$ satisfying $F_1 x_{t-1} + F_2 x_t + F_3 x_{t+1} = 0$ for all $t \ge 0$.[A-6] The dimension of this subspace is the number of roots strictly inside the unit circle, counted with algebraic multiplicity.

For any $a \in \mathbb{R}^n$, define $\widetilde{r} \in \ell^2$ by $\widetilde{r}_0 = -F_1 a$ and $\widetilde{r}_t = 0$ for $t \ge 1$. By the definition of J, $Jx = \widetilde{r}$ holds exactly when $F_1 x_{t-1} + F_2 x_t + F_3 x_{t+1} = 0$ for all $t \ge 0$ after setting $x_{-1} = a$. Since J is invertible, there is exactly one such $x \in \ell^2$ for each a. Thus $(x_0, x_{-1}) \mapsto x_{-1}$ is a bijection from the stable subspace onto $\mathbb{R}^n$. The stable subspace therefore has dimension n, so exactly n roots lie strictly inside the unit circle. For a basis Z_s as in appendix A.1, this projection is represented by Z_{-1}, so its bijectivity is exactly the rank condition (A.1). □

The proof also gives the following representation, which we use in appendix A.3.

Lemma 2 (Stable linear law of motion)**.** Under assumption 1, every path satisfying

$$F_1 x_{t-1} + F_2 x_t + F_3 x_{t+1} = 0$$

and $x_t \to 0$ follows a linear law of motion $x_t = P x_{t-1}$, where

$$F_1 + F_2 P + F_3 P^2 = 0, \qquad \rho(P) < 1. \tag{A.6}$$

Moreover, with

$$K \equiv F_2 + F_3 P, \qquad R \equiv -K^{-1} F_3, \tag{A.7}$$

K is invertible and $\rho(R) < 1$.

Proof. These claims were established in the forward direction of proposition 1. □

[A-6] To see this, let k be the number of roots inside the unit circle and apply an ordered real QZ decomposition to the stacked system $N\binom{x_{t+1}}{x_t} = M\binom{x_t}{x_{t-1}}$ from appendix A.1:

$$Q^\top N Z = \begin{pmatrix} S_s & S_{su} \\ 0 & S_u \end{pmatrix}, \qquad Q^\top M Z = \begin{pmatrix} T_s & T_{su} \\ 0 & T_u \end{pmatrix},$$

placing these k roots first. Write $Z = (Z_s \; Z_u)$, where Z_s and Z_u have k and $2n-k$ columns, and $\binom{x_t}{x_{t-1}} = Z\binom{y_t^s}{y_t^u}$. All roots in the second block lie outside the unit circle or at infinity, so T_u is invertible and $\rho(T_u^{-1}S_u) < 1$. The equations for y_t^u give $y_t^u = (T_u^{-1}S_u)^L y_{t+L}^u$. For $x \in \ell^2$, $y_{t+L}^u \to 0$, so $y_t^u = 0$. Conversely, S_s is invertible and $\rho(S_s^{-1}T_s) < 1$. Setting $y_t^u = 0$ and iterating $y_{t+1}^s = S_s^{-1}T_s y_t^s$ gives an ℓ^2 continuation from every initial pair in $\mathrm{span}(Z_s)$. Thus exactly the initial pairs in $\mathrm{span}(Z_s)$ admit an ℓ^2 continuation.

A.3 Proof of proposition 1 and lemma 1

Throughout the proof, we take the proposition's current date to be date 0, reserving t for dates along a continuation.

The main challenge for proposition 1 is that the date-0 equilibrium equation contains $\mathbb{E}_0 X_1$, so X_0 must be chosen together with a continuation $X_1, X_2, \ldots$ after every possible future innovation history. Our proof proceeds in three parts. First, building on appendices A.1 and A.2, we show that the linearized continuation problem has a unique bounded solution. Second, the implicit function theorem gives a nonlinear continuation whose date-0 value defines the locally stable state-space solution g, and we establish its local uniqueness. Third, we use g to construct the unique sequence-space solution $\mathcal{X}$ and establish the equivalence in (4). Lemma 1 then follows as a simple corollary of uniqueness.

1. The linear continuation problem. We use the matrices P, K, and R from lemma 2 to solve the linearized equilibrium equation for bounded continuations indexed by innovation histories.

Let $\mathcal{E}$ denote the bounded support of $\bar{\epsilon}_t$. Write $\bar{\epsilon}^t = (\bar{\epsilon}_1, \ldots, \bar{\epsilon}_t)$. By a *continuation* we mean a bounded sequence $x = \{x_t\}_{t \geq 0}$, where $x_0 \in \mathbb{R}^n$ and each $x_t : \mathcal{E}^t \to \mathbb{R}^n$ is measurable for $t \geq 1$, with norm

$$\|x\|_\infty \equiv \sup_{t \geq 0} \sup_{\bar{\epsilon}^t \in \mathcal{E}^t} |x_t(\bar{\epsilon}^t)|. \tag{A.8}$$

These continuations form a Banach space, and conditional expectation is a bounded linear operator on this space.

Lemma 3 (Linear continuation)**.** Under assumption 1, let $x_{-1} \in \mathbb{R}^n$ and let r be any uniformly bounded continuation. There is a unique uniformly bounded continuation x satisfying

$$F_1 x_{t-1} + F_2 x_t + F_3 \mathbb{E}_t x_{t+1} = r_t, \qquad t \geq 0, \tag{A.9}$$

after every history. Moreover,

$$\|x\|_\infty \leq C \left(|x_{-1}| + \|r\|_\infty \right) \tag{A.10}$$

for a constant C independent of x_{-1} and r.

Proof. Let P, K, and R be the matrices from lemma 2, and set $u_t \equiv x_t - P x_{t-1}$. Using (A.6) and (A.7) in (A.9) gives

$$u_t = K^{-1} r_t + R \mathbb{E}_t u_{t+1}.$$

After L successive substitutions, the remaining term is $R^{L+1} \mathbb{E}_t u_{t+L+1}$; for any bounded solution, it vanishes as $L \to \infty$. The law of iterated expectations therefore gives

$$u_t = \sum_{j=0}^{\infty} R^j K^{-1} \mathbb{E}_t r_{t+j}, \qquad x_t = P^{t+1} x_{-1} + \sum_{k=0}^{t} P^k u_{t-k}. \tag{A.11}$$

Equation (A.11) is the stochastic counterpart of (A.5): conditional on the date-t history, future residuals enter through $\mathbb{E}_t r_{t+j}$, while $P^{t+1}x_{-1}$ accounts for an arbitrary initial state. Because the powers of P and R decay geometrically and conditional expectation does not increase the sup norm, these formulas define a bounded continuation that solves (A.9) and satisfies (A.10). For uniqueness, the difference between two solutions has $x_{-1} = 0$ and satisfies $u_t = R\mathbb{E}_t u_{t+1}$. Iterating gives $u_t = R^j \mathbb{E}_t u_{t+j}$. Since u is bounded and $R^j \to 0$, this identity implies $u_t = 0$. Hence $x_t = Px_{t-1}$, and $x_{-1} = 0$ implies $x_t = 0$. □

2. The state-space solution. We now solve the nonlinear continuation problem. For a continuation $\mathbf{Y} = (Y_0, Y_1, \ldots)$, define $\mathcal{R}$ by

$$\mathcal{R}_0(\mathbf{Y}, X_{-1}, \epsilon_0, \sigma) = F\left(X_{-1}, Y_0, \mathbb{E}_0 Y_1, \epsilon_0\right), \tag{A.12}$$

$$\mathcal{R}_t(\mathbf{Y}, X_{-1}, \epsilon_0, \sigma)(\bar{\epsilon}^t) = F\left(Y_{t-1}(\bar{\epsilon}^{t-1}), Y_t(\bar{\epsilon}^t), \mathbb{E}_t Y_{t+1}(\bar{\epsilon}^t), \sigma\bar{\epsilon}_t\right), \qquad t \geq 1. \tag{A.13}$$

Thus $\mathcal{R} = 0$ requires $\mathbf{Y}$ to satisfy the equilibrium equation at date 0 and after every possible future innovation history. At $(X_{-1}, \epsilon_0, \sigma) = (X, 0, 0)$, the constant continuation $\mathbf{Y} = (X, X, \ldots)$ sets every residual to zero. Restrict $\mathbf{Y}$ and $(X_{-1}, \epsilon_0, \sigma)$ to a sufficiently small neighborhood, so that the arguments of F in (A.12)–(A.13) lie in a compact subset of the neighborhood on which F is continuously differentiable. The derivatives of F are then bounded and uniformly continuous, so pointwise application of F, together with the bounded linear expectation operator, makes $\mathcal{R}$ continuously differentiable. If F is $\mathcal{C}^r$, then so is $\mathcal{R}$.

The derivative of $\mathcal{R}$ with respect to $\mathbf{Y}$ at the steady state is

$$[D_{\mathbf{Y}}\mathcal{R}\, x]_t = F_1 x_{t-1} + F_2 x_t + F_3 \mathbb{E}_t x_{t+1}, \qquad x_{-1} = 0.$$

Lemma 3 shows that $D_{\mathbf{Y}}\mathcal{R}$ is invertible with a bounded inverse. The implicit function theorem therefore gives a unique nearby continuation $\mathbf{Y}(X_{-1}, \epsilon_0, \sigma)$ satisfying $\mathcal{R} = 0$, with the same differentiability as F. Define

$$g(X_{-1}, \epsilon_0, \sigma) \equiv Y_0(X_{-1}, \epsilon_0, \sigma). \tag{A.14}$$

In particular, $g(X, 0, 0) = X$.

After any realization $\bar{\epsilon}_1$, the continuation from date 1 onward, $(Y_1, Y_2, \ldots)$, satisfies $\mathcal{R} = 0$ with lagged state Y_0 and current shock $\sigma\bar{\epsilon}_1$. Uniqueness gives

$$Y_1(\bar{\epsilon}_1) = g(Y_0, \sigma\bar{\epsilon}_1, \sigma),$$

and repeating this argument gives $Y_t(\bar{\epsilon}^t) = g(Y_{t-1}(\bar{\epsilon}^{t-1}), \sigma\bar{\epsilon}_t, \sigma)$ for every $t \geq 1$. Together with $\mathcal{R} = 0$, these identities show that g is a state-space solution.

Set $\epsilon_0 = \sigma = 0$. For any $a \in \mathbb{R}^n$, let $x_t = D_{X_{-1}} Y_t(X, 0, 0)a$ for $t \geq 0$ and set $x_{-1} = a$. Differentiating $\mathcal{R} = 0$ shows that x is a bounded deterministic solution of (A.9) with $r = 0$. Equation (A.11) gives $x_t = P^{t+1}a$, so $D_{X_{-1}} Y_0(X, 0, 0)a = Pa$. Since a is arbitrary, (A.14) gives

$g_X(X,0,0) = P$. Thus $\rho(P) < 1$ implies that g is locally stable at X.

It remains to establish uniqueness among continuous locally stable solutions. Let $\widetilde{g}$ be another such solution, and let $\{X_t\}_{t\geq 0}$ denote the continuation it generates. Choose m with $\|\widetilde{g}_X^m\| < 1$. The m-fold iterate of $x \mapsto \widetilde{g}(x,0,0)$ has derivative $\widetilde{g}_X^m$ at X, so it maps a sufficiently small closed ball around X strictly inside itself. By continuity, $X_0 = \widetilde{g}(X_{-1}, \epsilon_0, \sigma)$ lies in this ball whenever X_{-1} is sufficiently close to X and ϵ_0 and σ are sufficiently small. Since $\mathcal{E}$ is bounded and the deterministic m-period iterate maps the ball strictly inside itself, this property persists for small σ, uniformly over all realizations of the next m innovations. With the ball chosen small enough, the intermediate states also remain within the neighborhood of X required by the implicit function theorem. Repeating this argument in blocks of m periods shows that the entire continuation lies in the neighborhood where $\mathcal{R} = 0$ has a unique solution. Because $\widetilde{g}$ satisfies the state-space equilibrium equation, its continuation satisfies $\mathcal{R} = 0$ and therefore equals $\mathbf{Y}(X_{-1}, \epsilon_0, \sigma)$. Comparing their date-0 values gives $\widetilde{g} = g$.

3. The sequence-space solution. Let ℓ_-^∞ denote the space of bounded sequences indexed by the nonpositive integers. Write $\boldsymbol{\epsilon} = (\epsilon_0, \epsilon_{-1}, \ldots)$ and $\mathbf{X} = (X_0, X_{-1}, \ldots)$. To express X_0 as a function of the entire shock history, we solve $X_j = g(X_{j-1}, \epsilon_j, \sigma)$ simultaneously for all $j \leq 0$. For $\mathbf{X}$ near the constant sequence $(X, X, \ldots)$ and $(\boldsymbol{\epsilon}, \sigma)$ near $(0,0)$, define

$$[\Phi(\mathbf{X}, \boldsymbol{\epsilon}, \sigma)]_j = X_j - g(X_{j-1}, \epsilon_j, \sigma), \qquad j \leq 0. \tag{A.15}$$

At the deterministic steady state, $\mathbf{X} = (X, X, \ldots)$ and $(\boldsymbol{\epsilon}, \sigma) = (0,0)$, so $\Phi = 0$. Since g is continuously differentiable, so is Φ on a sufficiently small sup-norm neighborhood. Its derivative with respect to $\mathbf{X}$ is

$$[D_{\mathbf{X}}\Phi\, h]_j = h_j - g_X h_{j-1},$$

where g_X is evaluated at $(X,0,0)$. For $r \in \ell_-^\infty$, a solution to $D_{\mathbf{X}}\Phi\, h = r$ is

$$h_j = \sum_{k=0}^{\infty} g_X^k r_{j-k}. \tag{A.16}$$

The series defines a bounded h because $\rho(g_X) < 1$ implies $\sum_{k\geq 0} \|g_X^k\| < \infty$. If $D_{\mathbf{X}}\Phi\, h = 0$ and h is bounded, then $h_j = g_X^k h_{j-k}$ for every $k \geq 1$, which implies $h_j = 0$. Thus $D_{\mathbf{X}}\Phi$ has a bounded inverse.

The implicit function theorem then gives a unique $\mathbf{X}(\boldsymbol{\epsilon}, \sigma)$ near $(X, X, \ldots)$ satisfying $\Phi = 0$ for every sufficiently small bounded $\boldsymbol{\epsilon}$ and sufficiently small σ. If F is $\mathcal{C}^r$, then g, Φ, and $\mathbf{X}$ are also $\mathcal{C}^r$. Define

$$\mathcal{X}(\sigma, \epsilon_0, \epsilon_{-1}, \ldots) \equiv X_0(\boldsymbol{\epsilon}, \sigma). \tag{A.17}$$

Thus $\mathcal{X}$ is $\mathcal{C}^r$ whenever F is $\mathcal{C}^r$.

Reindexing the tails by dates $0, -1, \ldots$, the state tail $(X_{-1}, X_{-2}, \ldots)$ solves $\Phi = 0$ for the shock

tail $(\epsilon_{-1}, \epsilon_{-2}, \ldots)$. Uniqueness therefore gives

$$X_{-1}(\boldsymbol{\epsilon}, \sigma) = \mathcal{X}(\sigma, \epsilon_{-1}, \epsilon_{-2}, \ldots).$$

Substituting this identity into the date-0 component of $\Phi = 0$ gives (4). After a future innovation $\bar{\epsilon}_1$, that equation gives

$$\mathcal{X}(\sigma, \sigma\bar{\epsilon}_1, \epsilon_0, \epsilon_{-1}, \ldots) = g\left(X_0, \sigma\bar{\epsilon}_1, \sigma\right).$$

Taking conditional expectations gives $\mathbb{E}_0 X_1 = \mathbb{E}_0 g(X_0, \sigma\bar{\epsilon}_1, \sigma)$. Because g is a state-space solution, $F(X_{-1}, X_0, \mathbb{E}_0 g(X_0, \sigma\bar{\epsilon}_1, \sigma), \epsilon_0) = 0$. The preceding identity therefore gives $F(X_{-1}, X_0, \mathbb{E}_0 X_1, \epsilon_0) = 0$, which is (1) for $\mathcal{X}$.

It remains to prove uniqueness. We first show that any other local sequence-space solution must obey the recursion generated by g; its past state history must then solve the same equation $\Phi = 0$. Let $\widetilde{\mathcal{X}}$ be another local sequence-space solution and fix $\boldsymbol{\epsilon} = (\epsilon_0, \epsilon_{-1}, \ldots)$. Set $Y_0 = \widetilde{\mathcal{X}}(\sigma, \epsilon_0, \epsilon_{-1}, \ldots)$ and, after future innovations $\bar{\epsilon}^t$, set

$$Y_t(\bar{\epsilon}^t) = \widetilde{\mathcal{X}}\left(\sigma, \sigma\bar{\epsilon}_t, \ldots, \sigma\bar{\epsilon}_1, \epsilon_0, \epsilon_{-1}, \ldots\right), \qquad t \geq 1.$$

Every shock history used to define Y_t has sup norm bounded by $\max\{\|\boldsymbol{\epsilon}\|_\infty, |\sigma| \sup_{e \in \mathcal{E}} |e|\}$. Continuity of $\widetilde{\mathcal{X}}$ therefore makes $\sup_t |Y_t - X|$ small for sufficiently small $\boldsymbol{\epsilon}$ and σ, so the continuation $\mathbf{Y} = (Y_0, Y_1, \ldots)$ lies in the implicit-function-theorem neighborhood where $\mathcal{R} = 0$ has a unique solution. Since $\widetilde{\mathcal{X}}$ solves (1), the continuation $\mathbf{Y}$ satisfies $\mathcal{R} = 0$ with lagged value $\widetilde{\mathcal{X}}(\sigma, \epsilon_{-1}, \epsilon_{-2}, \ldots)$ and current shock ϵ_0. Uniqueness of the continuation satisfying $\mathcal{R} = 0$ gives $Y_0 = g(\widetilde{\mathcal{X}}(\sigma, \epsilon_{-1}, \epsilon_{-2}, \ldots), \epsilon_0, \sigma)$. By the definition of Y_0, this is (4) with $\mathcal{X}$ replaced by $\widetilde{\mathcal{X}}$.

For the fixed shock history, define $\widetilde{X}_j = \widetilde{\mathcal{X}}(\sigma, \epsilon_j, \epsilon_{j-1}, \ldots)$ for $j \leq 0$. Substitution into (4) gives $\Phi(\widetilde{\mathbf{X}}, \boldsymbol{\epsilon}, \sigma) = 0$. Continuity of $\widetilde{\mathcal{X}}$ and $\|(\epsilon_j, \epsilon_{j-1}, \ldots)\|_\infty \leq \|\boldsymbol{\epsilon}\|_\infty$ make $\sup_{j \leq 0} |\widetilde{X}_j - X|$ small. Thus $\widetilde{\mathbf{X}}$ lies in the neighborhood where $\Phi(\cdot, \boldsymbol{\epsilon}, \sigma) = 0$ has the unique solution $\mathbf{X}$. Therefore $\widetilde{\mathbf{X}} = \mathbf{X}$. Since $\boldsymbol{\epsilon}$ and σ were arbitrary, $\widetilde{\mathcal{X}} = \mathcal{X}$ throughout the neighborhood. This completes the proof of proposition 1.

Symmetry in σ and proof of lemma 1. Given the continuation $\mathbf{Y}(X_{-1}, \epsilon_0, \sigma)$ constructed in Part 2, set $\widehat{Y}_0 = Y_0$ and $\widehat{Y}_s(\bar{\epsilon}^s) = Y_s(-\bar{\epsilon}^s)$ for $s \geq 1$. Symmetry of $\bar{\epsilon}_t$ implies that the conditional expectation of $\widehat{Y}_{s+1}$ at $\bar{\epsilon}^s$ equals that of Y_{s+1} at $-\bar{\epsilon}^s$. Together with $\sigma(-\bar{\epsilon}_s) = (-\sigma)\bar{\epsilon}_s$, this gives $\mathcal{R}(\widehat{\mathbf{Y}}, X_{-1}, \epsilon_0, -\sigma) = 0$. Uniqueness of the continuation gives $\widehat{\mathbf{Y}} = \mathbf{Y}(X_{-1}, \epsilon_0, -\sigma)$. Since $\widehat{Y}_0 = Y_0$, we obtain $g(X_{-1}, \epsilon_0, \sigma) = g(X_{-1}, \epsilon_0, -\sigma)$. Hence $\Phi(\mathbf{X}, \boldsymbol{\epsilon}, \sigma) = \Phi(\mathbf{X}, \boldsymbol{\epsilon}, -\sigma)$, and uniqueness for $\Phi = 0$ gives $\mathcal{X}(\sigma, \epsilon_t, \epsilon_{t-1}, \ldots) = \mathcal{X}(-\sigma, \epsilon_t, \epsilon_{t-1}, \ldots)$. Thus both functions are even in σ. Every derivative involving an odd number of differentiations with respect to σ is therefore zero at $\sigma = 0$, whenever that derivative exists.

A.4 Volterra form of the sequence-space Taylor expansion

Proposition 1 establishes that $\mathcal{X}$ is $\mathcal{C}^r$ on a neighborhood of $(0, \mathbf{0})$ in $\mathbb{R} \times \ell^\infty_-$ whenever F is $\mathcal{C}^r$. This gives a Fréchet Taylor expansion, but does not by itself give the Volterra expansions used in the main text. For instance, for a history with finitely many nonzero shocks, linearity expresses the derivative as the sum of the responses to those shocks. On ℓ^∞_-, however, this does not imply that the infinite sum for an arbitrary bounded history converges or equals the derivative. Here, we use the stable state-space recursion from appendix A.3 to close this gap, showing that the sums over shock dates converge absolutely for every bounded history and equal the full Fréchet derivatives.

Lemma 4 (Volterra representation)**.** Suppose that assumption 1 holds. If F is $\mathcal{C}^1$, then $\sum_{j\geq 0} \|\mathcal{X}_j\| < \infty$. If F is $\mathcal{C}^2$, then in addition $\sum_{j,k\geq 0} \|\mathcal{X}_{jk}\| < \infty$. If F is $\mathcal{C}^3$, then in addition $\sum_{j\geq 0} \|\mathcal{X}_{\sigma\sigma j}\| < \infty$ and $\sum_{j,k,k'\geq 0} \|\mathcal{X}_{jkk'}\| < \infty$.

For $r = 1, 2, 3$, respectively, the expressions in (8), (11), and (19) are then the rth-order Fréchet Taylor polynomials of $\mathcal{X}$ whenever F is $\mathcal{C}^r$. All their infinite sums converge absolutely for every bounded shock history, and the remainder is $o((|\sigma| + \|\boldsymbol{\epsilon}\|_\infty)^r)$ as $(\sigma, \boldsymbol{\epsilon}) \to (0, \mathbf{0})$.

Proof. Fix $r \in \{1, 2, 3\}$ and suppose that F is $\mathcal{C}^r$. For a history $\boldsymbol{\epsilon} = (\epsilon_0, \epsilon_{-1}, \ldots)$, equation (4) is

$$\mathcal{X}(\sigma, \epsilon_0, \epsilon_{-1}, \ldots) = g\big(\mathcal{X}(\sigma, \epsilon_{-1}, \epsilon_{-2}, \ldots), \epsilon_0, \sigma\big).$$

Let $P = g_X(X, 0, 0)$. The proof of proposition 1 shows that $\rho(P) < 1$, so $\sum_{\ell\geq 0} \|P^\ell\| < \infty$.

To differentiate this recursion, let $z_i = (\sigma_i, \boldsymbol{\epsilon}^i) \in \mathbb{R} \times \ell^\infty_-$ denote perturbations to anticipated risk and the shock history, where $\boldsymbol{\epsilon}^i = (\epsilon^i_0, \epsilon^i_{-1}, \ldots)$. All derivatives of $\mathcal{X}$ below are evaluated at $(0, \mathbf{0})$. Let S remove the current shock while leaving anticipated risk unchanged, so $Sz_i = (\sigma_i, (\epsilon^i_{-1}, \epsilon^i_{-2}, \ldots))$. Differentiating m times, for $m = 1, \ldots, r$, produces one term containing the mth derivative of $\mathcal{X}$. This term comes from differentiating the inner $\mathcal{X}$ m times and then applying $g_X = P$, and is $PD^m\mathcal{X}(0, \mathbf{0})[Sz_1, \ldots, Sz_m]$. Let Q_m denote the remaining chain-rule terms, which involve only derivatives of $\mathcal{X}$ of order strictly less than m. Then

$$D^m\mathcal{X}(0, \mathbf{0})[z_1, \ldots, z_m] = PD^m\mathcal{X}(0, \mathbf{0})[Sz_1, \ldots, Sz_m] + Q_m[z_1, \ldots, z_m]. \tag{A.18}$$

For example, for $m = 1$, $Q_1[z_1] = g_\epsilon \epsilon^1_0 + g_\sigma \sigma_1$.[A-7] Iterating (A.18) gives

$$D^m\mathcal{X}(0, \mathbf{0})[z_1, \ldots, z_m] = \sum_{\ell=0}^{L} P^\ell Q_m[S^\ell z_1, \ldots, S^\ell z_m] + P^{L+1} D^m\mathcal{X}(0, \mathbf{0})[S^{L+1}z_1, \ldots, S^{L+1}z_m].$$

Since S does not increase the norm, the last term is bounded by $\|P^{L+1}\| \|D^m\mathcal{X}(0, \mathbf{0})\| \prod_i \|z_i\|$ and

[A-7] At second order, write $a_i \equiv D\mathcal{X}(0, \mathbf{0})[Sz_i]$ for the first-order change in the lagged state under perturbation z_i. The second-order chain rule then gives $Q_2[z_1, z_2] = g_{XX}[a_1, a_2] + g_{X\epsilon}[a_1]\epsilon^2_0 + g_{X\epsilon}[a_2]\epsilon^1_0 + g_{\epsilon\epsilon}\epsilon^1_0\epsilon^2_0 + g_{\sigma\sigma}\sigma_1\sigma_2$, where all derivatives of g are evaluated at $(X, 0, 0)$ and symmetry in σ eliminates the terms involving $g_{X\sigma}$ and $g_{\epsilon\sigma}$.

converges to zero. Taking $L \to \infty$ therefore gives

$$D^m \mathcal{X}(0, \mathbf{0})[z_1, \ldots, z_m] = \sum_{\ell=0}^{\infty} P^\ell Q_m[S^\ell z_1, \ldots, S^\ell z_m]. \tag{A.19}$$

We now use (A.19) to express each derivative as an absolutely convergent sum over shock dates. First, for $m = 1$, write $z = (\sigma, \boldsymbol{\epsilon})$. Symmetry gives $g_\sigma = 0$, so $D\mathcal{X}(0, \mathbf{0})[z] = \sum_{j \geq 0} P^j g_\epsilon \epsilon_{-j}$. Thus $\mathcal{X}_j = P^j g_\epsilon$, and $\sum_j \|\mathcal{X}_j\| < \infty$ because the powers of P are summable.

For $m = 2$, we first show that the coefficients of Q_2 are absolutely summable. Substituting the first-order responses into the g_{XX} term gives $\sum_{j,k \geq 0} g_{XX}[\mathcal{X}_j, \mathcal{X}_k] \epsilon^1_{-j-1} \epsilon^2_{-k-1}$. Its coefficient norms sum to at most $\|g_{XX}\| (\sum_j \|\mathcal{X}_j\|)^2 < \infty$. The same reasoning applies to the $g_{X\epsilon}$ terms, using the absolute summability of the first-order coefficients, while the $g_{\epsilon\epsilon}$ and $g_{\sigma\sigma}$ terms have only finitely many coefficients. Thus Q_2 has absolutely summable coefficients.

It remains to show that the sum over ℓ in (A.19) preserves absolute summability. Replacing z_i by $S^\ell z_i$ shifts the shock dates but leaves the sum of coefficient norms for Q_2 unchanged. Multiplication by P^ℓ multiplies this bound by $\|P^\ell\|$, so summing over ℓ gives a finite bound because $\sum_{\ell \geq 0} \|P^\ell\| < \infty$. Thus $D^2\mathcal{X}$ also has absolutely summable coefficients.

For $m = 3$, the same argument applies to Q_3, whose terms involve only first- and second-order responses, current shocks, and risk perturbations. Substituting the series for these responses gives absolutely summable coefficients, since products of absolutely convergent series are again absolutely convergent. Equation (A.19) then gives the same property for $D^3\mathcal{X}$.

Equation (A.19) holds for every bounded history, so the absolutely convergent series obtained from it equal the full Fréchet derivatives. Their coefficients are the partial derivatives of $\mathcal{X}$, as we can see by taking each z_i to be a unit change in a single shock or in σ. Substituting these series into the Taylor expansion of $\mathcal{X}$ gives the claimed expansions and remainder, with symmetry in σ from lemma 1 eliminating the terms with an odd number of σ derivatives. □

A.5 Heterogeneous-agent model details

Here we provide more details about the computation and calibration of our two heterogeneous-agent models introduced in section 2.2.

Partial equilibrium (smooth SIM) model. We provide a high-level description of the equations here, with implementation details in appendix C.2. Optimal consumption and asset policy conditional on (a, e, β, ξ) are characterized by a first-order condition $u'(c) \geq \beta \mathbb{E}_t[V_{a,t+1}(a', e', \beta')|e, \beta]$, which holds with equality unless the borrowing constraint binds. The marginal value function $V_{a,t}$ is then given by the envelope condition $V_{a,t}(a, e, \beta) = (1 + r)\mathbb{E}_\xi[u'(c(a, e, \beta, \xi))]$.

For each e, β, let $\Phi_{t-1}(a, e, \beta)$ denote the CDF of assets a saved by households at the end of period $t - 1$, conditional on income state e and patience state β realized at the beginning of period t. Also let $\varpi_{e,\beta}$ denote the mass of agents in idiosyncratic state (e, β), which we assume to always be constant at its stationary value. Given the Φ_{t-1} from the previous period, the asset policy at

t, together with the fixed Π and distribution of ξ, determines Φ_t (see equations (A.46)–(A.49)). Finally, aggregate consumption at t is given by

$$C_t = \sum_{e,\beta} \omega_{e,\beta} \int \int c_t(a, e, \beta, \xi) f(\xi) d\xi d\Phi_{t-1}(a, e, \beta).$$

In practice, we represent $V_{a,t}(a, e, \beta)$ using its value at n_a asset gridpoints a, for each of the n_e income states e and n_β patience states β, and let $\boldsymbol{v}_t$ be the vector of these $n_a \times n_e \times n_\beta$ values. We represent $\Phi_t(a', e, \beta)$ with a cubic spline interpolating over the same gridpoints for each e, β, and let $\boldsymbol{\phi}_t$ be the vector of the $n_a \times n_e \times n_\beta$ spline coefficients.

To put the system in the notation of (1), we stack these vectors, along with aggregate C_t and Z_t, in the vector $X_t \equiv (\boldsymbol{v}_t, \boldsymbol{\phi}_t, C_t, Z_t)$. We then stack the $2(n_a n_e n_\beta + 1)$ equations into the system displayed in equation (7) in the main text, where $\mathcal{V}$, Λ, and $\mathcal{C}$ represent respectively the Bellman equation, the law of motion of the distribution, and the aggregation equation above, acting on the discretized representation (see appendix C.2 for the explicit construction of these functions).

Our calibration, which is at a quarterly frequency, mostly follows Auclert et al. (2024a), except for the addition of the smooth i.i.d. income shocks. We set the interest rate to $r = 0.005$ (2% annualized), and pick a two-point process for β with support (β_L, β_H), with half of the population in each state and probability of drawing a new β state of 1% quarterly. We set $Tr = 0.14$ and $Z = 0.67$ to be consistent with U.S. tax progressivity, as discussed in the next section.

We assume that the primitive log income process is the sum of two independent components, with the persistent component following an AR(1) with persistence $\rho_e = 0.91$ at the annual level. The innovations to both processes are normally distributed with standard deviation σ_e for the persistent component and σ_ξ for the transitory component. We solve for σ_e and σ_ξ so that the cross-sectional standard deviation of log income is 0.92 overall and 0.24 in quarterly changes, with the latter taken from Ganong et al. (2025). This yields $\sigma_e = 0.1964$ and $\sigma_\xi = 0.0964$.

We set $\underline{a} = 0$, and set the asset gridpoints on which to store the value function (which are also the knots for the cubic spline representing the distribution) to be $n_a = 200$ double-exponentially spaced points. In the deterministic steady state, we pick β_L and β_H so that the first-order aggregate marginal propensity to consume $\frac{\partial C_0}{\partial Z_0}$ is equal to 0.2 and aggregate assets are $A = 4$. This procedure yields $(\beta_L, \beta_H) = (0.96, 0.99)$: as is common for these models, some patience heterogeneity is required to simultaneously hit our high MPC and high asset targets. For our aggregate income process, we assume that aggregate income follows an i.i.d. income process with $\rho_Z = 0$ and quarterly standard deviation $\sigma = 0.05$. This generates a standard deviation of aggregate consumption relative to steady-state GDP (C_t / Y_{ss}) of 1.2%, similar to the one observed in the data.[A-8]

General equilibrium (IKC) model. Our calibration of the GE model exactly matches that of our PE model, as follows. We normalize income to $Y = 1$, set government spending at $G = 0.17$ and

[A-8] We compare this to aggregate consumption divided by potential GDP in the data, removing a linear trend to account for long-run shifts in the GDP composition. The resulting time series has standard deviation of 1.3%.

let the government transfer be $Tr = 0.14$. This implies that the tax rate is $\tau = rB + G + Tr = 0.02 + 0.17 + 0.14 = 0.33$, so net-of-tax aggregate income is $Z = 0.67$. In turn, $Tr = 0.14$ and $\tau = 0.33$ are consistent with the U.S. relationship between pre- and post-tax income estimated in Auclert and Rognlie (2018). For our aggregate bond process, we assume a quarterly persistence of $\rho_B = 0.95$.

In this model, positive shocks to B_t enable the government to temporarily lower tax revenue $T_t = \tau_t Y_t$ and therefore increase post-tax household income $Z_t = Y_t - T_t$. Since households are not Ricardian, this decline in T_t in turn affects consumption, which in equilibrium affects aggregate income Y_t and feeds back into consumption C_t. Alternatively, we can describe the process in the asset market: positive shocks to bond supply require an increase in post-tax income Z_t up to the point where household asset demand matches this higher bond supply. Formally, the model is described by the vector $X_t \equiv (\boldsymbol{v}_t, \boldsymbol{\phi}_t, Z_t, B_t)$, with

$$F\left(X_{t-1}, X_t, \mathbb{E}_t\left[X_{t+1}\right], \epsilon_t\right) \equiv \begin{bmatrix} \boldsymbol{v}_t - \mathcal{V}(Z_t, \mathbb{E}_t \boldsymbol{v}_{t+1}) \\ \boldsymbol{\phi}_t - \Lambda(Z_t, \mathbb{E}_t \boldsymbol{v}_{t+1}, \boldsymbol{\phi}_{t-1}) \\ B_t - \mathcal{A}(Z_t, \mathbb{E}_t \boldsymbol{v}_{t+1}, \boldsymbol{\phi}_{t-1}) \\ B_t - B_{ss} - \rho_B(B_{t-1} - B_{ss}) - \epsilon_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{A.20}$$

where $\mathcal{A}$ represents the aggregate asset function, whose explicit construction is described in appendix C.2. By Walras's law, the asset market clearing equation in (A.20) is equivalent to the goods market clearing condition $\mathcal{C}(Z_t, \mathbb{E}_t \boldsymbol{v}_{t+1}, \boldsymbol{\phi}_{t-1}) + G - Y_t = 0$, with $Y_t = Z_t + T_t$ and $T_t = (1+r)B_{t-1} + Tr + G - B_t$.[A-9]

In this model, the endogeneity of Z_t makes a global state-space solution substantially more difficult, since the distribution of household assets now affects equilibrium income.

B Appendix to section 3

To start, lemma 4 establishes that the first-, second-, and third-order expansions used in propositions 2, 3, and 4 are the corresponding Fréchet Taylor polynomials of $\mathcal{X}$, with sums over shock dates that converge absolutely for every bounded shock history. Appendices B.1–B.3 identify the coefficients in these expansions using the perfect-foresight calculations stated in each proposition.

B.1 Proof of proposition 2

By Lemma 1, $\mathcal{X}_\sigma = 0$ at the deterministic steady state. The first-order Taylor expansion of the sequence-space solution therefore has no term in σ, leaving

$$X_t \approx X + \sum_{j=0}^{\infty} \mathcal{X}_j \epsilon_{t-j}.$$

[A-9] We use the asset market formulation in practice because it is numerically more stable.

It remains only to identify coefficients $\mathcal{X}_j$ with perfect-foresight impulse responses. By definition, $X_j^{MIT}(\nu) = \mathcal{X}(0, 0, \ldots, 0, \nu, 0, \ldots)$, where ν is in the coordinate corresponding to ϵ_{-j}. Differentiating with respect to ν at $\nu = 0$ gives

$$(X_j^{MIT})'(0) = \mathcal{X}_j.$$

This proves the first-order certainty equivalence result.

B.2 Proof of proposition 3

MIT shock terms. The size and history terms follow directly from the definitions of the perfect-foresight impulse responses. By definition, $X_j^{MIT}(\nu) = \mathcal{X}(0, 0, \ldots, 0, \nu, 0, \ldots)$, where ν occupies shock coordinate j. Differentiating twice gives

$$(X_j^{MIT})''(0) = \mathcal{X}_{jj}.$$

For $l > 0$, the response $X_j^{MIT,l}(\nu, \epsilon)$ places the newest shock ν in coordinate j and the earlier shock ϵ in coordinate $j + l$. Consequently,

$$\left.\frac{\partial^2 X_j^{MIT,l}(\nu, \epsilon)}{\partial \nu \, \partial \epsilon}\right|_{\nu=\epsilon=0} = \mathcal{X}_{j,j+l}.$$

Taking $l = k - j$ gives (15). This proves the size and history parts of proposition 3. It remains to identify the risk term $\mathcal{X}_{\sigma\sigma}$.

Risk terms. Now set all realized shocks through date t to zero, but let agents anticipate future shocks with standard deviation σ. Then

$$X_t = \mathcal{X}(\sigma, 0, 0, \ldots) = X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2 + o(\sigma^2). \tag{A.21}$$

At date $t + 1$, the new shock is $\epsilon_{t+1} = \sigma\bar{\epsilon}_{t+1}$. Using $\mathbb{E}\bar{\epsilon}_{t+1} = 0$ and $\mathbb{E}\bar{\epsilon}_{t+1}^2 = 1$,

$$\mathbb{E}_t X_{t+1} = \mathbb{E}_t \mathcal{X}(\sigma, \sigma\bar{\epsilon}_{t+1}, 0, 0, \ldots) = X + \frac{1}{2}\left(\mathcal{X}_{\sigma\sigma} + \mathcal{X}_{00}\right)\sigma^2 + o(\sigma^2) \tag{A.22}$$

where the mixed $\sigma\epsilon$ term is zero by the symmetry in σ from lemma 1. The first-order shock term also drops out after taking the expectation.

Substitute (A.21) and (A.22) into

$$F(X_{t-1}, X_t, \mathbb{E}_t X_{t+1}, \epsilon_t) = 0.$$

Since the realized shock is zero and the path is constant through date t,

$$F\left(X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2,\ X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2,\ X + \frac{1}{2}\left(\mathcal{X}_{\sigma\sigma} + \mathcal{X}_{00}\right)\sigma^2,\ 0\right) = o(\sigma^2). \tag{A.23}$$

Equating coefficients on σ^2 in (A.23) gives $(F_1 + F_2 + F_3)\mathcal{X}_{\sigma\sigma} + F_3\mathcal{X}_{00} = 0$. The matrix $F_1 + F_2 + F_3$ is the characteristic matrix at $\lambda = 1$ and is therefore invertible because the root condition in assumption 1 excludes unit roots. Hence $\mathcal{X}_{\sigma\sigma} = -(F_1 + F_2 + F_3)^{-1}F_3\mathcal{X}_{00}$. The implicit function theorem then gives a unique differentiable solution $X^{pss}(\mathcal{X}_{00}\delta)$ for δ near zero, satisfying

$$F\left(X^{pss}(\mathcal{X}_{00}\delta),\ X^{pss}(\mathcal{X}_{00}\delta),\ X^{pss}(\mathcal{X}_{00}\delta) + \mathcal{X}_{00}\delta,\ 0\right) = 0 \tag{A.24}$$

where $X^{pss}(0) = X$. Differentiating at $\delta = 0$ gives $\frac{d}{d\delta}X^{pss}(\mathcal{X}_{00}\delta)\Big|_{\delta=0} = -(F_1 + F_2 + F_3)^{-1}F_3\mathcal{X}_{00}$, and hence

$$\mathcal{X}_{\sigma\sigma} = \frac{d}{d\delta}X^{pss}(\mathcal{X}_{00}\delta)\Big|_{\delta=0}. \tag{A.25}$$

Finally, $\mathcal{X}_{00} = (X_0^{MIT})''(0)$, so the perturbation in (A.24) is available from the perfect-foresight impact response. Taking $\delta = \sigma^2/2$ matches (A.21) to second order, completing the certainty correspondence for the risk term.

Alternative F formulation. Suppose instead that the model is written, as in Lan and Meyer-Gohde (2013b) and much of the rest of the literature, as

$$\mathbb{E}_t F(X_{t-1}, X_t, X_{t+1}, \epsilon_t) = 0. \tag{A.26}$$

Set realized shocks through date t to zero. Then at date t, we have the second-order approximation $X_t \approx X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2$. But the future variable inside F is random:

$$X_{t+1} \approx X + \mathcal{X}_0\epsilon_{t+1} + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2 + \frac{1}{2}\mathcal{X}_{00}\epsilon_{t+1}^2 + o(\sigma^2).$$

Substituting into (A.26), expanding, and taking expectations gives

$$(F_1 + F_2 + F_3)\mathcal{X}_{\sigma\sigma} + F_3\mathcal{X}_{00} + F_{33}[\mathcal{X}_0, \mathcal{X}_0] = 0.$$

The first two terms are the same as in (A.23), but the last term (where $F_{33}[\mathcal{X}_0, \mathcal{X}_0]$ denotes applying the Hessian F_{33} to $\mathcal{X}_0$ on both input dimensions) is new. It appears because we take the expectation of applying F to the random X_{t+1}, so that the curvature of F matters. Thus, in this formulation,

$$\mathcal{X}_{\sigma\sigma} = -(F_1 + F_2 + F_3)^{-1}\left(F_3\mathcal{X}_{00} + F_{33}[\mathcal{X}_0, \mathcal{X}_0]\right).$$

By contrast, in our formulation $F(X_{t-1}, X_t, \mathbb{E}_t X_{t+1}, \epsilon_t) = 0$, the random first-order term $\mathcal{X}_0\epsilon_{t+1}$ is averaged before F is applied, so the extra term does not appear. Indeed, Lan and Meyer-Gohde (2013b) have precisely this additional term in their equation (36); in their notation, it appears as $f_{y^{+2}}(y_0 \otimes y_0)\mathbb{E}_t(\varepsilon_{t+1} \otimes \varepsilon_{t+1})$. Without this term, their equation (36) would parallel our $\mathcal{X}_{\sigma\sigma}$.

Although either formulation is valid and can be used to solve a model when the equations are written in a compatible way, our approach makes possible the simple certainty-correspondence

statement of proposition 3, where the risk term $\mathcal{X}_{\sigma\sigma}$ can be obtained by perturbing expectations at the steady state with the size dependence of MIT shocks. With the (A.26) formulation, the extra term $F_{33}[\mathcal{X}_0, \mathcal{X}_0]$ cannot be obtained from size dependence alone.

B.3 Proof of proposition 4

MIT shock terms. For a single shock, $X_j^{MIT}(\nu)$ places ν in shock coordinate j, so

$$(X_j^{MIT})'''(0) = \mathcal{X}_{jjj}.$$

For $k > j$, the two-shock response $X_j^{MIT,k-j}(\nu, \epsilon)$ places ν in coordinate j and ϵ in coordinate k. Hence

$$\left.\frac{\partial^3 X_j^{MIT,k-j}(\nu,\epsilon)}{\partial \nu^2\, \partial \epsilon}\right|_{\nu=\epsilon=0} = \mathcal{X}_{jjk}, \qquad \left.\frac{\partial^3 X_j^{MIT,k-j}(\nu,\epsilon)}{\partial \nu\, \partial \epsilon^2}\right|_{\nu=\epsilon=0} = \mathcal{X}_{jkk}.$$

Finally, for $k' > k > j$, set $l = k - j$ and $l' = k' - j$. Then $0 < l < l'$, and $X_j^{MIT,k-j,k'-j}(\nu, \epsilon, \epsilon')$ places the newest, middle, and oldest shocks in coordinates j, k, and k', respectively. The gaps between the oldest and middle shocks and between the middle and newest shocks are $l' - l = k' - k$ and $l = k - j$. Therefore

$$\left.\frac{\partial^3 X_j^{MIT,k-j,k'-j}(\nu,\epsilon,\epsilon')}{\partial \nu\, \partial \epsilon\, \partial \epsilon'}\right|_{\nu=\epsilon=\epsilon'=0} = \mathcal{X}_{jkk'}.$$

These identities prove the pure MIT-shock parts of proposition 4. It remains to identify the mixed risk terms $\mathcal{X}_{j\sigma\sigma}$.

Mixed risk terms. The coefficient $\mathcal{X}_{j\sigma\sigma}$ records how background risk changes the response at horizon j. We first derive its linear equation and then separate its risky-steady-state and time-varying precautionary contributions.

Fixing $j \geq 0$, we will differentiate the date-t equilibrium condition

$$F(X_{t-1}, X_t, \mathbb{E}_t X_{t+1}, \epsilon_t) = 0$$

with respect to $\epsilon_{t-j}, \sigma, \sigma$, evaluating at the deterministic steady state, writing

$$u = (X_{t-1}, X_t, \mathbb{E}_t X_{t+1}, \epsilon_t)$$

for the stacked argument of F. Derivatives such as F_u and F_{uu} are taken with respect to this stacked argument and evaluated at $(X, X, X, 0)$, and we will use subscripts j and σ on u for derivatives with respect to ϵ_{t-j} and σ. Symmetry gives $u_\sigma = 0$ at the deterministic steady state. Differentiating $F(u) = 0$ therefore gives

$$F_u u_{j\sigma\sigma} + F_{uu}[u_{\sigma\sigma}, u_j] = 0 \tag{A.27}$$

where $F_{uu}[v, w]$ denotes the Hessian F_{uu} evaluated on vectors v and w.

We now evaluate u_j, $u_{\sigma\sigma}$, and $u_{j\sigma\sigma}$ as needed for (A.27). First, we have

$$u_j = \left(\mathcal{X}_{j-1},\ \mathcal{X}_j,\ \mathcal{X}_{j+1},\ \mathbf{1}_{j=0}\right),$$

with $\mathcal{X}_{-1} = 0$. Then

$$u_{\sigma\sigma} = \left(\mathcal{X}_{\sigma\sigma},\ \mathcal{X}_{\sigma\sigma},\ \mathcal{X}_{\sigma\sigma} + \mathcal{X}_{00},\ 0\right),$$

where we note that here the third entry includes $\mathcal{X}_{00}$ because the new shock at $t+1$ is $\sigma\bar{\epsilon}_{t+1}$, and we take the expectation of its square in $\mathbb{E}_t X_{t+1}$. Finally,

$$u_{j\sigma\sigma} = \left(\mathcal{X}_{j-1,\sigma\sigma},\ \mathcal{X}_{j\sigma\sigma},\ \mathcal{X}_{j+1,\sigma\sigma} + \mathcal{X}_{j+1,00},\ 0\right),$$

where $\mathcal{X}_{-1,\sigma\sigma} = 0$. At date $t+1$, the shock ϵ_{t-j} occupies lag $j+1$ and the new innovation occupies lag 0; differentiating twice with respect to σ and taking the expected square of the new innovation therefore contributes $\mathcal{X}_{j+1,00}$.

Substituting these expressions into (A.27) gives

$$F_1\mathcal{X}_{j-1,\sigma\sigma} + F_2\mathcal{X}_{j\sigma\sigma} + F_3\mathcal{X}_{j+1,\sigma\sigma} + F_{uu}[u_{\sigma\sigma}, u_j] + F_3\mathcal{X}_{j+1,00} = 0. \tag{A.28}$$

Lemma 4 implies that the two forcing sequences $\{F_{uu}[u_{\sigma\sigma}, u_j]\}_{j\geq 0}$ and $\{F_3\mathcal{X}_{j+1,00}\}_{j\geq 0}$ are absolutely summable. Proposition 1 therefore gives a unique solution $\{\mathcal{X}_{j\sigma\sigma}\}_{j\geq 0} \in \ell^2$ to (A.28), with $\mathcal{X}_{-1,\sigma\sigma} = 0$. We now show that solving this equation with each forcing sequence separately gives the RSS and TVP expressions in (26) and (27), respectively.

RSS term. The F_{uu} term captures how the first-order response changes when it is computed around the risky steady state. Let δ shift the expected-future term by $\mathcal{X}_{00}\delta$, as in (A.24), and let ν be the size of the one-time shock. Write

$$X_k^{pt,rss}(\nu, \delta) \equiv X_k^{pt}\left(\{\mathcal{X}_{00}\delta, \mathcal{X}_{00}\delta, \ldots\}, \nu\right).$$

By the definition in the main text, this path satisfies

$$F\left(X_{k-1}^{pt,rss}(\nu,\delta),\ X_k^{pt,rss}(\nu,\delta),\ X_{k+1}^{pt,rss}(\nu,\delta) + \mathcal{X}_{00}\delta,\ \nu\mathbf{1}_{k=0}\right) = 0, \qquad k \geq 0, \tag{A.29}$$

with

$$X_{-1}^{pt,rss}(\nu,\delta) = \lim_{k\to\infty} X_k^{pt,rss}(\nu,\delta) = X^{pss}(\mathcal{X}_{00}\delta),$$

since $\Delta_\infty = \mathcal{X}_{00}\delta$. After subtracting the terminal value $X^{pss}(\mathcal{X}_{00}\delta)$, the implicit function theorem and proposition 1 give a unique nearby path in ℓ^2 that is twice continuously differentiable in (ν, δ).

By stationarity, place the shock ϵ_{t-j} at date zero; the date-t equilibrium condition is then the $k = j$ equation of (A.29). At $\delta = 0$, the path is the ordinary MIT response, so the ν derivative of

the four arguments of F at zero is

$$u_j = \left(\mathcal{X}_{j-1},\ \mathcal{X}_j,\ \mathcal{X}_{j+1},\ \mathbf{1}_{j=0}\right),$$

with $\mathcal{X}_{-1} = 0$. At $\nu = 0$, the path is constant: $X_k^{pt,rss}(0,\delta) = X^{pss}(\mathcal{X}_{00}\delta)$. Using (17), the δ derivative of the four arguments at zero is therefore

$$u_{\sigma\sigma} = \left(\mathcal{X}_{\sigma\sigma},\ \mathcal{X}_{\sigma\sigma},\ \mathcal{X}_{\sigma\sigma} + \mathcal{X}_{00},\ 0\right).$$

Define, as in (26),

$$\left(\mathcal{X}_{j\sigma\sigma}\right)^{rss} \equiv \left.\frac{\partial^2 X_j^{pt,rss}(\nu,\delta)}{\partial\delta\,\partial\nu}\right|_{\delta=\nu=0}.$$

Differentiating (A.29) gives

$$F_1(\mathcal{X}_{j-1,\sigma\sigma})^{rss} + F_2(\mathcal{X}_{j\sigma\sigma})^{rss} + F_3(\mathcal{X}_{j+1,\sigma\sigma})^{rss} + F_{uu}[u_{\sigma\sigma}, u_j] = 0. \tag{A.30}$$

Although the boundary state moves with δ, it is independent of ν. Hence

$$\left(\mathcal{X}_{-1,\sigma\sigma}\right)^{rss} = \left.\frac{\partial^2 X^{pss}(\mathcal{X}_{00}\delta)}{\partial\delta\,\partial\nu}\right|_{\delta=\nu=0} = 0.$$

The mixed derivative of the centered path belongs to ℓ^2. Since the centering term is independent of ν, the same holds for $\{(\mathcal{X}_{j\sigma\sigma})^{rss}\}_{j\geq 0}$. Equation (A.30) is thus (A.28) with the same linear operator and boundary conditions, and only the F_{uu} forcing term.

TVP term. The TVP contribution allows the expectation adjustment $\delta\mathcal{X}_{00,k+1}$ to vary with k. Its forcing in (A.28) is $F_3\mathcal{X}_{j+1,00}$. To obtain it, write

$$X_k^{pt,tvp}(\delta) \equiv X_k^{pt}\left(\{\mathcal{X}_{00,1}\delta, \mathcal{X}_{00,2}\delta, \ldots\}, 0\right).$$

Lemma 4 implies that $\{\mathcal{X}_{00,k+1}\}_{k\geq 0}$ is absolutely summable. Thus $\Delta_\infty = 0$, and the boundary conditions in the main text become

$$X_{-1}^{pt,tvp}(\delta) = \lim_{k\to\infty} X_k^{pt,tvp}(\delta) = X^{pss}(0) = X.$$

The implicit function theorem and proposition 1 give a unique nearby differentiable path, with $\{X_k^{pt,tvp}(\delta) - X\}_{k\geq 0} \in \ell^2$, satisfying

$$F\left(X_{k-1}^{pt,tvp}(\delta),\ X_k^{pt,tvp}(\delta),\ X_{k+1}^{pt,tvp}(\delta) + \delta\mathcal{X}_{00,k+1},\ 0\right) = 0, \qquad k \geq 0.$$

Define, as in (27),

$$(\mathcal{X}_{j\sigma\sigma})^{tvp} \equiv \left.\frac{dX_j^{pt,tvp}(\delta)}{d\delta}\right|_{\delta=0}.$$

Differentiating at $\delta = 0$, and using symmetry of mixed partials to write $\mathcal{X}_{00,j+1} = \mathcal{X}_{j+1,00}$, gives

$$F_1(\mathcal{X}_{j-1,\sigma\sigma})^{tvp} + F_2(\mathcal{X}_{j\sigma\sigma})^{tvp} + F_3(\mathcal{X}_{j+1,\sigma\sigma})^{tvp} + F_3\mathcal{X}_{j+1,00} = 0. \tag{A.31}$$

Since the initial state is independent of δ, $(\mathcal{X}_{-1,\sigma\sigma})^{tvp} = 0$. The derivative path belongs to ℓ^2 because $\{X_k^{pt,tvp}(\delta) - X\}_{k\geq 0}$ is differentiable in ℓ^2. Equation (A.31) is therefore (A.28) with the same linear operator and boundary conditions, and only the $F_3\mathcal{X}_{j+1,00}$ forcing term.

Putting it together. Adding (A.30) and (A.31) gives (A.28). The RSS and TVP derivative sequences both vanish at index -1 and lie in ℓ^2, as shown above. The same holds for $\{\mathcal{X}_{j\sigma\sigma}\}_{j\geq 0}$ by lemma 4. Uniqueness in proposition 1 therefore gives

$$\mathcal{X}_{j\sigma\sigma} = (\mathcal{X}_{j\sigma\sigma})^{rss} + (\mathcal{X}_{j\sigma\sigma})^{tvp}. \tag{A.32}$$

This is the third-order certainty correspondence for the mixed $\epsilon_{t-j}\sigma^2$ term. The *rss* part captures how being at the risky steady state changes the first-order response to a shock, while the *tvp* part captures how anticipation of future mixed size-and-history effects $\mathcal{X}_{j+1,00}$ changes the same response.

B.4 Proof of propositions 5 and 6

Second-order GIRF. In the second-order expansion of X_{t+j}, the terms involving ϵ_t but no shock dated after t are

$$\mathcal{X}_j\epsilon_t + \frac{1}{2}\mathcal{X}_{jj}\epsilon_t^2 + \sum_{k>j}\mathcal{X}_{jk}\epsilon_t\epsilon_{t-(k-j)}. \tag{A.33}$$

Terms pairing ϵ_t with a shock dated after t have conditional expectation zero. Before observing ϵ_t, the expectation of (A.33) is $\mathcal{X}_{jj}\sigma^2/2$. Replacing ϵ_t by ν in (A.33) and subtracting the pre-observation expectation gives the second-order GIRF in (30):

$$IRF_j^{(2)}(\nu,\sigma) = \mathcal{X}_j\nu + \frac{1}{2}\mathcal{X}_{jj}(\nu^2 - \sigma^2) + \sum_{k>j}\mathcal{X}_{jk}\nu\epsilon_{t-(k-j)}.$$

Third-order GIRF. The third-order terms add a complication relative to the second-order proof: future shocks still have mean zero, but a future shock can now enter squared and be multiplied by ϵ_t. Thus, after taking expectations over shocks dated after t, those terms can depend on the realization of ϵ_t.

Beyond the second-order terms already handled above, the terms in the third-order expansion

of X_{t+j} whose expectations can change when ϵ_t is observed are

$$\frac{1}{2}\mathcal{X}_{\sigma\sigma j}\sigma^2\epsilon_t + \frac{1}{6}\mathcal{X}_{jjj}\epsilon_t^3 + \frac{1}{2}\sum_{k>j}\mathcal{X}_{jjk}\epsilon_t^2\epsilon_{t-(k-j)} + \frac{1}{2}\sum_{k>j}\mathcal{X}_{jkk}\epsilon_t\epsilon_{t-(k-j)}^2 + \sum_{k>j}\sum_{k'>k}\mathcal{X}_{jkk'}\epsilon_t\epsilon_{t-(k-j)}\epsilon_{t-(k'-j)} + \frac{1}{2}\sum_{i<j}\mathcal{X}_{iij}\epsilon_{t+j-i}^2\epsilon_t \quad \text{(A.34)}$$

where we note that the last sum is over shocks dated after t.

After observing $\epsilon_t = \nu$, the future shocks in the last term have variance σ^2, so the conditional expectation of these third-order terms is

$$\frac{1}{2}\mathcal{X}_{\sigma\sigma j}\sigma^2\nu + \frac{1}{6}\mathcal{X}_{jjj}\nu^3 + \frac{1}{2}\sum_{k>j}\mathcal{X}_{jjk}\nu^2\epsilon_{t-(k-j)} + \frac{1}{2}\sum_{k>j}\mathcal{X}_{jkk}\nu\epsilon_{t-(k-j)}^2 + \sum_{k>j}\sum_{k'>k}\mathcal{X}_{jkk'}\nu\epsilon_{t-(k-j)}\epsilon_{t-(k'-j)} + \frac{1}{2}\sigma^2\nu\sum_{i<j}\mathcal{X}_{iij}. \quad \text{(A.35)}$$

Before observing ϵ_t, symmetry gives $E_{t-1}\epsilon_t = E_{t-1}\epsilon_t^3 = 0$, while $E_{t-1}\epsilon_t^2 = \sigma^2$. Hence the only nonzero contribution from the displayed third-order terms is

$$\frac{1}{2}\sum_{k>j}\mathcal{X}_{jjk}\sigma^2\epsilon_{t-(k-j)}.$$

Subtracting this forecast from the forecast after observing $\epsilon_t = \nu$, and adding the second-order part already derived, gives the third-order GIRF in (31):

$$IRF_j^{(3)}(\nu,\sigma) = IRF_j^{(2)}(\nu,\sigma) + \frac{1}{2}\mathcal{X}_{\sigma\sigma j}\sigma^2\nu + \frac{1}{2}\sigma^2\nu\sum_{i<j}\mathcal{X}_{iij} + \frac{1}{6}\mathcal{X}_{jjj}\nu^3 + \frac{1}{2}\sum_{k>j}\mathcal{X}_{jjk}(\nu^2-\sigma^2)\epsilon_{t-(k-j)} + \frac{1}{2}\sum_{k>j}\mathcal{X}_{jkk}\nu\epsilon_{t-(k-j)}^2 + \sum_{k>j}\sum_{k'>k}\mathcal{X}_{jkk'}\nu\epsilon_{t-(k-j)}\epsilon_{t-(k'-j)}.$$

Ergodic moments. Fix $h \geq 0$. Write

$$X_t(\sigma) \equiv \mathcal{X}(\sigma, \sigma\bar{\epsilon}_t, \sigma\bar{\epsilon}_{t-1}, \ldots)$$

and let primes denote total derivatives of this expression with respect to σ, evaluated at $\sigma = 0$ while holding fixed the sequence $\{\bar{\epsilon}_t, \bar{\epsilon}_{t-1}, \ldots\}$. Then

$$X_t' = \sum_i \mathcal{X}_i \bar{\epsilon}_{t-i}$$
$$X_t'' = \mathcal{X}_{\sigma\sigma} + \sum_{i,j} \mathcal{X}_{ij} \bar{\epsilon}_{t-i} \bar{\epsilon}_{t-j}$$
$$X_t''' = 3 \sum_i \mathcal{X}_{\sigma\sigma i} \bar{\epsilon}_{t-i} + \sum_{i,j,k} \mathcal{X}_{ijk} \bar{\epsilon}_{t-i} \bar{\epsilon}_{t-j} \bar{\epsilon}_{t-k}.$$

Absolute summability from lemma 4 justifies interchanging these sums with the expectations below. To expand $\Sigma_h(\sigma)$, we first center $X_t(\sigma)$,

$$\tilde{X}_t(\sigma) \equiv X_t(\sigma) - \mu(\sigma), \qquad \mu(\sigma) \equiv \mathbb{E}[X_t(\sigma)],$$

so that $\Sigma_h(\sigma) = \mathbb{E}[\tilde{X}_t(\sigma)\tilde{X}_{t-h}(\sigma)^\top]$. Since $\mu' = \mathbb{E}[X_t'] = 0$ and $\tilde{X}_t(0) = 0$, we have

$$\tilde{X}_t(\sigma) = \sigma X_t' + \frac{\sigma^2}{2}(X_t'' - \mu'') + \frac{\sigma^3}{6}(X_t''' - \mu''') + \cdots,$$

where all derivatives are evaluated at $\sigma = 0$. Multiplying the expansions for t and $t-h$,

$$\begin{aligned}\Sigma_h(\sigma) = {} & \sigma^2 \mathbb{E}[X_t'(X_{t-h}')^\top] \\ & + \frac{\sigma^3}{2}\mathbb{E}\left[X_t'(X_{t-h}'' - \mu'')^\top + (X_t'' - \mu'')(X_{t-h}')^\top\right] \\ & + \sigma^4 \left\{\frac{1}{4}\mathbb{E}[(X_t'' - \mu'')(X_{t-h}'' - \mu'')^\top] + \frac{1}{6}\mathbb{E}[X_t'(X_{t-h}''' - \mu''')^\top] + \frac{1}{6}\mathbb{E}[(X_t''' - \mu''')(X_{t-h}')^\top]\right\} + \cdots.\end{aligned}$$

By lemma 1, $\mathcal{X}$ is even in σ; sign symmetry of the normalized shock history therefore implies that $(X_t(-\sigma), X_{t-h}(-\sigma))$ and $(X_t(\sigma), X_{t-h}(\sigma))$ have the same distribution. Hence $\Sigma_h(-\sigma) = \Sigma_h(\sigma)$, so the coefficients on σ^3 and σ^5 are zero and

$$\Sigma_h(\sigma) = \sigma^2 \Sigma_{h,2} + \sigma^4 \Sigma_{h,4} + O(\sigma^6).$$

Identifying these coefficients with terms from the expansion above, we get

$$\Sigma_{h,2} = \mathbb{E}[X_t'(X_{t-h}')^\top] = \sum_i \mathcal{X}_{h+i} \mathcal{X}_i^\top,$$

and

$$\Sigma_{h,4} = \frac{1}{4}\operatorname{Cov}(X_t'', X_{t-h}'') + \frac{1}{6}\mathbb{E}[X_t'(X_{t-h}''')^\top] + \frac{1}{6}\mathbb{E}[X_t'''(X_{t-h}')^\top].$$

The centering of X''' drops out in the last two terms because $\mathbb{E}[X_t'] = \mathbb{E}[X_{t-h}'] = 0$.

For independent normal $\bar{\epsilon}$s, Isserlis's theorem implies the identity

$$\mathbb{E}[\bar{\epsilon}_a\bar{\epsilon}_b\bar{\epsilon}_c\bar{\epsilon}_d] = \mathbf{1}_{a=b}\mathbf{1}_{c=d} + \mathbf{1}_{a=c}\mathbf{1}_{b=d} + \mathbf{1}_{a=d}\mathbf{1}_{b=c},$$

which continues to hold if the $\bar{\epsilon}$ are not normal but have the same moments as a normal up to the fourth. This, in turn, implies

$$\frac{1}{4}\operatorname{Cov}(X_t'', X_{t-h}'') = \frac{1}{2}\sum_{i,j}\mathcal{X}_{h+i,h+j}\mathcal{X}_{ij}^\top.$$

Substituting the expression for X_t''' above gives

$$\frac{1}{6}\mathbb{E}[X_t'(X_{t-h}''')^\top] = \frac{1}{2}\sum_i \mathcal{X}_{h+i}\mathcal{X}_{\sigma\sigma i}^\top + \frac{1}{2}\sum_{i,j}\mathcal{X}_{h+i}\mathcal{X}_{ijj}^\top,$$

and

$$\frac{1}{6}\mathbb{E}[X_t'''(X_{t-h}')^\top] = \frac{1}{2}\sum_i \mathcal{X}_{\sigma\sigma,h+i}\mathcal{X}_i^\top + \frac{1}{2}\sum_{i,j}\mathcal{X}_{h+i,jj}\mathcal{X}_i^\top.$$

Combining these terms gives the expression for $\Sigma_{h,4}$ in proposition 6:

$$\Sigma_{h,4} = \frac{1}{2}\sum_{i,j}\mathcal{X}_{h+i,h+j}\mathcal{X}_{ij}^\top + \frac{1}{2}\sum_i\left(\mathcal{X}_{h+i}\mathcal{X}_{\sigma\sigma i}^\top + \mathcal{X}_{\sigma\sigma,h+i}\mathcal{X}_i^\top\right) + \frac{1}{2}\sum_{i,j}\left(\mathcal{X}_{h+i}\mathcal{X}_{ijj}^\top + \mathcal{X}_{h+i,jj}\mathcal{X}_i^\top\right).$$

B.5 Equivalence to pruned state space

In this section, we work out the implications of the nonlinear equivalence between the state and the sequence-space solution in proposition 1 for the relationship between the derivatives $\mathcal{X}_i, \mathcal{X}_{ij}, \mathcal{X}_{ijk}, \mathcal{X}_{\sigma\sigma}, \mathcal{X}_{\sigma\sigma j}$ of the sequence-space solution on the one hand, and on the other the derivatives of the state space solution at the first order (g_X, g_ϵ), the second order ($g_{XX}, g_{X\epsilon}, g_{\epsilon\epsilon}, g_{\sigma\sigma}$), and the third order ($g_{XXX}, g_{XX\epsilon}, g_{X\epsilon\epsilon}, g_{\epsilon\epsilon\epsilon}, g_{X\sigma\sigma}, g_{\epsilon\sigma\sigma}$). We use this to establish the equivalence between the Volterra expansion and the so-called "pruned" state-space solution at orders 1, 2 and 3. In the next section, we will further use this result to validate our certainty correspondence propositions numerically on a known example with a small state space.

Coefficient recursions. We recall equation (4) from proposition 1:

$$\mathcal{X}(\sigma, \epsilon_t, \epsilon_{t-1}, \ldots) = g\left(\mathcal{X}(\sigma, \epsilon_{t-1}, \epsilon_{t-2}, \ldots), \epsilon_t, \sigma\right).$$

Differentiating this identity around the deterministic steady state gives us recursions that relate derivatives of the sequence-space solution $\mathcal{X}$ and the state-space solution g. For instance, differentiating (4) once with respect to ϵ_{t-j} gives $\mathcal{X}_j = g_X\mathcal{X}_{j-1} + g_\epsilon\mathbf{1}_{j=0}$, which provides a simple way to build the first-order sequence-space solution from the state-space solution. (Here, and also below, we adopt the convention that any coefficients with negative indices are zero.)

Differentiating twice with respect to σ, recalling that single σ derivatives are zero, gives an equation that can be solved for the second-order risk term:

$$\mathcal{X}_{\sigma\sigma} = g_X \mathcal{X}_{\sigma\sigma} + g_{\sigma\sigma} \qquad \iff \qquad \mathcal{X}_{\sigma\sigma} = (I - g_X)^{-1} g_{\sigma\sigma}. \tag{A.36}$$

Equation (A.36) corresponds to de Groot (2013)'s direct formula for computing the risky steady-state from the second-order state space solution coefficients g_X and $g_{\sigma\sigma}$.[A-10]

Differentiating with respect to ϵ_{t-j} and ϵ_{t-k} gives a recursion for the other second-order terms:

$$\mathcal{X}_{jk} = g_X \mathcal{X}_{j-1,k-1} + g_{XX}(\mathcal{X}_{j-1}, \mathcal{X}_{k-1}) + g_{X\epsilon}(\mathcal{X}_{j-1})\mathbf{1}_{k=0} + g_{X\epsilon}(\mathcal{X}_{k-1})\mathbf{1}_{j=0} + g_{\epsilon\epsilon}\mathbf{1}_{j=0}\mathbf{1}_{k=0}. \tag{A.37}$$

As usual, we write e.g. $g_{XX}(\mathcal{X}_{j-1}, \mathcal{X}_{k-1})$ to denote applying the vector-valued Hessian g_{XX} to $\mathcal{X}_{j-1}$ and $\mathcal{X}_{k-1}$. Here, we also write $g_{X\epsilon}(\mathcal{X}_{j-1})\mathbf{1}_{k=0}$ to denote applying $g_{X\epsilon}$ to $\mathcal{X}_{j-1}$ (for X) and the scalar $\mathbf{1}_{k=0}$ (for ϵ).

Next, for third order, taking a mixed partial with respect to ϵ_{t-j}, σ, and σ gives us a recursion for the mixed risk-shock term:

$$\mathcal{X}_{j\sigma\sigma} = g_X \mathcal{X}_{j-1,\sigma\sigma} + g_{XX}(\mathcal{X}_{j-1}, \mathcal{X}_{\sigma\sigma}) + g_{X\epsilon}(\mathcal{X}_{\sigma\sigma})\mathbf{1}_{j=0} + g_{X\sigma\sigma}(\mathcal{X}_{j-1}) + g_{\epsilon\sigma\sigma}\mathbf{1}_{j=0}. \tag{A.38}$$

Finally, to get a recursion for the general triple-shock third-order terms $\mathcal{X}_{jkl}$, we can differentiate with respect to ϵ_{t-j}, ϵ_{t-k}, and ϵ_{t-l}. This gives the following somewhat unwieldy recursion:

$$\begin{aligned}
\mathcal{X}_{jkl} = {} & g_X \mathcal{X}_{j-1,k-1,l-1} \\
& + g_{XX}(\mathcal{X}_{j-1,k-1}, \mathcal{X}_{l-1}) + g_{XX}(\mathcal{X}_{j-1,l-1}, \mathcal{X}_{k-1}) + g_{XX}(\mathcal{X}_{k-1,l-1}, \mathcal{X}_{j-1}) \\
& + g_{X\epsilon}(\mathcal{X}_{j-1,k-1})\mathbf{1}_{l=0} + g_{X\epsilon}(\mathcal{X}_{j-1,l-1})\mathbf{1}_{k=0} + g_{X\epsilon}(\mathcal{X}_{k-1,l-1})\mathbf{1}_{j=0} \\
& + g_{XXX}(\mathcal{X}_{j-1}, \mathcal{X}_{k-1}, \mathcal{X}_{l-1}) \\
& + g_{XX\epsilon}(\mathcal{X}_{j-1}, \mathcal{X}_{k-1})\mathbf{1}_{l=0} + g_{XX\epsilon}(\mathcal{X}_{j-1}, \mathcal{X}_{l-1})\mathbf{1}_{k=0} + g_{XX\epsilon}(\mathcal{X}_{k-1}, \mathcal{X}_{l-1})\mathbf{1}_{j=0} \\
& + g_{X\epsilon\epsilon}(\mathcal{X}_{j-1})\mathbf{1}_{k=0}\mathbf{1}_{l=0} + g_{X\epsilon\epsilon}(\mathcal{X}_{k-1})\mathbf{1}_{j=0}\mathbf{1}_{l=0} + g_{X\epsilon\epsilon}(\mathcal{X}_{l-1})\mathbf{1}_{j=0}\mathbf{1}_{k=0} \\
& + g_{\epsilon\epsilon\epsilon}\mathbf{1}_{j=0}\mathbf{1}_{k=0}\mathbf{1}_{l=0}.
\end{aligned} \tag{A.39}$$

Note, however, that (A.39) is long mostly to handle all the base cases where one or more of the indices is zero.

Equivalence with pruning. The coefficient recursions above become the pruned state-space recursions from Andreasen et al. (2018) once we evaluate the sequence-space expansions along a realized shock history. For instance, at second order, write the first- and second-order contribu-

[A-10] More specifically, de Groot (2013) suggests solving for the difference between the risky and the deterministic steady state x^* by solving the quadratic $C + Bx^* + A\text{vec}(x^*x^{*\prime}) = 0$, where $C = g_{\sigma\sigma}\frac{\sigma^2}{2}$ and $B = -(I - g_X)$. Since the term $A\text{vec}(x^*x^{*\prime})$ is $O(\sigma^4)$, to second-order this gives $x^* = (I - g_X)^{-1} g_{\sigma\sigma}\frac{\sigma^2}{2}$. Since further $x^* = \mathcal{X}_{\sigma\sigma}\frac{\sigma^2}{2}$, we recover (A.36).

tions to (11) as

$$X_t^{[1]} \equiv \sum_{j\geq 0} \mathcal{X}_j \epsilon_{t-j}, \qquad X_t^{[2]} \equiv \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2 + \frac{1}{2}\sum_{j,k\geq 0} \mathcal{X}_{jk}\epsilon_{t-j}\epsilon_{t-k},$$

so that the full second-order approximation is $X_t^{(2)} = X + X_t^{[1]} + X_t^{[2]}$.

Summing the first-order coefficient recursion over the realized shocks gives

$$X_t^{[1]} = g_X X_{t-1}^{[1]} + g_\epsilon \epsilon_t.$$

To get the second-order equation, multiply (A.36) by $\sigma^2/2$, multiply (A.37) by $\epsilon_{t-j}\epsilon_{t-k}/2$, and sum over $j, k \geq 0$. This gives

$$X_t^{[2]} = g_X X_{t-1}^{[2]} + \frac{1}{2}g_{XX}(X_{t-1}^{[1]}, X_{t-1}^{[1]}) + g_{X\epsilon}(X_{t-1}^{[1]})\epsilon_t + \frac{1}{2}g_{\epsilon\epsilon}\epsilon_t^2 + \frac{1}{2}g_{\sigma\sigma}\sigma^2.$$

This is the standard second-order pruned state-space recursion: we calculate new second-order interactions from the first-order path, and the result is then propagated forward linearly by g_X.

For third order, let $X_t^{[3]}$ denote the terms that first appear in (19). Multiplying (A.38) by $\sigma^2\epsilon_{t-j}/2$, multiplying (A.39) by $\epsilon_{t-j}\epsilon_{t-k}\epsilon_{t-l}/6$, and summing over their indices gives

$$\begin{aligned} X_t^{[3]} = {} & g_X X_{t-1}^{[3]} + g_{XX}(X_{t-1}^{[1]}, X_{t-1}^{[2]}) + g_{X\epsilon}(X_{t-1}^{[2]})\epsilon_t \\ & + \frac{1}{6}g_{XXX}(X_{t-1}^{[1]}, X_{t-1}^{[1]}, X_{t-1}^{[1]}) + \frac{1}{2}g_{XX\epsilon}(X_{t-1}^{[1]}, X_{t-1}^{[1]})\epsilon_t \\ & + \frac{1}{2}g_{X\epsilon\epsilon}(X_{t-1}^{[1]})\epsilon_t^2 + \frac{1}{6}g_{\epsilon\epsilon\epsilon}\epsilon_t^3 + \frac{1}{2}g_{X\sigma\sigma}(X_{t-1}^{[1]})\sigma^2 + \frac{1}{2}g_{\epsilon\sigma\sigma}\epsilon_t\sigma^2. \end{aligned}$$

This is Andreasen et al. (2018)'s third-order pruned recursion. Because g_X is stable, this recursion has a unique bounded causal solution on the infinite shock history. Together with the second-order calculation, this shows that, for $k = 2, 3$, the kth-order sequence-space and pruned state-space solutions coincide exactly.[A-11]

This is not a new result: Lombardo and Uhlig (2018) previously established it in a very related context, and used it as a motivation for pruning in the state space in the first place.[A-12] The upshot is that in the sequence space, there is no need to add an auxiliary pruning procedure on top of the perturbation solution; instead, the stability given by pruning is already built into the perturbation solution itself (see also Bhandari et al. 2023).

B.6 Validating the certainty correspondence

The equivalence between the pruned state-space and the sequence-space solution established in section B.5 provides us with a procedure for validating our certainty correspondence results.

[A-11] Although we do not formally demonstrate it here, this presumably continues to hold for all orders $k > 3$.

[A-12] Their concept differs from ours by assuming only a finite history of shocks, starting at date 0, so that the sequence-space solution is not stationary, but the logic of why sequence-space and pruned state-space perturbation are equivalent is essentially unchanged.

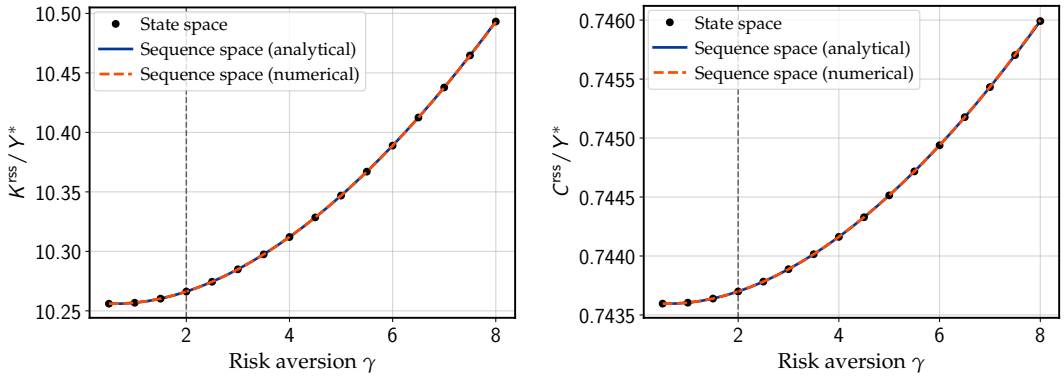
Consider a particular small-scale model of the form (1) with known derivatives. First, use standard algorithms (e.g., Schmitt-Grohé and Uribe 2004, de Groot 2013, and Andreasen et al. 2018), together with equations (A.36)–(A.39), to obtain the coefficients $\mathcal{X}_j^{state}, \mathcal{X}_{jk}^{state}, \mathcal{X}_{jkk'}^{state}, \mathcal{X}_{\sigma\sigma}^{state}, \mathcal{X}_{\sigma\sigma j}^{state}$ that prevail in the third-order Volterra expansion implied by the state-space solution. Separately, using perfect-foresight impulse responses and perturbed-expectations steady states and transitions, calculate $\mathcal{X}_j^{seq}, \mathcal{X}_{jk}^{seq}, \mathcal{X}_{jkk'}^{seq}, \mathcal{X}_{\sigma\sigma}^{seq}, \mathcal{X}_{\sigma\sigma j}^{seq}$ implied by propositions 2, 3 and 4—that is, by the certainty correspondence at orders up to 3. Finally, validate numerically the equality of all the coefficients $\mathcal{X}^{state} = \mathcal{X}^{seq}$, for all coefficients j, k, k' in some range.

Note that such a test simultaneously validates the certainty-correspondence results in propositions 2, 3 and 4, the equivalence between sequence-space and pruned state-space solution in section B.5, and our implementation routines. The computational procedures used to calculate the two sets of $\mathcal{X}$ coefficients are completely distinct: the first require standard state-space machinery (solving a stable matrix quadratic for the first-order, and generalized Sylvester equations for higher-order state-space coefficients, then performing the coefficient recursions implied by pruning), while the second require perfect-foresight routines (perfect-foresight transitions and perturbed expectations steady-state and transitions).

As an illustration, we perform this test on the neoclassical growth model—our Example 1 in section 2.2.[A-13] For our baseline simulation, we use a quarterly calibration with $\alpha = 0.36$, $\delta = 0.025$, and $\beta = 0.99$, and a log TFP process with persistence $\rho_a = 0.979$ and innovation standard deviation $\sigma = 0.0072$, in the spirit of King and Rebelo (1999). We pick a baseline value of $\gamma = 2$ but also experiment with alternative values of γ in the range of 0.5 to 8. To calculate the sequence-space derivatives, we use two approaches, paralleling those we advocate in section 4: first, a numerical differentiation of nonlinear perfect-foresight routines with symmetric differences ("numerical") and second, a set of analytical perfect-foresight derivatives, up to the numerical inversion of a sequence-space Jacobian ("analytical"). In all our perfect-foresight exercises, we chose the truncation horizon to be $T = 800$. For numerical derivatives, we follow section 4.4 in doing symmetric numerical differentiation and using polarization formulas for mixed derivatives. We choose step sizes of $h_1 = 10^{-4}$ for the first order, $h_2 = 10^{-3}$ for the second and $h_3 = 2 \times 10^{-2}$, as well as $h_e = 10^{-4}$ when differentiating the perturbed-expectations steady-state and transition routines.

Figure B.1 shows the second-order approximation to the risky steady state level of capital and consumption normalized by the deterministic level of output, $\frac{X}{Y} + \mathcal{X}_{\sigma\sigma}\frac{\sigma^2}{2Y}$, obtained from all three approaches. Note that, in the deterministic steady state, we have $\frac{K}{Y} = \frac{\alpha}{1/\beta - 1 + \delta} = 10.2561$ and $\frac{C}{Y} = 1 - \delta\frac{K}{Y} = 0.7436$ irrespective of risk aversion γ. As risk aversion γ increases, holding patience β and therefore the steady-state real interest rate fixed, precautionary savings rises, resulting in a higher risky steady-state level of capital and therefore also consumption. Figure B.1 shows that these effects are relatively modest for our calibrated value of σ. Importantly, while all three lines

[A-13] We have also performed similar tests of our routines on two other standard models: an RBC model with endogenous labor, and a model with consumption habit formation and quadratic capital adjustment costs.



(a) Capital K (b) Consumption C

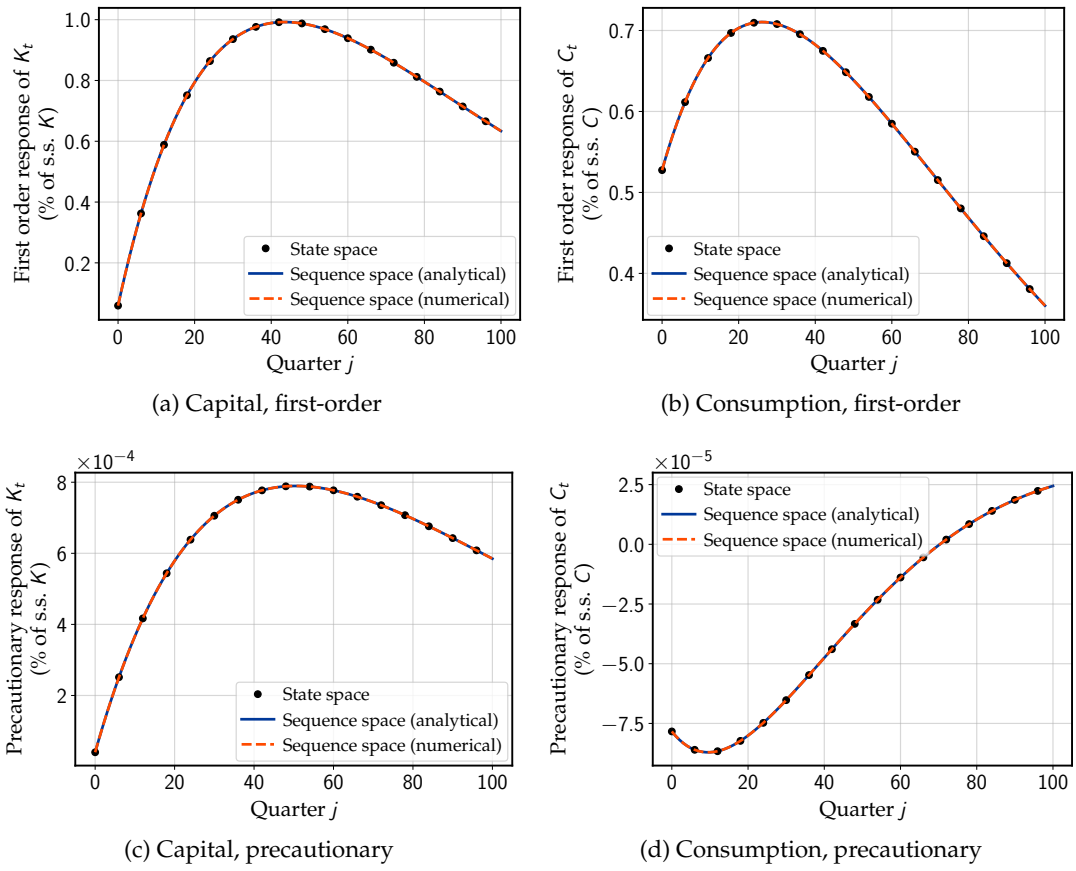
Figure B.1: Effect of risk aversion on risky steady-state capital and consumption in the neoclassical model

on each graph involve very different approaches to calculating $\mathcal{X}_{\sigma\sigma}$, they all produce the same outcome, validating this part of the certainty correspondence.

The top row of figure B.2 shows the first order impulse response of consumption and capital after a 1% innovation to TFP (i.e. a shock ϵ of about 1.39 standard deviations). Those exhibit well-known shapes and magnitudes (see e.g. King and Rebelo 1999), with consumption rising by a little more than 0.5% on impact and peaking around quarter 25, and capital peaking at an increase of about 1% around quarter 45. There is very close agreement between all three approaches to calculating the first-order impulse responses, reflecting certainty equivalence.

The bottom row of figure B.2 then reports the impulse response of precautionary consumption and capital $\mathcal{X}_{\sigma\sigma j}$ after the same 1% innovation to TFP. Concretely, these graphs trace out the path, in (19), of $\mathcal{X}_{\sigma\sigma j}\frac{\sigma^2\epsilon}{2X}$ for our calibrated $\sigma = 0.72\%$ and $\epsilon = 1\%$, reported in percent. This includes the *rss* term–the contribution to the impulse response from starting at the risky steady state rather than the deterministic steady state–and the *tvp* term, that is, the response of precautionary savings to the TFP innovation. Capital responds positively and with a similar hump-shape as in the first-order solution, reflecting a positive contribution from both terms. By contrast, consumption responds negatively initially, reflecting a dominant negative contribution of the *tvp* term (precautionary savings rises and consumption falls) more than offsetting a positive contribution of the *rss* term. Note that these terms are very small compared to the first-order impulse response: the time-varying precautionary terms are at most 0.1% of the corresponding first-order response. This is consistent with existing results in the literature that the first-order solution approximates the non-linear dynamics of the neoclassical growth model in standard calibrations (e.g. Christiano 1990, Taylor and Uhlig 1990, King and Rebelo 1999). Note also that all three approaches to calculating the $\mathcal{X}_{\sigma\sigma j}$ term deliver visually indistinguishable results, validating this part of the third-order certainty correspondence.

The numerical agreement presented in figures B.1 and B.2 is not isolated to the terms $\mathcal{X}_j$, $\mathcal{X}_{\sigma\sigma}$ and $\mathcal{X}_{\sigma\sigma j}$, although the latter two are perhaps the most surprising part of the certainty correspondence. Indeed, we have checked that all coefficients $\mathcal{X}_j$ and $\mathcal{X}_{\sigma\sigma j}$ agree for indices $j = 0, \ldots, 100$,



(a) Capital, first-order

(b) Consumption, first-order

(c) Capital, precautionary

(d) Consumption, precautionary

Figure B.2: Impulse responses to 1% TFP shock: first-order and precautionary (neoclassical growth model)

that all coefficients $\mathcal{X}_{jk}$ agree for indices $j, k = 0, \ldots, 30$, and all coefficients $\mathcal{X}_{jkk'}$ agree for indices $j, k, k' = 0, \ldots, 10$. The maximal absolute error between the state-space and the analytical sequence-space terms is 6.4×10^{-10}. The maximal absolute error between the state-space and the numerical sequence-space terms is 7.3×10^{-4}, reflecting finite-difference error in the third-order and mixed risk-history calculations.

B.7 Multiple exogenous shocks

The scalar-shock notation in section 2 extends directly to multiple independent exogenous shocks. Let the last argument of F be $\epsilon_t = (\epsilon_{t,1}, \ldots, \epsilon_{t,m}) \in \mathbb{R}^m$, with $\epsilon_{t,a} = \sigma \bar{\epsilon}_{t,a}$.[A-14] The assumptions on F, local existence, and differentiability are otherwise unchanged. Assume that the $\bar{\epsilon}_{t,a}$ are symmetric around zero, have unit variance and bounded support, and are independent across dates and components:

$$\mathbb{E}\bar{\epsilon}_{t,a} = 0, \qquad \mathbb{E}\bar{\epsilon}_{t,a}^2 = 1, \qquad \{\bar{\epsilon}_{t,a}\}_{t,a} \text{ are mutually independent.}$$

[A-14] The common σ is only a normalization: any differences in variance can be absorbed into how ϵ_t enters F.

We add a component index to each shock derivative: $\mathcal{X}^a_j$ is the derivative with respect to shock a dated $t-j$, while $\mathcal{X}^{ab}_{jk}$ is the derivative with respect to shocks a and b dated $t-j$ and $t-k$, and similarly at higher order. With this convention, (11) and (19) keep the same structure, with sums over shock components added to the sums over dates.

Since $\mathbb{E}[\epsilon_{t,a}\epsilon_{s,b}] = \sigma^2 \mathbf{1}_{t=s}\mathbf{1}_{a=b}$, each expected squared scalar shock in the risk terms is replaced by the sum over componentwise variances. In (17), the expectation perturbation $\mathcal{X}_{00}\delta$ becomes $\delta \sum_a \mathcal{X}^{aa}_{00}$, so (17) holds with this replacement. At third order, (24) holds for each component a:

$$\mathcal{X}^a_{\sigma\sigma j} = (\mathcal{X}^a_{\sigma\sigma j})^{rss} + (\mathcal{X}^a_{\sigma\sigma j})^{tvp}.$$

The risky steady state is common to every *rss* term, but the response is component-specific. For the *tvp* term, the perturbed-expectations transition for component a uses the forcing $\delta \sum_b \mathcal{X}^{abb}_{j+1,00}$, since the future squared shock can be in any component b.

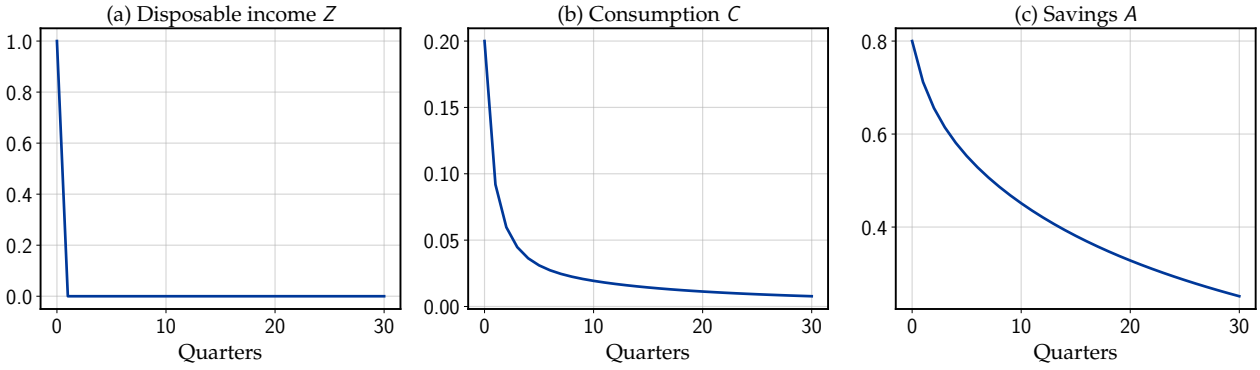
The GIRF formulas change in a similar way. If the date-t innovation is $\nu = (\nu_1, \ldots, \nu_m)$, then (30) becomes

$$IRF^{(2)}_j(\nu, \sigma) = \sum_a \mathcal{X}^a_j \nu_a + \frac{1}{2}\sum_{a,b} \mathcal{X}^{ab}_{jj}(\nu_a\nu_b - \sigma^2 \mathbf{1}_{a=b}) + \sum_{k>j}\sum_{a,b} \mathcal{X}^{ab}_{jk}\nu_a \epsilon_{t-(k-j),b}.$$

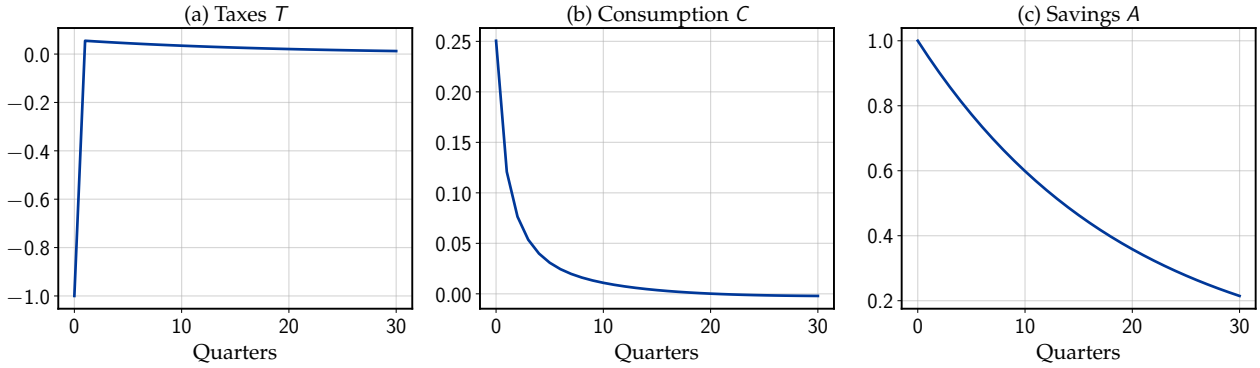
The third-order formula (31) is obtained in the same way. For example, the centered term $\mathcal{X}_{jjk}(\nu^2 - \sigma^2)\epsilon_{t-(k-j)}$ becomes $\sum_{a,b,c} \mathcal{X}^{abc}_{jjk}(\nu_a\nu_b - \sigma^2\mathbf{1}_{a=b})\epsilon_{t-(k-j),c}$.

The moment formulas in proposition 6 also extend by matching both dates and components. The leading covariance term becomes $\Sigma_{h,2} = \sum_{i,a} \mathcal{X}^a_{h+i}(\mathcal{X}^a_i)^\top$. In the fourth-order term, each repeated shock index is replaced by a repeated date-component pair: for example, $\sum_{i,j} \mathcal{X}_{h+i,h+j}\mathcal{X}^\top_{ij}$ becomes $\sum_{i,j,a,b} \mathcal{X}^{ab}_{h+i,h+j}(\mathcal{X}^{ab}_{ij})^\top$, and $\sum_{i,j} \mathcal{X}_{h+i}\mathcal{X}^\top_{ijj}$ becomes $\sum_{i,j,a,b} \mathcal{X}^a_{h+i}(\mathcal{X}^{abb}_{ijj})^\top$, with the transposed terms adjusted analogously. For the fourth-order term in proposition 6, we also assume the corresponding fourth-moment condition $\mathbb{E}\bar{\epsilon}^4_{t,a} = 3$; together with independence, this gives the same fourth-moment identities used in the scalar case.

B.8 Additional quantitative results: general equilibrium IKC model



(a) Disposable income shock (PE model)



(b) Government bonds shock (GE model)

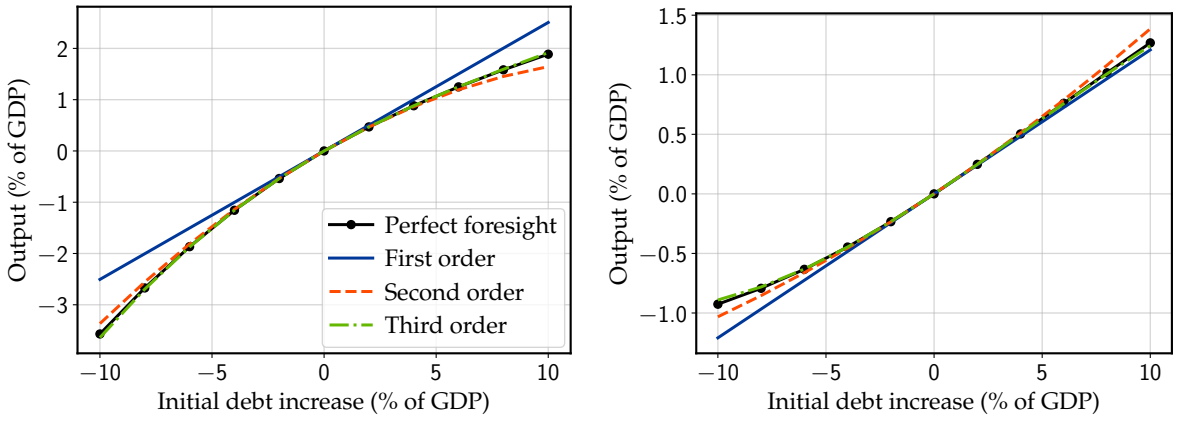
Figure B.3: First-order impulse response functions

Figure B.3 displays the shocks and resulting impulses in our heterogeneous-agent models. The top panel shows the PE model, with the shock to aggregate income Z on the left. The bottom panel shows the GE model, with the shock to aggregate B with persistence $\rho_B = 0.95$ (which must also be aggregate assets in general equilibrium) on the right. As can be seen by comparing the top and bottom panel, the underlying shocks to net labor income Z are very similar in the two models—in particular, the impact effect on aggregate taxes T is the same. However, the income shock has present discounted value of one in the PE model, while the tax shock has zero present discounted value in the GE model, as it must to satisfy the intertemporal budget constraint.

Figures B.4–B.7 provide the counterpart of figures 2–5 in the general equilibrium IKC model (the exception is figure 6, which requires a global solution which we do not have for that model). The general patterns are very similar as those from the PE model, because the same underlying effects are operative in both.

One exception is that the risky steady state in figure B.5 is now expressed in terms of an effect on aggregate income. Here, as risk increases, agents attempt to save more and consume less, but in

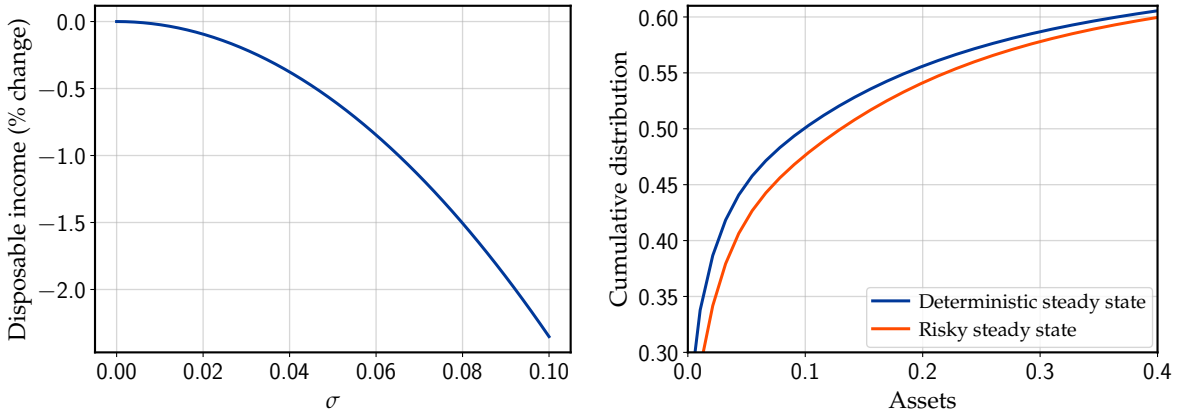
general equilibrium monetary policy maintains the real interest rate constant at the deterministic steady state r, requiring a decline in equilibrium income to clear the goods and asset market. This effect is nontrivial (a little under 2.5% for a shock standard deviation of 0.1); this effect is mitigated in our quantitative HANK model of section 5.1, since there we have a more responsive monetary policy rule.



(a) Response at $t = 0$, GE

(b) Response at $t = 1$, GE

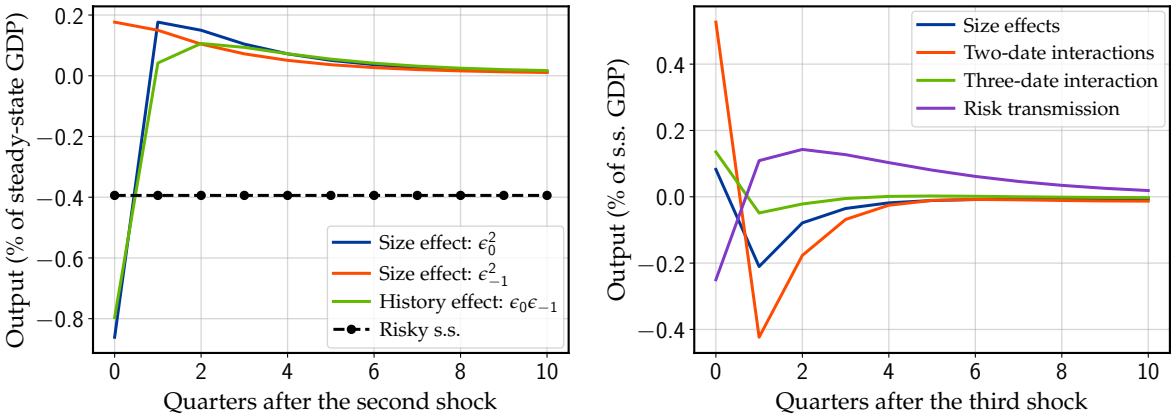
Figure B.4: Size dependence in the response of output to bond shocks (GE model)



(a) Disposable income

(b) Wealth distribution

Figure B.5: The risky steady state: aggregate income and aggregate uncertainty (GE model)



(a) Second order

(b) Third order

Figure B.6: Decomposing second and third order impulse on consumption (GE model)

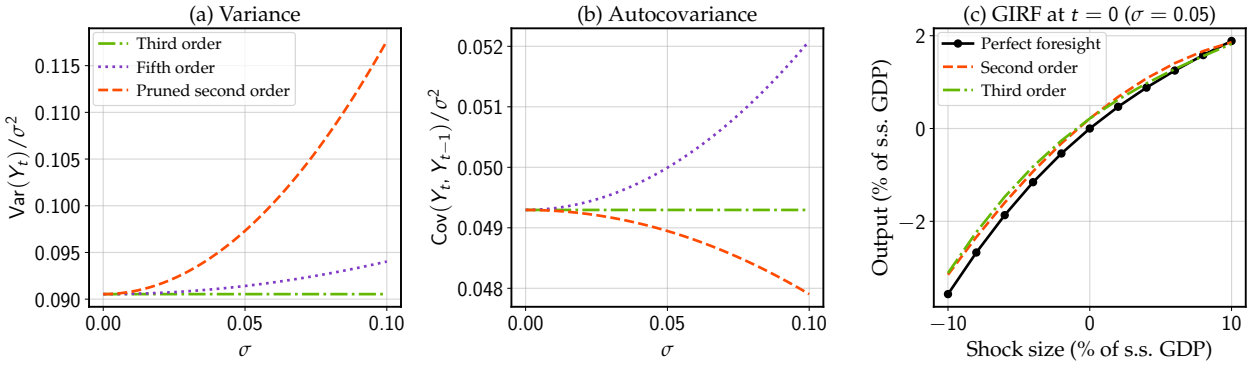


Figure B.7: Ergodic moments and GIRF (IKC model)

C Smooth standard incomplete markets (SIM) model

C.1 Kinks and higher-order differentiation: the problem with non-smooth models

C.1.1 Motivation

To motivate our smooth version of the standard incomplete markets model, we first discuss why the standard incomplete markets model can be problematic for higher-order differentiation without smoothing. The central issue is that there is a kink in the policy function at the asset level where the borrowing constraint becomes binding. At this kink, the second derivative of the consumption function is infinite. Furthermore, when income risk is discrete—including when a continuous income process is discretized via Rouwenhorst or a similar procedure—this kink leads to *secondary kinks* at other asset levels. These occur at the asset levels where, if a certain discrete income path is realized, the borrowing constraint will become exactly binding in $n > 1$ periods.[A-15]

[A-15] Even in discretizations with a decent number of points, these secondary kinks are often quite pronounced, because there is still a high probability of remaining at the same income state for the next several periods. One benefit of

Much of the overall nonlinearity of the consumption function is concentrated at these kinks, which makes differentiation potentially problematic: the second derivative at these kinks is infinite, while the second derivative away from these kinks will understate the overall curvature of the function. If we integrate over a smooth, continuous distribution of households, however, the resulting aggregate consumption (as a function, say, of a lump-sum transfer) is nevertheless well-behaved, and its higher derivatives will reflect the nonlinearity from kinks.

Unfortunately, typical calibrations of the standard incomplete markets model, with a discrete Markov chain representing idiosyncratic income, do not have such a distribution over assets. Indeed, their stationary distributions consist of countably many mass points: with probability 1, every household has hit the borrowing constraint at some previous date, and households' current asset levels are then determined by the finite history of income states received since the most recent visit to the constraint.[A-16] In practice, a large fraction of the distribution often concentrates at a limited number of mass points—including the large mass point at the constraint itself.[A-17] In implementation, this is often made worse by the standard "lottery" method of discretizing the distribution (Young, 2010), which constrains the distribution to consist entirely of mass points.[A-18] This makes higher derivatives of aggregate consumption unreliable. In the unlikely event that a mass of households is at a kink point of the policy function, then these derivatives will be infinite; otherwise, they will understate the curvature of aggregate consumption.

In appendix C.2, we show how to fix these pathologies in practice with a "smooth model", which includes a smoothly distributed iid component of income, and where we carefully perform numerical integration with respect to this component separately on both sides of the primary kink. This approach eliminates secondary kinks, and ensures a smooth distribution of cash on hand, resulting in a smooth aggregate consumption function that is amenable to perturbation.[A-19]

For the remainder of this section, we will illustrate these concepts with a stark, analytically tractable model.

C.1.2 Analytical illustration

Consider a household with log utility, discount factor $\beta < 1$, no income risk, income normalized to 1, facing a constant interest rate of $r = 0$ and a borrowing constraint. Given incoming assets

continuous-time solutions is that they avoid these secondary kinks—although there is still an important kink at the constraint. Also see Iskhakov, Jørgensen, Rust and Schjerning (2017) for more discussion of secondary kinks, which in their context are kinks in the value function rather than the policy function.

[A-16] Indeed, Le Grand and Ragot (2022) have used this insight as part of a computational strategy.

[A-17] See some discussion and visualization of mass points away from the constraint in the Lecture 1 notebook at `https://github.com/shade-econ/nber-workshop-2023/`.

[A-18] See Bhandari et al. (2023) for a critique of this method in the context of higher-order perturbation, as well as Bilal (2023). Bayer et al. (2026b) offer an alternative approach to iterating on the distribution that is better-suited to higher-order perturbation.

[A-19] For an example of using heterogeneity to smooth out a kink and facilitate perturbation in a very different context, see Schmitt-Grohé and Uribe (2022)'s model of downward nominal wage rigidity.

a_{t-1}, this household chooses consumption c_t and assets a_t carried into the next period:

$$\max_{\{c_t, a_t\}_{t\geq 0}} \sum_{t=0}^{\infty} \beta^t \log c_t \qquad \text{subject to} \qquad c_t + a_t = 1 + a_{t-1}, \quad a_t \geq 0, \quad a_{-1} = a.$$

Given the net interest rate of $r = 0$, the interior Euler equation is $c_{t+1} = \beta c_t$. The household is impatient relative to the interest rate, so it eventually runs down its assets and reaches the constraint.[A-20]

Suppose a household with initial assets a first reaches the constraint after n periods. Consumption follows a path $c_t = \beta^t c_0$ for $t = 0, \ldots, n-1$, and all initial assets have been spent by the end of these n periods. Defining

$$S_n = 1 + \beta + \cdots + \beta^{n-1},$$

the corresponding budget constraint is $S_n c_0 = a + n$. Current consumption is therefore

$$c(a) = \frac{a+n}{S_n}, \qquad a \in [a^{n-1}, a^n].$$

The kink point a^n is the boundary at which the household arrives at zero assets after n periods and its consumption path implies $c_n = 1$, exactly equal to next period's income. Thus $c_0 = \beta^{-n}$, which together with the budget constraint gives

$$a^n = \frac{\beta^{-n} - 1}{1 - \beta} - n.$$

Consumption is continuous and piecewise linear, with MPC $1/S_n$ on the nth segment. Note in this example, the *entire* nonlinearity of the consumption function is at the kink points.

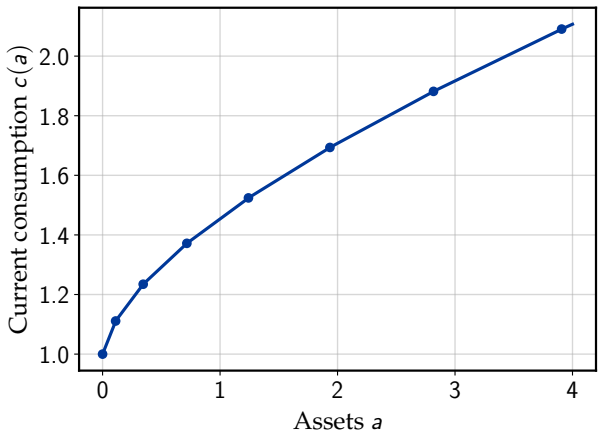
Figure C.1(a) plots this consumption function for an assumed $\beta = 0.9$, with solid dots denoting the kink points. For the aggregation below, we will further suppose that initial assets are lognormally distributed with mean 0.5 and log standard deviation 0.5; panel (b) overlays this density on the individual MPC schedule.

Aggregation and discretization. An aggregate transfer T shifts each household from assets a to $a + T$. If f is the smooth lognormal asset density, aggregate current consumption is

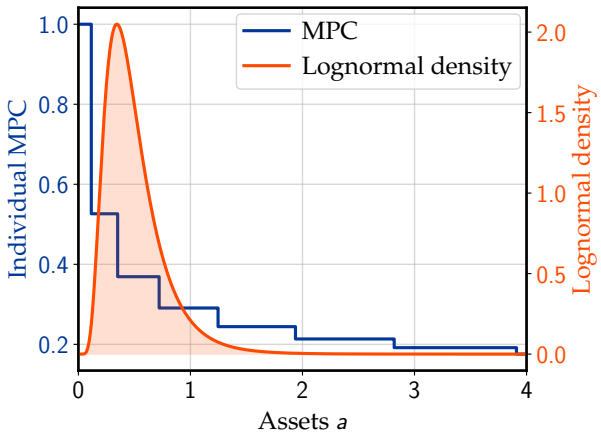
$$C(T) = \int_0^{\infty} c(a + T) f(a)\, da.$$

We compare this distribution in figure C.2 with a 21-point Rouwenhorst discretization that matches mean assets and the standard deviation of log assets. Note that the nonlinear consumption functions $C(T)$ implied by the smooth density and discretization are visually indistinguishable in panel (b).

[A-20] Note this model is not stationary, so contrary to the rest of the paper, its derivatives are taken around an arbitrary initial distribution rather than around the steady state.

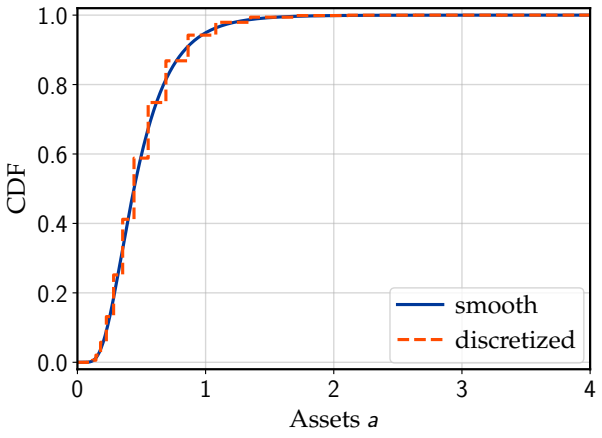


(a) Individual consumption

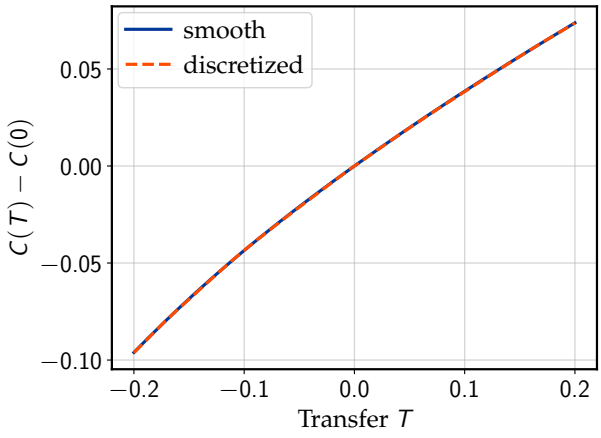


(b) Individual MPC and asset density

Figure C.1: Individual consumption function and marginal propensity to consume

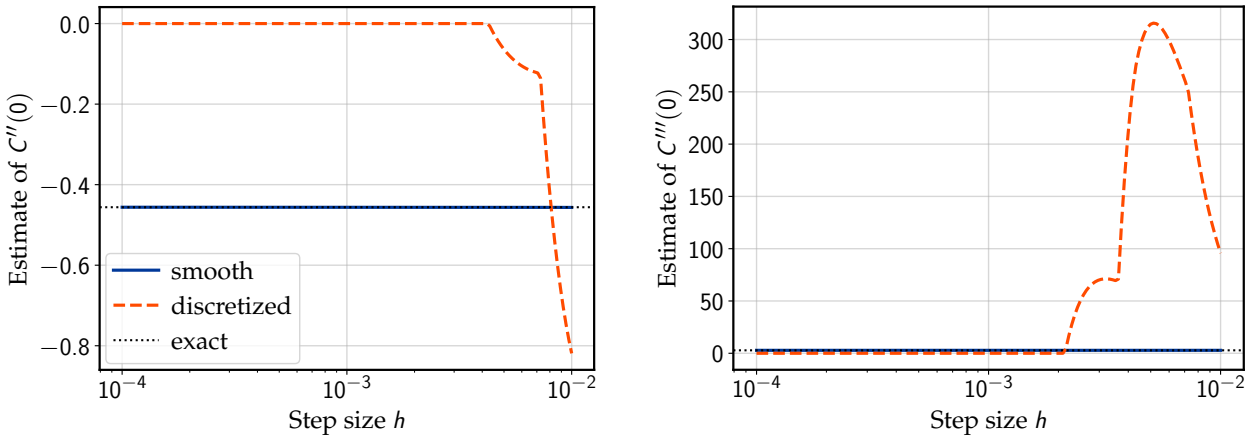


(a) Asset CDFs



(b) Aggregate consumption $C(T)$

Figure C.2: Asset distributions and implied aggregate consumption



(a) Second derivative $C''(0)$

(b) Third derivative $C'''(0)$

Figure C.3: Numerical estimates of higher derivatives of aggregate consumption

Higher-order numerical differentiation. Let F and f denote the CDF and density of assets. At kink a^q, the individual MPC changes by

$$\Delta m_q = \frac{1}{S_{q+1}} - \frac{1}{S_q}.$$

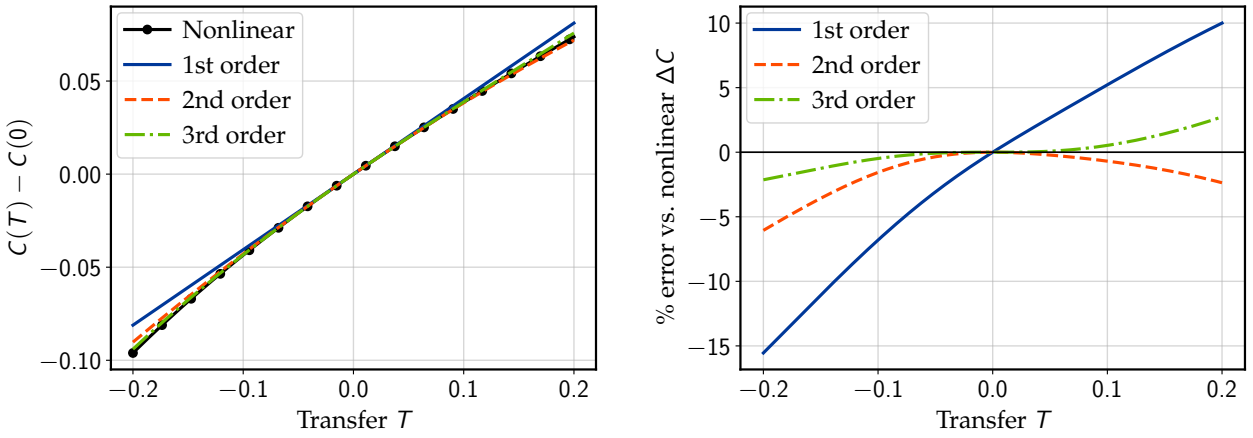
The aggregate derivatives at $T = 0$ are therefore

$$C''(0) = \sum_{q \geq 1} \Delta m_q f(a^q), \qquad C'''(0) = -\sum_{q \geq 1} \Delta m_q f'(a^q).$$

which we can calculate directly using the smooth density f.

Figure C.3 compares these exact derivatives with derivatives obtained using numerical differentiation of $C(T)$, implemented using centered finite differences over various step sizes $h \in [10^{-4}, 10^{-2}]$ as in figure 8, when $C(T)$ is calculated using either the smooth density or its discretization. Although these two versions of $C(T)$ were visually indistinguishable in figure C.2(b), their higher-order numerical derivatives are not: the derivatives for the discretized version of $C(T)$ are unreliable and erratic, while the derivatives for the smooth $C(T)$ agree with the exact computation.

The reason is closely related to our earlier discussion: when $C(T)$ is calculated using a discretized distribution, its higher-order numerical derivatives at 0 depend on whether any of the mass points a are close to a kink. If they are not at a kink, so that the finite differences (i.e. $a + h$ and $a - h$ for second order) do not cross any kinks for small h, the estimated derivatives will be exactly zero. This is indeed the case for small h in both panels of figure C.3. But once h becomes sufficiently large, then the finite differences start to cross kinks, at which higher derivatives are locally infinite. This results in erratic estimates for $C''(0)$ and $C'''(0)$, with the latter being especially extreme.



(a) Nonlinear response and approximations

(b) Approximation errors

Figure C.4: Taylor approximations to aggregate consumption

Taylor approximation. Figure C.4(a) compares first-, second-, and third-order Taylor approximations around $T = 0$ with the nonlinear change in consumption over transfers from -0.2 to 0.2. Panel (b) reports the corresponding errors relative to the nonlinear consumption change $\Delta C = C(T) - C(0)$. The first-order approximation is visibly inaccurate for larger transfers, with errors reaching roughly 10–15% at the endpoints. The second-order approximation captures most of the curvature, while the third-order approximation keeps the absolute error below about 3% throughout the range, although errors begin to grow near the boundaries.[A-21]

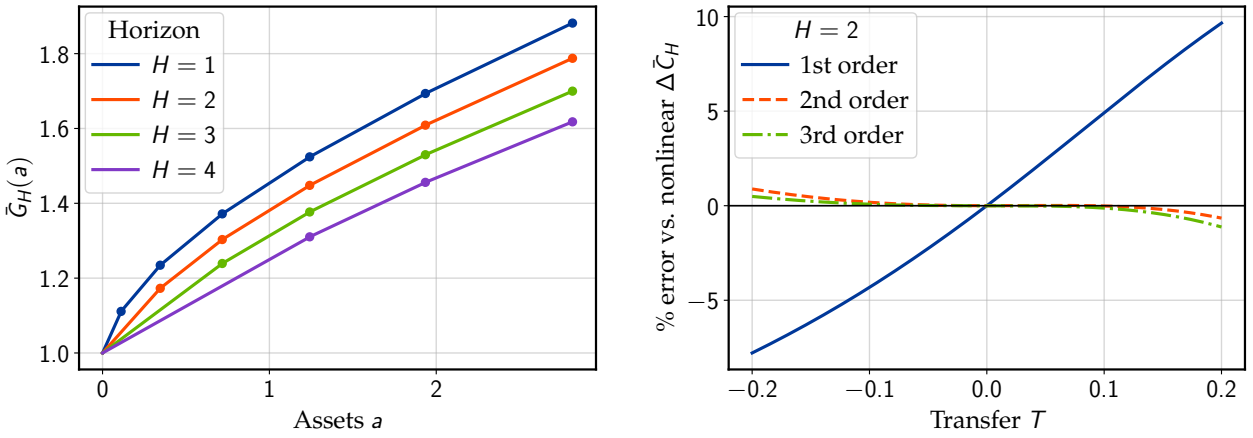
Average consumption over several periods. The preceding results concern current consumption. We can also ask how well the Taylor approximations perform for *cumulative* consumption over several periods. For a household starting with assets a, define cumulative and average consumption over the next H periods as

$$G_H(a) = \sum_{t=0}^{H-1} c_t(a), \qquad \bar{G}_H(a) = \frac{G_H(a)}{H}.$$

If the household reaches the constraint within the horizon, it spends all of its initial assets during these H periods. Cumulative consumption is then initial assets plus total income, so $G_H(a) = a + H$, regardless of exactly when the constraint is reached. If $n > H$, the household remains unconstrained throughout the horizon, so consumption continues to satisfy $c_t = \beta^t c_0$ and

$$G_H(a) = S_H \frac{a+n}{S_n}, \qquad a \in [a^{n-1}, a^n].$$

[A-21] Interestingly, the Taylor series around $T = 0$ has radius of convergence of only $a^1 = \beta^{-1} - 1 \simeq 0.11$. This is because at $T = a^1$, the first consumption kink reaches the lower endpoint of the asset distribution (0), where the lognormal density is smooth but not analytic. An infinite series therefore diverges for $|T| > a^1$, but the accuracy shown here illustrates that finite Taylor approximations can still work well beyond this radius when the relevant lower-order derivatives are sufficiently well-behaved.



(a) Individual averages by horizon

(b) Approximation errors for $H = 2$

Figure C.5: Consumption averaged over multiple periods

Hence, the segments on which the constraint is reached within the horizon all collapse to the same line, with slope $1/H$ for average consumption. The kinks at $a^1, \ldots, a^{H-1}$ disappear from $\bar{G}_H$, and its first remaining kink is at a^H. Figure C.5(a) shows how increasing the horizon successively removes these low-asset kinks.

Finally, define aggregate average consumption as

$$\bar{C}_H(T) = \int_0^\infty \bar{G}_H(a+T) f(a)\, da.$$

Figure C.5(b) repeats the Taylor-approximation exercise for $H = 2$, which is average consumption over the first two periods rather than one. The linear approximation is better but remains noticeably inaccurate; the second- and third-order errors, however, are now around 1% or less over nearly the entire range of transfers. Intuitively, as the kink points near the lognormal distribution's extreme curvature just above zero (figure C.1(b)) are removed, aggregate consumption becomes much better-behaved: although it is still nonlinear, it is regular enough to be well-captured by a second- or third-order Taylor expansion.[A-22] This helps explain why, in appendix D, we find that perturbation does even better for cumulative outcomes over a few periods.

Taking stock. Kinks in the individual consumption function, which emerge naturally in models with a borrowing constraint, do not by themselves rule out higher-order approximation: when households are smoothly distributed, aggregation produces a smooth consumption function whose derivatives capture households crossing the kinks. The problem is the combination of kinked policies and a distribution with mass points. As figures C.2 and C.3 show, a discretization can match aggregate consumption almost exactly in levels, while producing unreliable higher derivatives. By contrast, with a smooth distribution, second- and third-order approximations can capture substan-

[A-22] This improved approximation can be understood in part via the observation in footnote A-21: averaging over H periods expands the radius of convergence from a^1 to a^H.

tial nonlinearities, especially for cumulative outcomes over which the most important individual kinks disappear.

C.2 Smooth SIM model: solution and implementation

We now describe the smooth standard incomplete markets model used in the paper. The key addition to the usual model is a continuously distributed i.i.d. component of income. This component is realized and integrated out within the period, so it smooths the distribution of cash on hand without adding a state variable. It thereby eliminates the mass points in cash on hand that caused the difficulty in Appendix C.1, as well as the secondary kinks in the policy function: consumption is smooth everywhere except where the borrowing constraint starts to bind.[A-23]

Household problem and timing. Let $s = (e, \beta)$ denote the household's discrete state and let $\Pi_{ss'}$ be its transition matrix. At the beginning of period t, the household enters with assets a, observes s and the transitory shock ξ, and then chooses consumption c and assets a'. In the notation of equation (6), the Bellman equation is

$$\begin{aligned} V_t(a,s) &= \mathbb{E}_\xi \left[\max_{a' \geq \underline{a}} \left\{ u(c) + \beta \sum_{s'} \Pi_{ss'} \mathbb{E}_t V_{t+1}(a', s') \right\} \right], \\ c + a' &= (1 + r_t)a + Z_t e \xi + Tr_t. \end{aligned} \tag{A.40}$$

We retain time subscripts on r_t and Tr_t because both vary in some of the exercises in the paper. To be consistent with equation (7) in the main text, however, we suppress them as arguments of the household maps below. In the baseline calibration, $u(c) = \log c$, $\underline{a} = 0$, and

$$\log \xi \sim N(-\sigma_\xi^2/2, \sigma_\xi^2), \qquad \mathbb{E}\xi = 1.$$

We solve the household problem in three backward stages. This decomposition makes explicit the construction of the map $\mathcal{V}(Z_t, \mathbb{E}_t \boldsymbol{v}_{t+1})$ in the stacked household block in (7).

1. Continuation values. We iterate on the marginal value of incoming assets, $v_t(a,s) \equiv V_{a,t}(a,s)$. Let $w_t(a',s) \equiv \beta \mathbb{E}_t[V_{t+1}(a',s')|s]$ denote the discounted expected continuation value, so that the objective inside (A.40) is $u(c) + w_t(a',s)$. Its derivative is

$$w_{a,t}(a',s) = \beta \sum_{s'} \Pi_{ss'} \mathbb{E}_t v_{t+1}(a',s'). \tag{A.41}$$

2. Endogenous-grid solution conditional on cash on hand. Write cash on hand as

$$m = (1 + r_t)a + Z_t e \xi + Tr_t.$$

[A-23] Code and slides for the implementation of the model, although the notation differs, is available at lecture 9 of https://github.com/shade-econ/nber-workshop-2025.

Conditional on s and m, the household solves

$$\max_{a' \geq \underline{a}} \{u(m - a') + w_t(a', s)\}.$$

At each asset gridpoint a_i,[A-24] the interior first-order condition gives

$$c_{s,i} = (u')^{-1}(w_{a,t}(a_i, s)), \qquad m_{s,i} = a_i + c_{s,i}. \tag{A.42}$$

For our assumed CRRA utility with coefficient γ, this is $c_{s,i} = w_{a,t}(a_i, s)^{-1/\gamma}$. Thus the endogenous gridpoint method of Carroll (2006) directly supplies observations $(m_{s,i}, c_{s,i})$. We interpolate these observations with a cubic spline. If $m_s^* \equiv m_{s,0}$ is the first endogenous gridpoint, the complete policy is

$$c_t(m, s) = \begin{cases} m - \underline{a}, & m \leq m_s^*, \\ \widetilde{c}_t(m, s), & m > m_s^*, \end{cases} \tag{A.43}$$

where $\widetilde{c}_t$ is the spline. Hence, for each discrete state s, the consumption policy has a single kink at m_s^*, which is the point at which the borrowing constraint ceases to bind.

3. Integration over the transitory shock. The envelope condition returns the marginal value at the original asset gridpoints:

$$v_t(a, s) = (1 + r_t)\mathbb{E}_\xi \left[u'(c_t((1 + r_t)a + Z_t e\xi + Tr_t, s))\right]. \tag{A.44}$$

This expectation must be evaluated carefully because there is a kink at m_s^*. Let $d_t(a) = (1 + r_t)a + Tr_t$, $x = \log(Z_t e\xi)$, and $\mu_s = \log(Z_t e) - \sigma_\xi^2/2$. When $d_t(a) < m_s^*$, set $x_s^*(a) = \log(m_s^* - d_t(a))$. Then

$$\begin{aligned} \frac{v_t(a, s)}{1 + r_t} = &\int_{-\infty}^{x_s^*(a)} u'(d_t(a) + e^x - \underline{a})\, p_s(x)\, dx \\ &+ \int_{x_s^*(a)}^{\infty} u'(\widetilde{c}_t(d_t(a) + e^x, s))\, p_s(x)\, dx, \end{aligned} \tag{A.45}$$

where p_s is the density of a normal random variable with mean μ_s and standard deviation σ_ξ. If $d_t(a) \geq m_s^*$, only the second integral is needed. Numerically, we apply a fixed, high-order Gauss-Legendre rule separately to each nonempty interval. Splitting at the kink is important: a single quadrature rule whose nodes move across the kink is much less accurate and can lead to erratic numerical derivatives, similar to those we will see in the non-smooth model in appendix C.3.[A-25]

The combination of equations (A.41)–(A.45) defines the map $\mathcal{V}(Z_t, \mathbb{E}_t v_{t+1})$ in (7).

[A-24] Our implementation uses 200 double-exponentially spaced asset gridpoints on $[0, 1000]$. If u_i is uniformly spaced from zero to $\log[1 + \log(1 + 1000)]$, then $a_i = \exp(\exp(u_i) - 1) - 1$.

[A-25] Concretely, we truncate at $\mu_s \pm 8\sigma_\xi$ and use 40 Gauss–Legendre nodes on each nonempty side. In cases where we are far from the kink, we could also use Gauss-Hermite with a smaller number of nodes, but we stick to Gauss-Legendre for simplicity.

Forward iteration of the distribution. The smooth shock also permits a continuous-distribution implementation of the forward step. For each discrete state, we represent the conditional CDF of assets on the same asset grid used for the backward iteration. The CDF may have a mass point at $\underline{a}$, but it is smooth above the constraint. We interpolate the stored CDF values with a cubic spline whenever the CDF is evaluated between gridpoints. Thus the vector $\boldsymbol{\phi}_t$ in the main text may equivalently be viewed as CDF values at the knots or as coefficients of the cubic spline representation; the two are related by a fixed linear transformation.

To see the forward step, hold the current state s fixed and let $\Phi_{t-1,s}(a) \equiv \Pr(a_{t-1} \leq a \mid s_t = s)$ denote the CDF of incoming assets at t. The CDF of cash on hand is the convolution

$$\Phi^{\text{coh}}_{t,s}(m) = \int \Phi_{t-1,s}\left(\frac{m - Tr_t - Z_t e^{\xi}}{1 + r_t}\right) f_{\xi}(\xi)\, d\xi, \tag{A.46}$$

where $\Phi_{t-1,s}$ is understood to be zero below the borrowing limit. Next, denote by $\widehat{\Phi}_{t,s}(a_i)$ the post-saving CDF over assets conditional on the current state s. Since the saving policy is monotone and equation (A.42) maps $m_{s,i}$ into a_i, we can obtain $\widehat{\Phi}_{t,s}(a_i)$ without inversion or lottery:[A-26]

$$\widehat{\Phi}_{t,s}(a_i) = \Phi^{\text{coh}}_{t,s}(m_{s,i}). \tag{A.47}$$

In particular, $\widehat{\Phi}_{t,s}(\underline{a}) = \Phi^{\text{coh}}_{t,s}(m^*_s)$ records the mass at the constraint. Equation (A.46) is evaluated with the same quadrature rule in log transitory income.

The last forward stage applies the transition of the discrete state. If ϖ is the stationary distribution of s, defining the reverse conditional transition matrix as

$$\Pi^{\text{rev}}_{s's} = \Pr(s_t = s \mid s_{t+1} = s') = \frac{\varpi_s \Pi_{ss'}}{\varpi_{s'}}. \tag{A.48}$$

we have that the CDF carried into the next period, conditional on s', is

$$\Phi_{t,s'} = \sum_s \Pi^{\text{rev}}_{s's} \widehat{\Phi}_{t,s}. \tag{A.49}$$

The combination of equations (A.46)–(A.49) defines the map $\Lambda(Z_t, \mathbb{E}_t \boldsymbol{v}_{t+1}, \boldsymbol{\phi}_{t-1})$ in (7).[A-27]

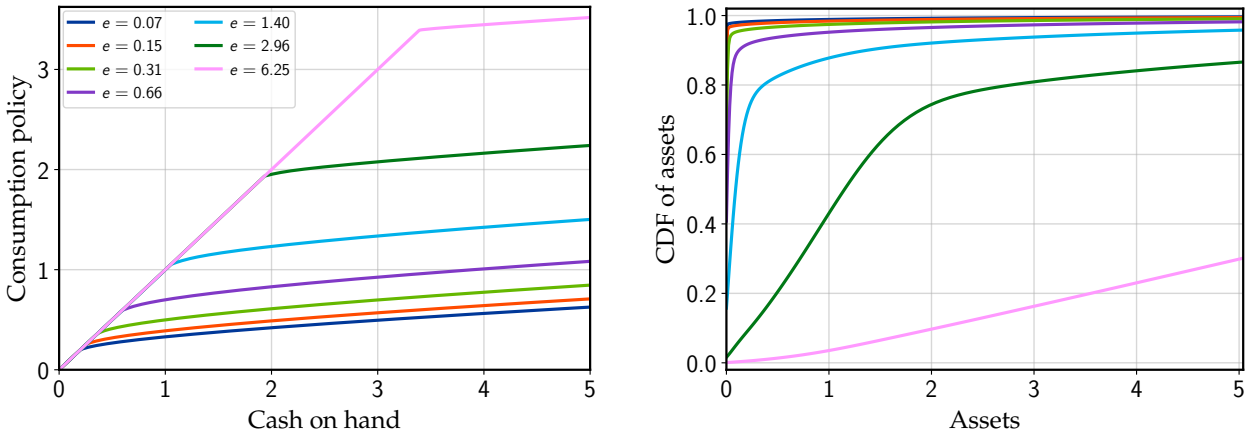
Aggregate consumption can be computed by integrating the policy over incoming assets, current discrete states, and ξ, as described in appendix A.5. In the implementation we use the equivalent aggregate budget identity

$$C_t = Z_t + (1 + r_t) A_{t-1} + Tr_t - A_t, \qquad A_t = \sum_s \varpi_s \int_{\underline{a}}^{\bar{a}} (1 - \Phi_{t,s}(a))\, da, \tag{A.50}$$

where $\bar{a}$ is the upper endpoint of the numerical asset grid and the baseline borrowing limit is $\underline{a} = 0$.

[A-26] This is also the key insight behind the distributional EGM method of Bayer et al. (2026b).

[A-27] To avoid instability in the spline near the boundary, our current implementation also smoothly forces the CDF to one near the upper end of the asset grid, which we choose so that it is far from much probability mass.



(a) Consumption policy

(b) Conditional asset CDFs

Figure C.6: Policies and conditional distributions in the smooth SIM model (low β type)

The survival functions are integrated as cubic splines. Equation (A.50) defines $\mathcal{C}(Z_t, \mathbb{E}_t \boldsymbol{v}_{t+1}, \boldsymbol{\phi}_{t-1})$, as well as the corresponding asset equation $\mathcal{A}(Z_t, \mathbb{E}_t \boldsymbol{v}_{t+1}, \boldsymbol{\phi}_{t-1})$. Although the individual policy retains a kink and the asset distribution an atom at the borrowing constraint, integration over the continuous ξ shock makes $\mathcal{V}$, Λ, and $\mathcal{C}$ smooth functions of their aggregate arguments.

All parameter values follow the calibration reported in appendix A.5. We discretize persistent income using the Rouwenhorst method (Kopecky and Suen, 2010) and compute the steady state by standard backward and forward iteration until convergence.[A-28]

Figure C.6 shows the two objects that result from the backward and forward iterations. To make the kink most visible, both panels condition on the lower discount-factor type. The consumption function has one kink for each persistent-income state and is smooth thereafter. The conditional asset CDFs have an atom at zero but no other jumps.

C.3 Smooth vs. non-smooth model

We next compare the smooth model with a conventional finite-state approximation and non-smooth implementation of an otherwise identical model.

Discrete non-smooth benchmark. We now approximate the i.i.d. shock ξ with a five-state, zero-persistence Rouwenhorst process, following the same construction used for the persistent component of income (Kopecky and Suen, 2010). The process is normalized to match the mean and level variance of the continuous lognormal.[A-29]

[A-28] The convergence criteria are a maximum change below 10^{-9} in the endogenous cash-on-hand grid and below 10^{-11} in the conditional CDFs.

[A-29] A five-point Gauss–Hermite approximation, treating its quadrature nodes as discrete i.i.d. shock states, produces the same qualitative results as in figure C.7: the first numerical derivative is stable, while the second and especially the third vary sharply with the finite-difference step size.

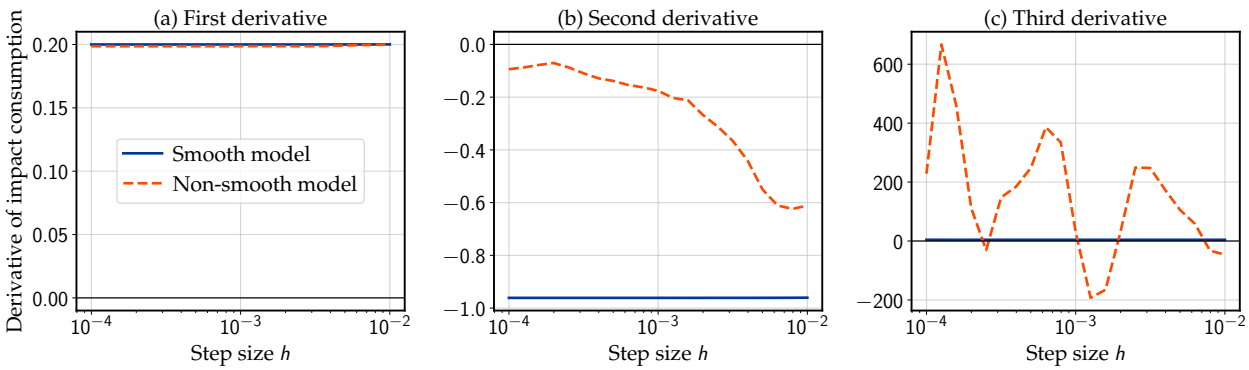


Figure C.7: Numerical derivatives of impact consumption with respect to a one-time income Z shock

We then solve the non-smooth model using the endogenous-grid method of Carroll (2006) with linear interpolation and the lottery method of Young (2010), following our typical approach (e.g., Auclert et al., 2021). Holding the smooth model's discount factors fixed gives $A = 4.01314$ and an impact income MPC of 0.19889, compared with targets of 4 and 0.2. These small gaps suggest that the two models are very close in levels and first-order behavior.[A-30]

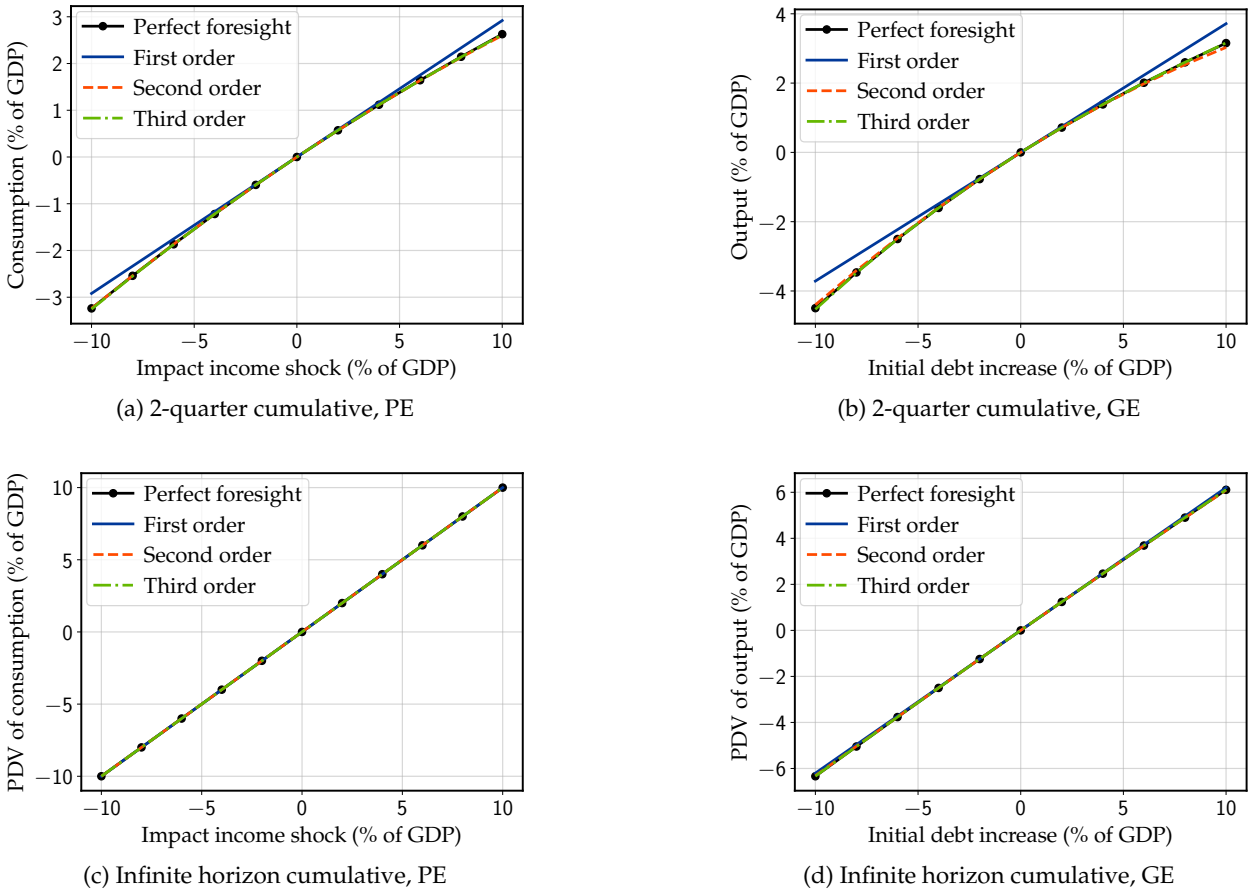
Numerical derivatives. The distinction becomes sharp at higher order. Let $C_0(z)$ denote impact consumption when $Z_0 = Z_{ss} + z$, with subsequent Z_t and transfers held at their baseline values. For each step size h, we estimate its first three derivatives at $z = 0$ using centered finite differences. Figure C.7 reports these estimates over $h \in [10^{-4}, 10^{-2}]$, holding the smooth discount factors fixed in both models. The first derivatives are stable and nearly identical. In the smooth model, both the second and third derivatives are stable around $h = 10^{-3}$, where their estimated values are approximately -0.963 and 4.11. In the non-smooth model, the second derivative varies substantially with h, while the third derivative is erratic and becomes very large at small step sizes.

The mechanism is the one described in Appendix C.1: mass points in the non-smooth distribution interact with kinks in the individual policy. The problem is more severe here because linear interpolation creates many closely spaced kinks throughout the policy function. As a result, the estimated third derivative is unstable and remains far from the smooth benchmark throughout the range of h shown in Figure C.7.

D Quantifying nonlinearities

Figures 1, 2 and B.4 show that third order very accurately captures the nonlinear solution in our model, for aggregate income shocks of up to $\pm 10\%$ of GDP, in both partial and general equilibrium. Bianchi and Kaplan (2026) instead argue that the third order approximation works poorly: in their words, "local methods of any order are unlikely to be reliable in settings where a fail-

[A-30] Recalibrating the non-smooth model to the same targets instead gives $(\beta_L^{\text{disc}}, \beta_H^{\text{disc}}) = (0.9597140, 0.9904922)$, changes of -0.0002504 and -0.0000026 relative to the smooth calibration.



(a) 2-quarter cumulative, PE

(b) 2-quarter cumulative, GE

(c) Infinite horizon cumulative, PE

(d) Infinite horizon cumulative, GE

Figure D.1: Cumulative impulse response of consumption

ure of Ricardian equivalence from high MPCs is important." Here, we show that a third order solution continues to work well for many variants of our model—all featuring lack of Ricardian equivalence from high MPCs—and also explore the types of situations in which it works less well.

Smooth vs unsmooth model. As discussed in appendix C, implementation of a third-order approximation via numerical differentiation requires a smooth underlying model, which Bianchi and Kaplan (2026) do not have. We therefore focus on our smooth model. Appendix C discusses at length the types of problems one encounters with perturbation when using a non-smooth model, so we do not discuss this issue further here.

Impact vs cumulative multipliers. As figures 2 and B.4 clearly illustrate, the size effect of income on consumption is concave on impact but convex later on, irrespective of the nature of the shock. As discussed in section 3.2, this is due to the intertemporal budget constraint, which holds nonlinearly. Hence, concavity is largest when we look at the impact effect: when we consider the cumulative effect on consumption from an income shock over longer horizons, we should expect the nonlinearity to progressively disappear.

Figure D.1 illustrates this pattern. Panel (a) shows the cumulative consumption response over the first two quarters, $\sum_{t=0}^{1}(C_t - C_{ss})$, following our baseline income shock in the PE model. The concavity and convexity in figure 2 partly offset, bringing the first-, second-, and third-order predictions much closer to the nonlinear perfect-foresight response. Panel (b) shows that this is also true in the GE model. Panel (c) shows the discounted cumulative response over the infinite horizon: in the PE model, we get all responses on a line with slope 1, as predicted by the intertemporal budget constraint. In the GE model, there is not such a clean theoretical benchmark, but the nonlinear responses are nevertheless all very close to linear, with slope equal to the cumulative transfer multiplier (which is around 0.6 in our model).

We consider the partial equilibrium model first, and then turn to the general equilibrium model, in which the same broad conclusions hold.

Lump-sum vs proportional transfers. Our baseline model has a shock to transfers that is distributed proportionately to household income (via Z in the PE model or τ in the GE model), as would be the case for changes in tax rates. In practice, some transfers are distributed more progressively, the classic case being that of lump-sum transfers. Figure D.2 considers the effect of a one-off lump-sum transfer, where the shock is to Tr_t rather than Z_t in equation 6. As the perfect-foresight lines show, nonlinearities are bigger for these transfers: there is both more concavity in the first quarter (left panel) and more convexity in the second quarter (middle panel), implying that large positive shocks have a lower impact and more persistent effect on consumption relative to their size. This type of nonlinearity is intuitive: very large lump-sum shocks have a highly disproportionate effect on households with low income and low assets. In fact, for agents with approximately zero productivity and zero incoming assets, a negative income shock that exceeds the steady state Tr is not even compatible with positive consumption, whereas a positive shock of that size would lead to large savings.

In conclusion, lump-sum transfers do create larger nonlinearities than proportional transfers, and these imply that the second and even the third order perturbation can fail to capture the nonlinearities correctly at very large shock sizes—for instance, the third order does poorly with shocks of +10% of quarterly GDP, although it does much better with shocks of -10% of quarterly GDP (recall that steady state transfers are 14% of quarterly GDP, so -10% is close to the maximum feasible reduction). Even then, these approximation errors partly offset when we sum consumption over the first two quarters (right panel of Figure D.2).

Transitory vs persistent transfers. So far, we have only considered fully transitory transfers, both proportional and lump-sum ($\rho_Z = \rho_{Tr} = 0$). In practice, however, large fiscal shocks are often persistent. Persistence, in turn, matters for the nonlinearity of the impact response of a transfer on consumption, because individuals at the borrowing constraint are unable to borrow against anticipated future transfers while those further away from the constraint can to some extent, making effective marginal propensities to consume out of transfer receipts less unequal.

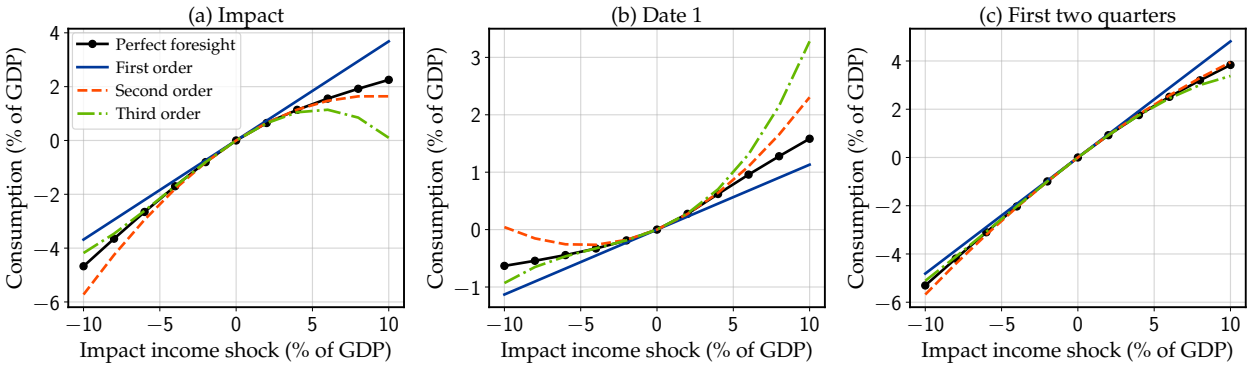


Figure D.2: Response to a one-time lump-sum transfer (PE model)

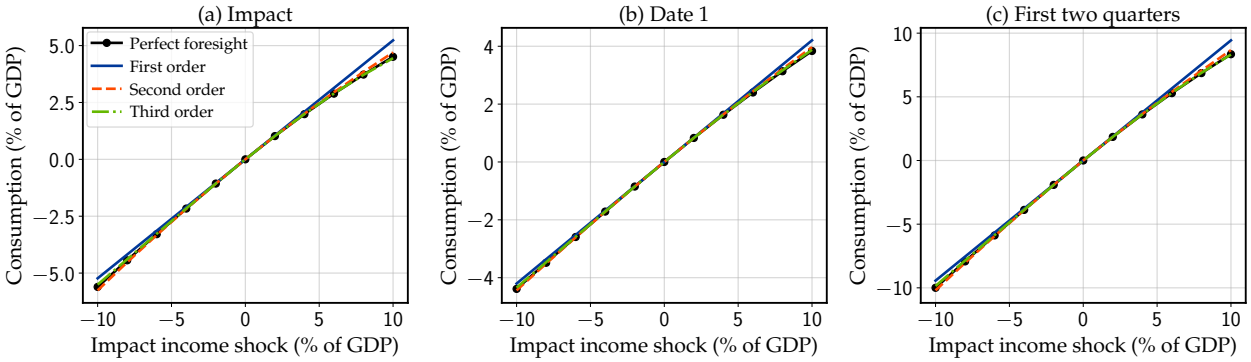


Figure D.3: Response to a lump-sum transfer with persistence 0.7 (PE model)

Indeed, figure D.3 shows that for aggregate, lump-sum income shocks of persistence $\rho_{Tr} = 0.7$, much of the nonlinearity in both the impact and the delayed effect on consumption goes away—in fact, with this amount of persistence, the $t = 1$ response is still mildly concave in size, with convexity appearing at longer horizons; the cumulative response over the first two quarters is already close to linear. Interestingly, even though higher persistence, conditional on the same impact transfer shock, makes the cumulative transfer much *larger*, the nonlinearity is nevertheless much *smaller*.

General equilibrium vs partial equilibrium. We have seen that, in our partial equilibrium model, nonlinearities are greatest when shocks are lump-sum, have no persistence, and their effect is considered over a single quarter. These results carry over to the general equilibrium version of the model, which features amplification from consumption to income across periods but does not fundamentally alter the nature of the nonlinearities in the effect of the fiscal shock itself.

Indeed, figures D.4 and D.5 reproduce D.2 and D.3 in the general equilibrium version of our model, with our background fiscal rule in place to bring bonds back to the steady state at rate $\rho_B = 0.95$. For persistent transfers, each additional transfer is financed by borrowing, while existing debt continues to be repaid at the same rate. While the amplification effect from the intertemporal Keynesian cross alters the slope of the first-order lines, similar nonlinearities as in the partial

equilibrium model exist in the perfect foresight solution, and the second- and third-order solutions do a similar job at capturing them: the third order solution performs poorly for very large one-off lump-sum transfers, but does better once we sum the effects over the first two quarters, consider a more persistent shock, or both.

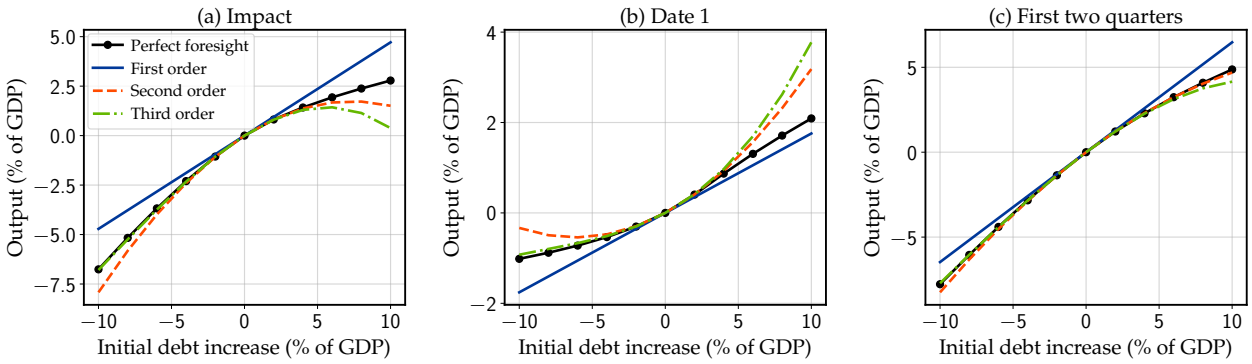


Figure D.4: Response to a one-time lump-sum transfer (GE model)

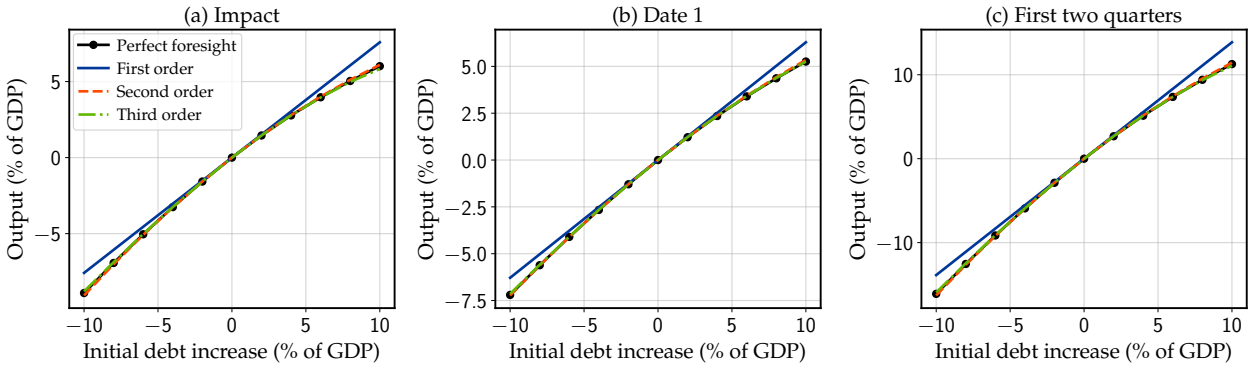
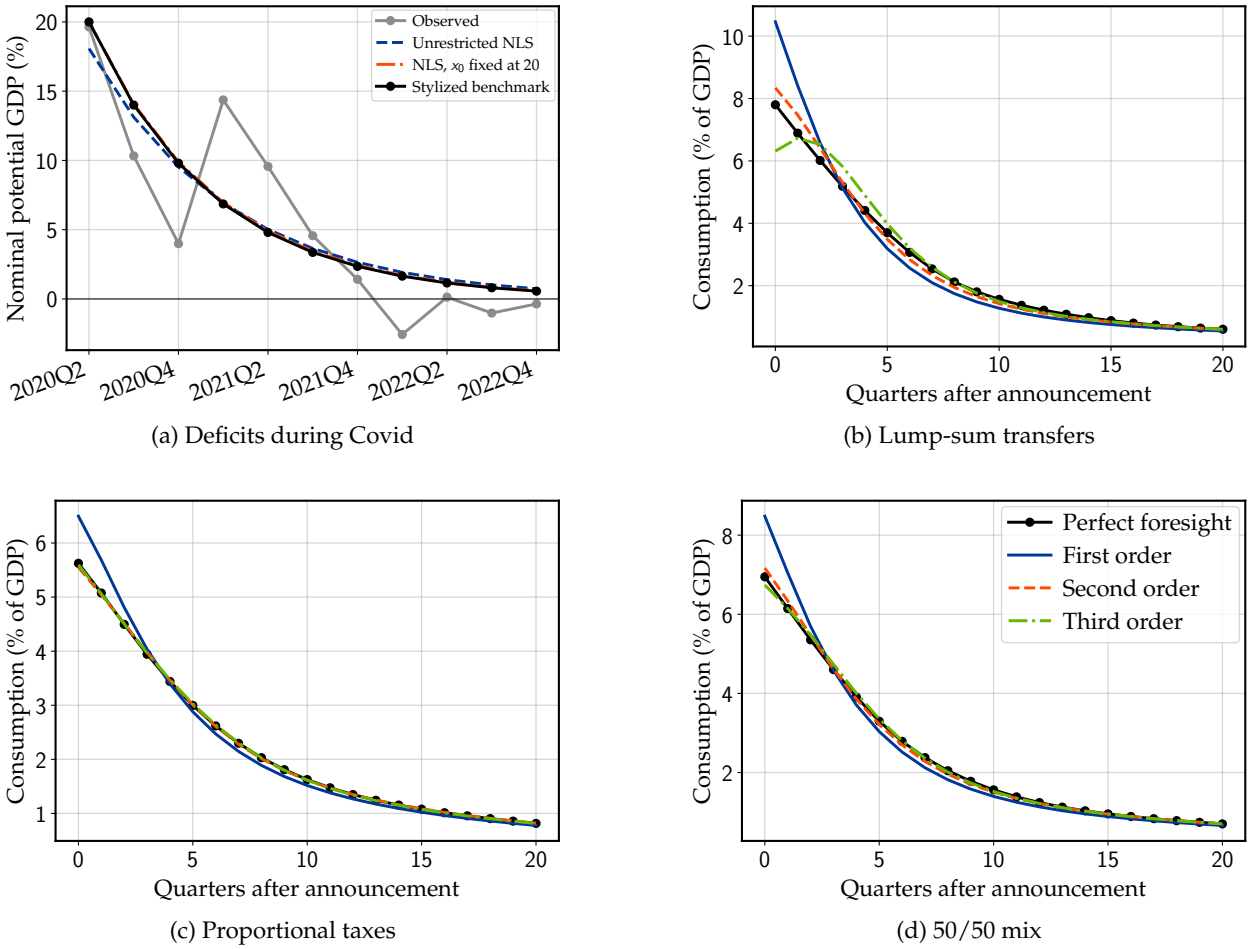


Figure D.5: Response to a lump-sum transfer with persistence 0.7 (GE model)

Covid-sized shock. Covid was a unique historical episode, with the largest-ever fiscal transfers seen in the United States. This provides a natural setting to judge the plausible magnitude and persistence of fiscal shocks.

The left panel of figure D.6 shows U.S. federal net borrowing over potential GDP, net of its pre-pandemic baseline.[A-31] The figure overlays an AR(1) shock with impact size 0.2 and persistence 0.7: this is very close to the best nonlinear least squares AR(1) fit (with impact 0.181 and persistence 0.726, plotted in the dashed blue line). Indeed, the quarterly fiscal deficit was very high in both 2020Q2 and 2021Q1; it only fell gradually back to its pre-Covid level over the following two years. A first takeaway from this exercise is that a fiscal shock persistence of $\rho = 0.7$, as in figures D.5 and D.3, is a natural approximation, and we have seen that this kind of persistence mitigates the nonlinearities in the impact effect of shocks.

[A-31] Federal net borrowing is from the BEA (FRED: AD02RC1Q027SBEA), potential GDP is from the CBO (FRED: NGDPPOT), and we take the baseline to be 2019Q4 value of the ratio.



(a) Deficits during Covid

(b) Lump-sum transfers

(c) Proportional taxes

(d) 50/50 mix

Figure D.6: Response to Covid shock

It is unclear whether we should model the effect of this shock as a lump-sum or a proportional transfer. Very far from all of the deficit went to pay for direct fiscal transfers to households, for which a lump sum would be appropriate: for instance, a significant fraction also financed transfers to firms to make payroll, which eventually benefited households in proportion to their income. A 50/50 mix of lump-sum and proportional transfers seems like a plausible assumption, though we also consider pure lump-sum and pure proportional transfers for completeness.

Panels b, c and d of figure D.6 show the nonlinear partial equilibrium consumption response to shocks with impact size 20% of quarterly GDP and persistence 0.7. For pure lump-sum transfers, even the third order solution has trouble matching the perfect-foresight solution on impact. However, third and even second-order appear quite accurate when we consider pure proportional transfers, or even a 50/50 mix of proportional and lump-sum transfers.

We view this good performance as a favorable result from a demanding stress test of perturbation. Given that the Covid pandemic was an exceptional, unanticipated event in modern US history, we would feel comfortable advising nonlinear MIT shocks rather than perturbation—whose main advantage is capturing how ongoing aggregate risk affects behavior—for a pandemic-sized

shock. But even at this unusual magnitude, third-order perturbation nevertheless acquits itself well for proportional transfers and for the plausible 50/50 mix.

Takeaway. We find that third order perturbation used correctly on a smooth model is likely to be reliable in most settings where a failure of Ricardian equivalence from high MPCs is important. In our experiments, significant errors in the third order approximation relative to the nonlinear perfect foresight solution only appear when most or all of the following features are present: 1) pure lump-sum rather than proportional (or a mix of proportional and lump-sum) transfers, 2) transitory rather than persistent transfers, 3) looking at the impact effect, rather than cumulative effect over a few quarters, and 4) shocks that are outside of a plausible range for business-cycle fluctuations, such as a Covid-sized shock.

E Appendix to section 4

E.1 Proof of the one-shot principle

Let $\boldsymbol{\epsilon} = (\epsilon_t, \epsilon_{t-1}, \ldots)$ and let $\mathbf{U}(\boldsymbol{\epsilon})$ denote the exact solution of (34). Set

$$A = \mathbf{H}_U(\mathbf{U}_{ss}, \mathbf{0}) \qquad \text{and} \qquad Q_{\boldsymbol{\epsilon}}(\mathbf{V}) = \mathbf{V} - A^{-1}\mathbf{H}(\mathbf{V}, \boldsymbol{\epsilon}).$$

Thus $\mathbf{U}(\boldsymbol{\epsilon})$ is a fixed point of $Q_{\boldsymbol{\epsilon}}$, while (35) sets $\mathbf{U}^{[k]} = Q_{\boldsymbol{\epsilon}}(\mathbf{U}^{[k-1]})$, starting from $\mathbf{U}^{[0]} = \mathbf{U}_{ss}$.

At the deterministic steady state,

$$\frac{\partial Q_{\mathbf{0}}}{\partial \mathbf{V}}(\mathbf{U}_{ss}) = I - A^{-1}A = 0.$$

Smoothness therefore implies, for $\boldsymbol{\epsilon}$ sufficiently small and for $\mathbf{V}$ within order $\|\boldsymbol{\epsilon}\|$ of the steady state,

$$\left\|Q_{\boldsymbol{\epsilon}}(\mathbf{V}) - Q_{\boldsymbol{\epsilon}}(\mathbf{U}(\boldsymbol{\epsilon}))\right\| \leq C\|\boldsymbol{\epsilon}\|\left\|\mathbf{V} - \mathbf{U}(\boldsymbol{\epsilon})\right\|. \tag{A.51}$$

The implicit function theorem gives $\mathbf{U}(\boldsymbol{\epsilon}) - \mathbf{U}_{ss} = O(\|\boldsymbol{\epsilon}\|)$. Hence the initial error is of first order, and inequality (A.51) raises its order by one at every iteration. Induction gives

$$\mathbf{U}^{[k]} - \mathbf{U}(\boldsymbol{\epsilon}) = O(\|\boldsymbol{\epsilon}\|^{k+1}). \tag{A.52}$$

The second line of the recursion in proposition 7 defines the approximate outcome as

$$X_t^{[k]} = \widetilde{\mathcal{X}}\left(\mathbf{U}^{[k]}, \boldsymbol{\epsilon}\right),$$

while the exact perfect-foresight outcome is

$$\mathcal{X}(0, \boldsymbol{\epsilon}) = \widetilde{\mathcal{X}}\left(\mathbf{U}(\boldsymbol{\epsilon}), \boldsymbol{\epsilon}\right).$$

Since $\widetilde{\mathcal{X}}$ is smooth, its change between these two evaluations is bounded by a constant times the change in its first argument. Equation (A.52) therefore gives

$$\left\| X_t^{[k]} - \mathcal{X}(0, \boldsymbol{\epsilon}) \right\| \leq C \left\| \mathbf{U}^{[k]} - \mathbf{U}(\boldsymbol{\epsilon}) \right\| = O(\|\boldsymbol{\epsilon}\|^{k+1}) = o(\|\boldsymbol{\epsilon}\|^k),$$

which proves the proposition. The same argument, with the expectational shifter in place of $\boldsymbol{\epsilon}$, proves the steady-state version.

E.2 Practical implementation details

The MIT-shock routine $\tilde{X}_j^{MIT}(\nu, X_-, \mathbf{U})$ takes as its initial condition the state X_- at date -1, immediately before the date-zero shock, and returns the outcome at date $j \geq 0$.

Equations (36)–(38) follow by applying this routine successively. Normalize the date of the newest shock ν to zero. If the earlier shock ϵ arrives at date $-l$, the state immediately before the newest shock is $X_{l-1}^{MIT}(\epsilon)$, giving (37). With a third shock ϵ' at date $-l'$, where $0 < l < l'$, the state immediately before the middle shock is $X_{l'-l-1}^{MIT}(\epsilon')$. Starting the middle-shock transition from this state and evaluating it $l-1$ periods later gives the state $X_{l-1}^{MIT,l'-l}(\epsilon, \epsilon')$ immediately before the newest shock, as in (38). Each shock is unanticipated when it arrives, and no subsequent shock is anticipated in any of these transitions. Uniqueness of the perfect-foresight continuation therefore ensures that the recursively constructed paths coincide with the shock responses defined in section 3.

By (A.52), a derivative of total order q can be calculated after replacing each nonlinear equilibrium solution $\mathbf{U}$ by $\mathbf{U}^{[q]}$. This replacement applies both to the final transition and to the earlier transitions used to construct its inherited state. Smooth composition preserves agreement through order q.

Equation (39) gives the second-order formula for a single shock. Its third-order analogue is

$$\mathcal{X}_{jjj} = \frac{d^3}{d\nu^3} \tilde{X}_j^{MIT} \left(\nu, X, \mathbf{U}^{[3]}(\nu, X) \right) \Big|_{\nu=0}. \tag{A.53}$$

To write the remaining expressions compactly, for $h \geq 1$ define

$$S_h^{[q]}(\epsilon) \equiv \tilde{X}_{h-1}^{MIT} \left(\epsilon, X, \mathbf{U}^{[q]}(\epsilon, X) \right). \tag{A.54}$$

Thus $S_h^{[q]}(\epsilon)$ approximates the state immediately before a new shock arriving h periods after ϵ. In particular, when $h = 1$, the inherited state is the impact outcome of the earlier shock.

For $l > 0$, the second-order coefficient from two shocks is

$$\mathcal{X}_{j,j+l} = \frac{\partial^2}{\partial \nu \, \partial \epsilon} \tilde{X}_j^{MIT} \left(\nu, S_l^{[2]}(\epsilon), \mathbf{U}^{[2]} \left(\nu, S_l^{[2]}(\epsilon) \right) \right) \Big|_{\nu=\epsilon=0}. \tag{A.55}$$

The two third-order coefficients from the same experiment are

$$\mathcal{X}_{j,j,j+l} = \frac{\partial^3}{\partial \nu^2\, \partial \epsilon} \tilde{X}_j^{MIT}\left(\nu, S_l^{[3]}(\epsilon), \mathbf{U}^{[3]}\left(\nu, S_l^{[3]}(\epsilon)\right)\right)\Big|_{\nu=\epsilon=0}, \tag{A.56}$$

$$\mathcal{X}_{j,j+l,j+l} = \frac{\partial^3}{\partial \nu\, \partial \epsilon^2} \tilde{X}_j^{MIT}\left(\nu, S_l^{[3]}(\epsilon), \mathbf{U}^{[3]}\left(\nu, S_l^{[3]}(\epsilon)\right)\right)\Big|_{\nu=\epsilon=0}. \tag{A.57}$$

Finally, for $0 < l < l'$, define the state inherited from two earlier shocks by

$$S_{l,l'}^{[3]}(\epsilon, \epsilon') \equiv \tilde{X}_{l-1}^{MIT}\left(\epsilon, S_{l'-l}^{[3]}(\epsilon'), \mathbf{U}^{[3]}\left(\epsilon, S_{l'-l}^{[3]}(\epsilon')\right)\right). \tag{A.58}$$

Here $S_{l'-l}^{[3]}(\epsilon')$ is the state immediately before the middle shock at date $-l$. The outer routine returns the outcome $l-1$ periods after that shock, at date -1, immediately before the newest shock. Hence $S_{l,l'}^{[3]}(\epsilon, \epsilon')$ approximates $X_{l-1}^{MIT,l'-l}(\epsilon, \epsilon')$.

The third-order coefficient from three shocks is therefore

$$\mathcal{X}_{j,j+l,j+l'} = \frac{\partial^3}{\partial \nu\, \partial \epsilon\, \partial \epsilon'} \tilde{X}_j^{MIT}\left(\nu, S_{l,l'}^{[3]}(\epsilon, \epsilon'), \mathbf{U}^{[3]}\left(\nu, S_{l,l'}^{[3]}(\epsilon, \epsilon')\right)\right)\Big|_{\nu=\epsilon=\epsilon'=0}. \tag{A.59}$$

All expressions apply for $j \geq 0$. Each equilibrium solve uses two fixed-Jacobian iterations for a second-order derivative and three for a third-order derivative, so no nonlinear fixed point is required.

For the risk-related terms, write $U(\delta)$ for the steady-state unknowns when expectations are shifted by $\mathcal{X}_{00}\delta$. They solve

$$H\big(U(\delta), \mathcal{X}_{00}\delta\big) = 0, \qquad U(0) = U_{ss}.$$

Differentiating at zero gives

$$H_U U'(0) + \frac{\partial H(U_{ss}, \mathcal{X}_{00}\delta)}{\partial \delta}\Big|_{\delta=0} = 0.$$

Consequently,

$$U'(0) = -H_U^{-1}\, \frac{\partial H(U_{ss}, \mathcal{X}_{00}\delta)}{\partial \delta}\Big|_{\delta=0},$$

which is (40) when $U_{\sigma\sigma}$ denotes $U'(0)$. Substituting this derivative into the reconstruction of the steady state gives $\mathcal{X}_{\sigma\sigma}$. Equivalently, because this is a first derivative with respect to δ, it is obtained by one fixed-Jacobian iteration.

The mixed risk term can be constructed just as explicitly. Let

$$\tilde{X}_j^{pt}(\{\Delta_i\}, \nu, X^b, \mathbf{U})$$

denote the reconstruction of the date-j outcome of (25), given the expectation shifts, the date-zero

shock, the common initial and terminal state X^b, and the unknown path $\mathbf{U}$. Let $\mathbf{U}^{pt,[q]}(\{\Delta_i\}, \nu, X^b)$ denote q fixed-Jacobian iterations for the corresponding reduced equilibrium system, starting from the steady-state path. The second-order risky-steady-state result implies

$$X^{pss}(\mathcal{X}_{00}\delta) = X + \mathcal{X}_{\sigma\sigma}\delta + o(\delta).$$

Only this first-order change in the boundary state affects the mixed derivative below.

For the risky-steady-state component, the one-shot construction is

$$\left(\mathcal{X}_{\sigma\sigma j}\right)^{rss} = \frac{\partial^2}{\partial\delta\,\partial\nu}\tilde{X}_j^{pt}\left(\{\mathcal{X}_{00}\delta, \mathcal{X}_{00}\delta, \ldots\}, \nu, X + \mathcal{X}_{\sigma\sigma}\delta, \right. \\ \left. \mathbf{U}^{pt,[2]}\left(\{\mathcal{X}_{00}\delta, \mathcal{X}_{00}\delta, \ldots\}, \nu, X + \mathcal{X}_{\sigma\sigma}\delta\right)\right)\Big|_{\delta=\nu=0}. \quad \text{(A.60)}$$

This uses two iterations because the derivative is of total order two.

For the time-varying precautionary component, set

$$\Delta_i^{tvp}(\delta) = \mathcal{X}_{00,i+1}\delta, \qquad i \geq 0.$$

There is no realized shock and both boundary conditions remain at X, so the corresponding construction is

$$\left(\mathcal{X}_{\sigma\sigma j}\right)^{tvp} = \frac{d}{d\delta}\tilde{X}_j^{pt}\left(\{\Delta_i^{tvp}(\delta)\}, 0, X, \mathbf{U}^{pt,[1]}\left(\{\Delta_i^{tvp}(\delta)\}, 0, X\right)\right)\Big|_{\delta=0}. \quad \text{(A.61)}$$

Only one iteration is needed because this is a first derivative. Adding (A.60) and (A.61) gives $\mathcal{X}_{\sigma\sigma j}$ by (24), without solving a nonlinear fixed point.

E.3 Choice of truncation details

We now describe the calculations underlying figure 7 in more detail. Throughout, we assume $\bar{\epsilon}_t$ has the same moments (up to 6th) as a standard normal and set $\sigma = 0.05$. We normalize most measures of error by the standard deviation $\sigma\sqrt{\Sigma_{0,2}}$ implied by the first-order solution, where $\Sigma_{0,2} \equiv \sum_j \mathcal{X}_j^2$ as in proposition 6.

We adopt two notational conventions. First, $L_{\text{low}} < L_{\text{high}}$ denote the two history lengths being compared ($L_{\text{high}} = 300$ for the benchmark comparisons in (a), and $L_{\text{high}} = L_{\text{low}} + 20$ for the successive Ls used in extrapolation in (b) and (c)). Second, for convenience we use $\mathcal{X}_j$, $\mathcal{X}_{ij}$, etc. to denote scalar coefficients for the outcome of interest (here, output), rather than the full vector $\mathcal{X}$.

Second order. Define $\mathcal{X}_{ij}(L)$ to equal $\mathcal{X}_{ij}$ when $\min(i,j) < T$ and $|i-j| < L$, and zero otherwise, and then $\Delta\mathcal{X}_{ij} \equiv \mathcal{X}_{ij}(L_{\text{high}}) - \mathcal{X}_{ij}(L_{\text{low}})$. The difference between the two truncated approximations is then $\Delta X_t^{(2)} = \frac{1}{2}\sum_{i,j}\Delta\mathcal{X}_{ij}\epsilon_{t-i}\epsilon_{t-j}$. Using the Gaussian moments and noting that $\Delta\mathcal{X}_{ii} = 0$, its exact root mean square is $\sigma^2[\frac{1}{2}\sum_{i,j}(\Delta\mathcal{X}_{ij})^2]^{1/2}$. Normalizing by $\sigma\sqrt{\Sigma_{0,2}}$, we report $\sigma[\frac{1}{2}\sum_{i,j}(\Delta\mathcal{X}_{ij})^2]^{1/2}/\sqrt{\Sigma_{0,2}}$. The maximum-coefficient diagnostic similarly reports $\sigma\max_{i,j}|\Delta\mathcal{X}_{ij}|/\sqrt{\Sigma_{0,2}}$,

which turns out to decay at a similar rate.

For the second-order moment lines, let $\Sigma^{(2)}_{h,4}(L)$ denote the first term in $\Sigma_{h,4}$ in Proposition 6, evaluated at history length L, and define $\Delta\Sigma^{(2)}_{h,4} = \Sigma^{(2)}_{h,4}(L_{\text{high}}) - \Sigma^{(2)}_{h,4}(L_{\text{low}})$. At $h = 0$, the resulting difference in variance is $\sigma^4|\Delta\Sigma^{(2)}_{0,4}|$. We convert this locally into a difference in standard deviations using $d\sqrt{V} = dV/(2\sqrt{V})$ around the first-order approximation $V = \sigma^2\Sigma_{0,2}$. This gives $\sigma^3|\Delta\Sigma^{(2)}_{0,4}|/(2\sqrt{\Sigma_{0,2}})$, which after normalizing by the first-order standard deviation becomes the reported error $\sigma^2|\Delta\Sigma^{(2)}_{0,4}|/(2\Sigma_{0,2})$. At $h = 5$, the difference in autocovariances is $\sigma^4|\Delta\Sigma^{(2)}_{5,4}|$; dividing directly by the first-order variance gives the reported error $\sigma^2|\Delta\Sigma^{(2)}_{5,4}|/\Sigma_{0,2}$.

Third order. We continue to hold $T = 300$ fixed and also hold the second-order solution fixed at $L = 100$. Define $\mathcal{X}_{ijk}(L)$ to equal $\mathcal{X}_{ijk}$ when $\min(i,j,k) < T$ and $\max(i,j,k) - \min(i,j,k) < L$, and zero otherwise, and let $\Delta\mathcal{X}_{ijk} \equiv \mathcal{X}_{ijk}(L_{\text{high}}) - \mathcal{X}_{ijk}(L_{\text{low}})$. For $\mathcal{X}_{\sigma\sigma i}$, only the time-varying precautionary component depends on L; for each L, we recompute it using the third-order histories truncated at L, while holding the risky-steady-state component fixed, and then define $\Delta\mathcal{X}_{\sigma\sigma i} \equiv \mathcal{X}_{\sigma\sigma i}(L_{\text{high}}) - \mathcal{X}_{\sigma\sigma i}(L_{\text{low}})$. The difference between the two third-order approximations is then $\Delta X^{(3)}_t = \frac{1}{6}\sum_{i,j,k}\Delta\mathcal{X}_{ijk}\epsilon_{t-i}\epsilon_{t-j}\epsilon_{t-k} + \frac{1}{2}\sigma^2\sum_i \Delta\mathcal{X}_{\sigma\sigma i}\epsilon_{t-i}$. Using the assumption that $\bar{\epsilon}$ has Gaussian moments, its mean square can be shown to be

$$\mathbb{E}[(\Delta X^{(3)}_t)^2] = \sigma^6\left\{\frac{1}{6}\sum_{i,j,k}(\Delta\mathcal{X}_{ijk})^2 + \frac{1}{4}\sum_i\left(\Delta\mathcal{X}_{\sigma\sigma i} + \sum_j \Delta\mathcal{X}_{ijj}\right)^2\right\}. \tag{A.62}$$

To illustrate (A.62), note that after squaring the cubic part of $\Delta X^{(3)}_t$ and expanding, a product has nonzero expectation only if every dated shock appears an even number of times. One contribution matches the three shock factors in one copy with the three factors in the other; there are $3! = 6$ such matchings. Each contributes a product of three variances, σ^6, so symmetry gives $6\sigma^6\sum_{i,j,k}(\Delta\mathcal{X}_{ijk})^2$; multiplication by $(1/6)^2$ gives the first term on the right above. The other nonzero terms in the expansion, involving repeated shock dates or $\Delta\mathcal{X}_{\sigma\sigma i}$, produce the second term on the right.[A-32]

After taking the square root of (A.62), we normalize by $\sigma\sqrt{\Sigma_{0,2}}$ as at second order. We also calculate the moment errors as at second order, now using the change between L_{low} and L_{high} in the final two terms of $\Sigma_{h,4}$ in proposition 6, which involve $\mathcal{X}_{\sigma\sigma i}$ and $\mathcal{X}_{ijj}$. As noted in the main text, these terms are linear in the third-order coefficients rather than squared, explaining why they decay more slowly with L in figure 7(c) than the second-order moment errors in (b).

Extrapolation. For each diagnostic, let q_0 denote its value when comparing history lengths $L-40$ and $L-20$, and let q_1 denote its value when comparing $L-20$ and L. Setting $r = q_1/q_0$, we assume, when $0 < r < 1$, that further 20-lag increases contribute $q_1 r, q_1 r^2, \ldots$. For second-order root mean square, terms from disjoint ranges of $|i-j|$ are uncorrelated, so the extrapolated to-

[A-32] We have also validated (A.62) using Monte Carlo simulation; since (A.62) is exact and lacks Monte Carlo noise, it is preferable for practical truncation diagnostics.

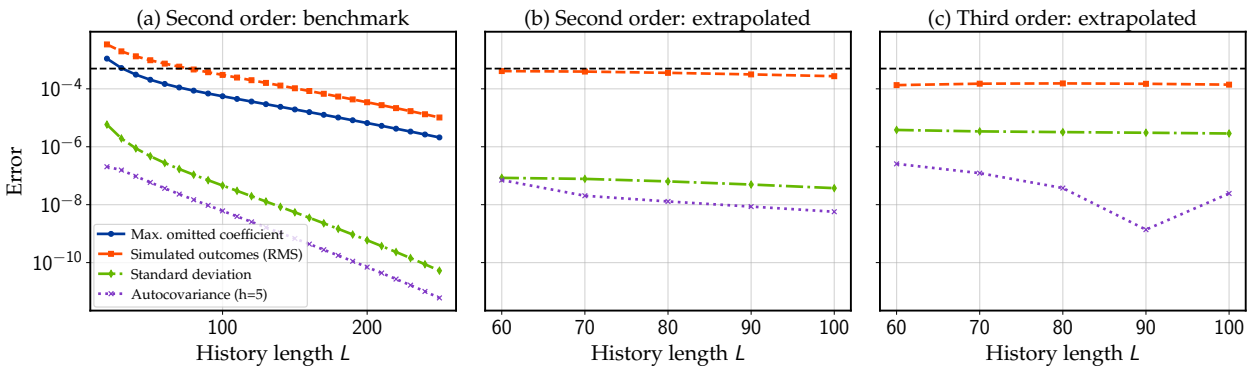


Figure E.1: Truncation error in the smooth SIM (PE) solution by history length L

tal error at L relative to the infinite-history expansion is $q_1 r/\sqrt{1-r^2}$. For moment errors and all third-order diagnostics, we sum absolute projected changes instead, giving a conservative estimate $q_1 r/(1-r)$ of this total error.

Partial equilibrium. Figure E.1 applies identical diagnostics to consumption in the PE smooth SIM model, giving the PE counterpart of figure 7. At $L = 100$, every diagnostic is below our 5×10^{-4} target; the largest is the direct second-order root mean squared error, approximately 3.0×10^{-4}. This supports the same choice of $L = 100$ as in the GE IKC model.

E.4 Numerical differentiation details

In our implementation, we first apply the finite-difference formulas to the computationally costly household transition, then differentiate the aggregate equilibrium conditions and solve for the associated general-equilibrium response using the one-shot procedure in appendices E.1 and E.2. The main text gives the centered formulas for first- and second-order derivatives in equations (41) and (42). For third-order size dependence, we use the seven-point centered stencil

$$\mathcal{X}_{jjj} \simeq \frac{X_j^{MIT}(-3h) - 8X_j^{MIT}(-2h) + 13X_j^{MIT}(-h) - 13X_j^{MIT}(h) + 8X_j^{MIT}(2h) - X_j^{MIT}(3h)}{8h^3} \tag{A.63}$$

Assuming sufficient differentiability, this has error of order h^4. We use this wider stencil—as opposed to the simpler five-point stencil—because third-order numerical differentiation is demanding numerically, and a more accurate formula is necessary to achieve sufficiently low error.

For shocks at different dates, let l and l' be the lags of the middle and oldest shocks relative to the newest shock, with $0 < l < l'$. The relevant polarization identity can be written as

$$\mathcal{X}_{j,j+l,j+l'} = \frac{1}{6}\frac{d^3}{d\eta^3}\bigg|_{\eta=0} \big[X_j^{MIT,l,l'}(\eta,\eta,\eta) - X_j^{MIT,l}(\eta,\eta) - X_j^{MIT,l'}(\eta,\eta) - X_{j+l}^{MIT,l'-l}(\eta,\eta) + X_j^{MIT}(\eta) + X_{j+l}^{MIT}(\eta) + X_{j+l'}^{MIT}(\eta) - X\big].$$

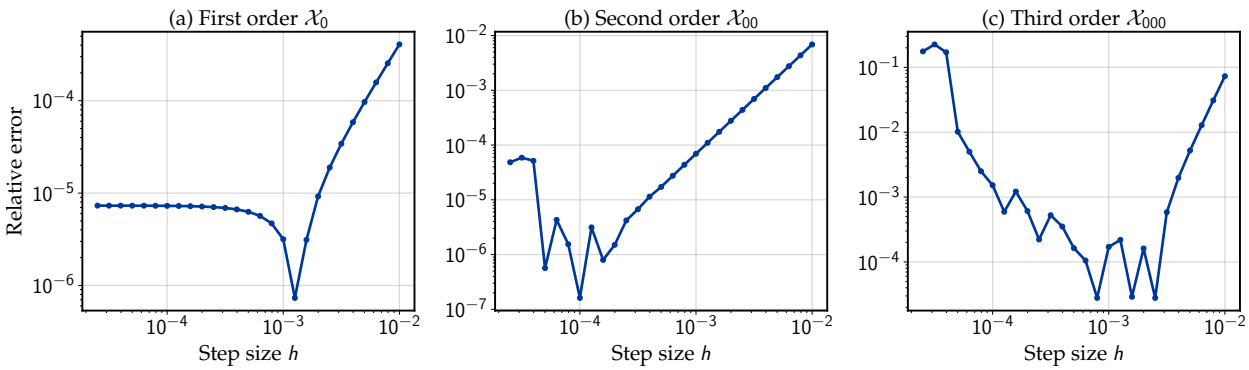


Figure E.2: Relative discrepancies between numerical and analytical impact output derivatives (IKC model)

We evaluate the derivative of the bracket using the seven-point stencil above. The displayed formula assumes three distinct shock dates. When the newest two indices coincide, its three-shock term is instead evaluated as $X_j^{MIT,l'}(2\eta, \eta)$. When the oldest two indices coincide, the corresponding term is $X_j^{MIT,l}(\eta, 2\eta)$. (When all three coincide, we simply have a size effect.) To limit repeated nonlinear household transition computations, we reuse each one-shock transition wherever it appears, and each two-shock transition across coefficient calculations with the same separation between shocks.

Figure E.2 plots the relative discrepancies between the numerical and analytical impact output coefficients, displayed in figure 8, on a logarithmic scale. As our defaults throughout the paper, we choose points roughly near the middle of the stable region for numerical differentiation: $h_1 = 10^{-4}$ for first order, $h_2 = 2 \times 10^{-4}$ for second order, and $h_3 = 10^{-3}$ for third order. The figure confirms that numerical differentiation is accurate relative to the analytical benchmark in these regions—albeit with somewhat higher error, as usual for numerical differentiation, for second and third vs. first order.

E.5 Analytical differentiation details

- Use the construction of the functions $\mathcal{V}, *$ in C.2 to give details of what we say in text for first and second order

E.6 Exponential mixture approximation and extended timings

- Fitting the first-order exponential
- Give the second-order recursions
- Explain the third order
- Add extended speed table with exponential + quantitative model

Model	Object to compute	Type of differentiation		
		Numerical	**Analytical**	**Exponential**
Pre-computation *(needed for PE & GE)*	Steady state	4.0 s	4.0 s	4.0 s
	Other setup	0.0 s	1.2 s	1.2 s
	First-order solution	40.4 s	1.5 s	1.5 s
SIM model (PE)	RSS and ergodic mean	16.0 s	0.5 s	0.5 s
	Second-order solution	4.7 m	10.6 s	0.5 s
	Moments to fifth order	12.7 m	14.7 s	0.9 s
	Third-order solution	3.7 h	4.5 m	1.7 s
IKC model (GE)	RSS and ergodic mean	15.3 s	0.4 s	1.2 s
	Second-order solution	4.7 m	10.7 s	1.3 s
	Moments to fifth order	15.4 m	16.8 s	17.3 s
	Third-order solution	3.9 h	4.4 m	18.5 s
Quantitative model	RSS and ergodic mean	36.7 s	1.1 s	4.0 s
	Second-order solution	41.4 m	95.8 s	5.7 s
	Moments to fifth order	2.7 h	3.5 m	1.3 m
	Third-order solution	103.1 h	2.0 h	3.4 m

Note: Times are cumulative within each block.

Table E.1: Computation speed for sequence space solution: extended version

E.7 Sequence space vs. state space: the N country neoclassical growth model

In this section, we consider an N country extension of the Neoclassical Growth Model in Example 1. Each country is indexed by $i = 1, \ldots, N$, and has its own risk aversion parameter γ_i , capital share α_i, loading ζ_i on the world's aggregate TFP, and initial level of capital. Goods and capital are freely mobile, and markets are complete. The planner's problem, given country weights $\omega_i > 0$, is to solve:

$$\max_{\{C_{it}, K_{it}\}} \mathbb{E} \sum_{t=0}^{\infty} \beta^t \sum_{i=1}^{N} \omega_i \frac{C_{it}^{1-\gamma_i}}{1-\gamma_i} \tag{A.64}$$

subject to technology

$$Y_{it} = e^{\zeta_i a_t} K_{i,t-1}^{\alpha_i} \tag{A.65}$$

and the world resource constraint

$$\sum_{i=1}^{N} (C_{it} + K_{it} - (1-\delta_i) K_{i,t-1} - Y_{it}) = 0 \tag{A.66}$$

The process for aggregate log TFP is still given by $a_t = \rho_a a_{t-1} + \epsilon_t$, with each country's own log TFP being $\zeta_i a_t$.

The optimum in this model is characterized by the risk-sharing conditions

$$\omega_i C_{it}^{-\gamma_i} = \mu_t, \qquad i = 1, \ldots, N. \tag{A.67}$$

The capital first-order condition is

$$\mu_t = \beta \mathbb{E}_t \left[\left(1 - \delta_i + \alpha_i \exp(\zeta_i a_{t+1}) K_{it}^{\alpha_i - 1} \right) \mu_{t+1} \right] \qquad i = 1, \ldots, N \tag{A.68}$$

so that the relevant risk-adjusted return is equalized across all countries. To put this model in the notation of (1), we define as in the one-country model the country-specific variable $\lambda_{it} \equiv \left(1 - \delta_i + \alpha_i \exp(\zeta_i a_t) K_{i,t-1}^{\alpha_i - 1} \right) \mu_t$. Then, we stack $X_t = (a_t, \mu_t, \{\lambda_{it}\}, \{K_{it}\}, \{C_{it}\})$, and define

$$F\left(X_{t-1}, X_t, \mathbb{E}_t\left[X_{t+1}\right], \epsilon_t\right) = \begin{bmatrix} \lambda_{it} - \left(1 - \delta_i + \alpha_i \exp(\zeta_i a_t) K_{i,t-1}^{\alpha_i - 1} \right) \mu_t \\ \mu_t - \omega_i C_{it}^{-\gamma_i} \\ \mu_t - \beta \mathbb{E}_t \left[\lambda_{i,t+1} \right] \\ \sum_{i=1}^N \left(C_{it} + K_{it} - (1 - \delta_i) K_{i,t-1} - Y_{it} \right) \\ a_t - \rho_a a_{t-1} - \epsilon_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{A.69}$$

The model therefore has dimension $n = 3N + 2$.

Deterministic and perturbed-expectations steady-state. At the deterministic steady state with $\epsilon = 0$, world TFP is $a = 0$. The capital first-order condition (A.68) gives $1 - \delta_i + \alpha_i K_i^{\alpha_i - 1} = \frac{1}{\beta}$, and hence

$$K_i = \left(\frac{\alpha_i}{\beta^{-1} - 1 + \delta_i} \right)^{1/(1-\alpha_i)}, \qquad Y_i = K_i^{\alpha_i}.$$

World steady-state consumption is therefore

$$C = \sum_{i=1}^N (Y_i - \delta_i K_i)$$

Individual-country consumption and λ's are given by

$$C_i = \left(\frac{\omega_i}{\mu} \right)^{1/\gamma_i}, \qquad \lambda_i = \frac{\mu}{\beta},$$

with the multiplier on the world resource constraint μ solving

$$\sum_{i=1}^N \left(\frac{\omega_i}{\mu} \right)^{1/\gamma_i} = C.$$

To solve for the risky steady-state in the sequence space, we let there be a constant shift in expectations λ_i^e, so that $\mathbb{E}_t\left[\lambda_{i,t+1}\right] = \lambda_i^{rss} + \lambda_i^e$ in (A.69). Then the capital first-order condition has a risk-premium-like wedge

$$1 - \delta_i + \alpha_i \left(K_i^{rss} \right)^{\alpha_i - 1} = \frac{1}{\beta} - \frac{\lambda_i^e}{\mu^{rss}}$$

so that

$$K_i^{rss} = \left(\frac{\alpha_i}{\beta^{-1} - 1 + \delta_i - \frac{\lambda_i^e}{\mu^{rss}}} \right)^{1/(1-\alpha_i)}$$

The risky steady-state level of the world multiplier μ^{rss} still solves the fixed point equation

$$C^{rss} = \sum_{i=1}^{N} \left((K_i^{rss})^{\alpha_i} - \delta_i K_i^{rss} \right) = \sum_i \left(\frac{\omega_i}{\mu^{rss}} \right)^{1/\gamma_i}$$

from which we recover $\lambda_i^{rss} = \mu^{rss}/\beta - \lambda_i^e$.

Calibration. For our calibration, we use $\beta = 0.99, \delta_i = 0.025, \omega_i = 1$ and $\rho_a = 0.9$. For any chosen country size N, we draw independently for each country

$$\alpha_i \sim U[0.3, 0.4], \qquad \gamma_i \sim U[1, 3], \qquad \zeta_i \sim U[0.5, 1.5].$$

we take the draws before giving it to various solvers so that we can check the solution is identical across solvers.

Solution in the sequence space. Consider perfect foresight. Assuming the paths for world log TFP $\mathbf{a} \equiv (a_0, a_1, \ldots)$ and the world multiplier $\boldsymbol{\mu} \equiv (\mu_0, \mu_1, \ldots)$ are known, (A.67) implies that country i's consumption at t is

$$C_{it} = \left(\frac{\omega_i}{\mu_t} \right)^{1/\gamma_i} \tag{A.70}$$

while (A.68) implies that capital at t is

$$K_{it} = \left[\frac{\alpha_i \exp(\zeta_i a_{t+1})}{\mu_t/(\beta \mu_{t+1}) - 1 + \delta_i} \right]^{1/(1-\alpha_i)}. \tag{A.71}$$

and we can construct $NX_{it} \equiv e^{\zeta_i a_t} K_{i,t-1}^{\alpha_i} - K_{it} + (1 - \delta_i) K_{i,t-1} - C_{it}$ as net exports at time t. Letting $\mathcal{N}_{it}(\boldsymbol{\mu}, \mathbf{a})$ denote that function, the relevant market clearing condition determining $\boldsymbol{\mu}$ given the path of $\mathbf{a}$ is

$$\sum_{i=1}^{N} \mathcal{N}_{it}(\boldsymbol{\mu}, \mathbf{a}) = 0 \tag{A.72}$$

We can construct country-level sequence-space Jacobians $\mathbf{N}_{\mu i} = \frac{\partial \mathcal{N}_{it}}{\partial \mu_s}$ and $\mathbf{N}_{ai} = \frac{\partial \mathcal{N}_{it}}{\partial a_s}$ around the steady state, and then to first order the path of the world multiplier given a change in $d\mathbf{a}$ (with $d\mathbf{a} = (1, \rho_a, \rho_a^2, \ldots)\, \epsilon$ after an innovation ϵ given the TFP process) is

$$d\boldsymbol{\mu} = -\left(\sum_{i=1}^{N} \mathbf{N}_{\mu i} \right)^{-1} \left(\sum_{i=1}^{N} \mathbf{N}_{ai} \right) d\boldsymbol{a}$$

Type	Object	Max normalized error
Volterra coefficients	First-order $\mathcal{X}_i$	8.09×10^{-14}
	Second-order $\mathcal{X}_{ij}$	2.52×10^{-13}
	Second-order risk $\mathcal{X}_{\sigma\sigma}$	3.17×10^{-12}
	Third-order $\mathcal{X}_{ijk}$	6.15×10^{-12}
	Third-order risk $\mathcal{X}_{\sigma\sigma j}$	3.02×10^{-12}
Pruned moments	Second-order ergodic mean	1.30×10^{-11}
	Second-order covariance matrix	1.87×10^{-10}
Pruned simulations	Second-order simulation with Dynare burn-in	1.59×10^{-12}
	Third-order simulation with Dynare burn-in	4.10×10^{-13}

Table E.2: Comparison of sequence space vs Dynare outcomes (10 country model)

This gives the first-order aggregate unknown $\mathbf{U}$, from which we can compute the first-order solution of all variables following (A.70) and (A.71). This gives us the first-order solution in the sequence space.

To get the higher-order solutions, we use the one-shot principle: we first obtain the higher-order solutions $d^{(2)}\boldsymbol{\mu}$, $d^{(3)}\boldsymbol{\mu}$ to (A.72) by reequilibrating the nonlinear error from the lower-order solution with the sequence space Jacobian $-\left(\sum_{i=1}^{N}\mathbf{N}_{\mu i}\right)^{-1}$, and then use the resulting solution to compute the size and history terms in C_{it}, K_{it} and λ_{it}. For the second-order risk effect, we solve a perturbed expectations steady state: we use the second-order size effect λ_{i00} as our expectations perturbation λ_i^e and solve for the resulting perturbed-expectations steady state. For the third order mixed risk terms, we proceed similarly, computing both impulse responses starting from the risky steady state and the perturbed-expectations transitions induced by the mixed size-history term λ_{i00t}. We then store the Volterra coefficients $\left\{\mathcal{X}_j, \mathcal{X}_{jk}, \mathcal{X}_{jkk'}, \mathcal{X}_{\sigma\sigma}, \mathcal{X}_{\sigma\sigma j}\right\}$. We use symbolic differentiation and truncate the sequence-space systems at $T = 800$.

Solution in the state space. We use Dynare 7.1 to solve the model in the state space up to third order. We give Dynare the equations in (A.69) and use its stoch_simul routines to solve for the state-space laws of motions at orders 1, 2 and 3. Dynare uses a QZ decomposition to solve the first-order law of motion and a specialized generalized-Sylvester algorithm to solve for the higher-order law of motion, following Kameník (2003); these are the bottleneck operations as n gets large. Dynare also stores the tensors characterizing the third-order law of motion in an economical way by exploiting symmetry. We also use Dynare's pruning routines to obtain ergodic moments,[A-33] and to do pruned simulations given an arbitrary sequence of ϵ_t's.[A-34]

Verification. We verify that the outputs from the Dynare solution and the sequence space are identical in three ways. First, from the Dynare state-space law of motion, we recover the

[A-33] Called using stoch_simul(order=2,pruning), with the default periods=0 to avoid Monte-Carlo simulation
[A-34] Called using simult_(M_,options), with options.pruning set to True

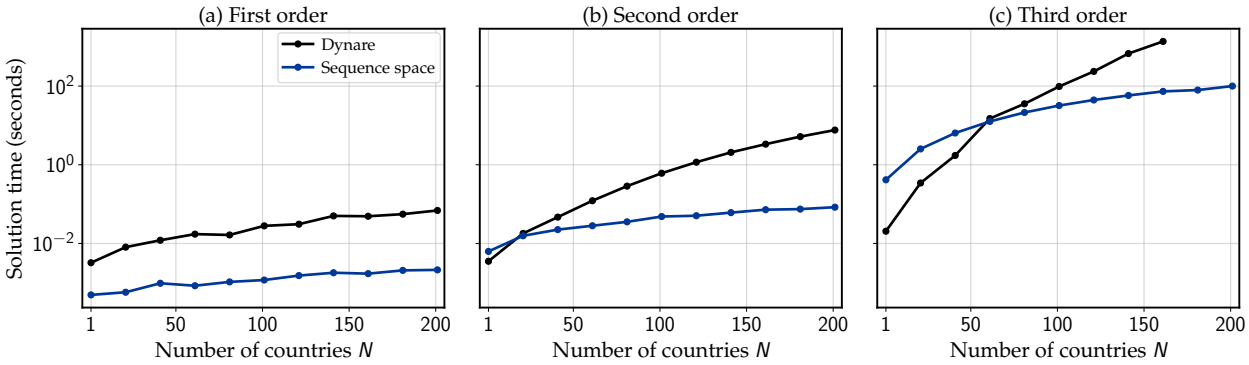


Figure E.3: Time to solution in N country model: Dynare vs sequence-space

$\left\{\mathcal{X}^{state}_{j}, \mathcal{X}^{state}_{jk}, \mathcal{X}^{state}_{jkk'}, \mathcal{X}^{state}_{\sigma\sigma}, \mathcal{X}^{state}_{\sigma\sigma j}\right\}$ and verify that they coincide with the coefficients from our sequence-space procedure. Second, we compare the pruned moments from Dynare's pruned state space solver—both ergodic mean and the pruned nth order variance matrix—to the formulas we obtain for those objects. Third, we do a pruned simulation for an arbitrary sequence $\{\epsilon_t\}$ and verify that this gives the same answer as the Volterra series on this sequence. Since Dynare starts its pruned simulations at the deterministic steady state,we allow for a 2000 burn-in period with ϵ's set to zero before feeding in our sequence, this allows time for the deterministic transition between the deterministic and the risky steady state.

Table E.2 does this verification for our $N = 10$ country simulation. The solutions agree up to a factor 10^{-10} or better.

Scaling. Figure E.3 displays the time it takes to compute the sequence-space and the Dynare (state-space) solution for first order, second order, and third order as a function of N. For first order, both are very fast, with less than 0.1 seconds taken to solve even a 200 country model, though sequence space is faster at all N. For second order, Dynare starts out faster, but the solution time grows much more rapidly with N, so that for a 200 country model Dynare takes a few seconds while sequence space remains below 0.1 seconds. The crossing point is around 20 countries. For third order, the pattern is similar, but the computation times are higher and grow faster with N. For a single country, sequence-space starts around 1 second and Dynare is 10 times faster. For 160 countries, sequence space is about 100 seconds while Dynare is 10 times slower. The crossing point is around $N = 60$, corresponding to a state space of about $n = 200$.

Figures E.4 displays the storage for the resulting solution. We only keep track of the storage used for the state-space laws of motion by Dynare, discarding additional output used for the pruning solution. The storage outcomes correspond closely to the dimension analysis in table 3. Beyond $N = 160$ (a state space of size about $n = 500$), Dynare cannot allocate enough memory to store the third order state-space law of motion. This suggests that a Dynare approach to solving heterogeneous-agent models with $n \gg 1000$ to third order would be infeasible.

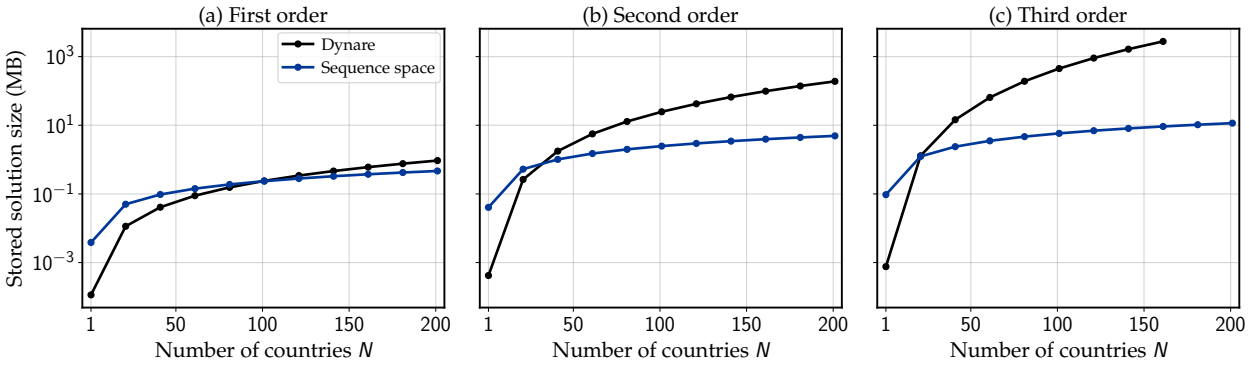


Figure E.4: Solution storage in N country model: Dynare vs sequence space

F Appendix to section 5

F.1 Quantitative HANK model details

F.1.1 General setup

The household side of the model is as in section 2.2, except that real assets pay the time-varying return r_t. Households solve the Bellman equation

$$\begin{aligned} V_t(a,e,\beta) &= \mathbb{E}_\xi\left[\max_{c,a'} u(c) + \beta \mathbb{E}_t[V_{t+1}(a',e',\beta')|e,\beta]\right] \\ \text{s.t.} \qquad & a' + c = (1+r_{t-1})a + Z_t \xi e + Tr; \qquad a' \geq \underline{a} \end{aligned} \tag{A.73}$$

where $Z_t = (1-\tau_t)\, w_t N_t$ is aggregate post-tax real wage income. Firms produce competitively under flexible prices with production function $Y_t = N_t$, so the real wage is $w_t = 1$, and post-tax income is $Z_t = (1-\tau_t)\, Y_t$.

We now allow for exogenous shocks to government spending G_t. The government budget constraint is now:

$$B_t + \tau_t Y_t = (1+r_{t-1})B_{t-1} + Tr + G_t$$

F.1.2 Monetary policy and interest rates

We assume that the central bank follows the rule in (44),

$$1 + i_t = (1+r)\,(1+\pi_t)^\phi$$

where r is the deterministic steady state interest rate and $\pi_t = P_t/P_{t-1} - 1$ is the rate of price inflation.

Following Auclert et al. (2024a), appendix A.2, we assume that households can only hold real bonds, while financial market arbitrageurs can trade both real bonds and one-period nominal bonds (reserves). The central bank sets the nominal return on reserves, which are in zero net

supply in equilibrium. We assume that arbitrageurs are risk neutral and face no borrowing or portfolio constraints.

Let r_t be the real interest rate promised at t for payment at $t+1$. Investing one unit of the consumption good in a real bond at t delivers $1+r_t$, while investing in a nominal bond delivers $\frac{(1+i_t)P_t}{P_{t+1}} = \frac{1+i_t}{1+\pi_{t+1}}$. Given risk neutrality and absence of constraints, arbitrageurs choose to borrow q_t of the consumption good to invest in nominal bonds solve

$$\max_{q_t \in \mathbb{R}} \; q_t \left\{ \mathbb{E}_t \left[\frac{1+i_t}{1+\pi_{t+1}} \right] - (1+r_t) \right\}$$

Equilibrium imposes the Fisher equation (45),

$$1 + r_t = \mathbb{E}_t \left[\frac{1+i_t}{1+\pi_{t+1}} \right]$$

and $q_t = 0$ at all time, so that the nominal bond market clears. This setting implies that realized profits of arbitrageurs are zero in every date and state, so their portfolios and profits do not enter other equilibrium conditions.

F.1.3 Wage setting

The wage setting microfoundation is similar to Auclert et al. (2024a), appendix A.3. A unit mass of unions, indexed by $k \in [0,1]$, supplies differentiated labor services. A competitive labor packer combines them using

$$N_t = \left(\int_0^1 N_{kt}^{\frac{\epsilon-1}{\epsilon}} \, dk \right)^{\frac{\epsilon}{\epsilon-1}}$$

for $\epsilon > 1$. The labor packer minimizes $\int_0^1 W_{kt} N_{kt} \, dk$ subject to delivering N_t. This delivers labor demand facing the union, as well as the wage index:

$$N_{kt} = \left(\frac{W_{kt}}{W_t} \right)^{-\epsilon} N_t, \qquad W_t = \left(\int_0^1 W_{kt}^{1-\epsilon} \, dk \right)^{\frac{1}{1-\epsilon}} \tag{A.74}$$

Each union is atomistic and takes the aggregates $\{C_t, N_t, W_t, P_t, \tau_t\}$ as given. We define the consumption index used by union k to be

$$\widetilde{C}_{kt} = C_t + \frac{1}{\mu_t}(1-\tau_t) \left(\frac{W_{kt}}{P_t} N_{kt} - w_t N_t \right). \tag{A.75}$$

where $w_t = W_t / P_t$ is the real wage. This index is equal to aggregate consumption C_t in a symmetric allocation, and values changes in wages as if income and savings are held fixed, modulo an exogenous multiplicative wedge μ_t. Given its initial nominal wage $W_{k,-1}$, we assume that union k

chooses a path of nominal wages to solve

$$\max \mathbb{E}_0 \sum_{t=0}^{\infty} \bar{\beta}^t \left(u(\widetilde{C}_{kt}) - v(N_{kt}) - \frac{\psi}{2} \left(\frac{W_{kt}}{W_{k,t-1}} - 1 \right)^2 \right) \tag{A.76}$$

subject to labor demand (A.74) and the consumption index (A.75).

Write $x_{kt} \equiv \log W_{kt}$ and $\pi^w_{kt} \equiv W_{kt}/W_{k,t-1} - 1$. Consider a change in x_{kt} at time t, holding other wage choices and aggregate fixed at all dates. Labor demand (A.74) and the consumption index equation (A.75) implies

$$\frac{\partial N_{kt}}{\partial x_{kt}} = -\epsilon N_{kt}, \qquad \frac{\partial \widetilde{C}_{kt}}{\partial x_{kt}} = (1-\epsilon) \frac{1-\tau_t}{\mu_t} \frac{W_{kt}}{P_t} N_{kt}$$

so the derivative of the flow payoff from consumption and labor is

$$\frac{\partial}{\partial x_{kt}} \left[u(\widetilde{C}_{kt}) - v(N_{kt}) \right] = \epsilon N_{kt} \left[v'(N_{kt}) - \frac{\epsilon - 1}{\epsilon} \frac{1-\tau_t}{\mu_t} \frac{W_{kt}}{P_t} u'(\widetilde{C}_{kt}) \right]$$

The date-t wage also enters the adjustment costs at t and $t+1$ in (A.76). Since

$$\frac{\partial \pi^w_{kt}}{\partial x_{kt}} = 1 + \pi^w_{kt}, \qquad \frac{\partial \pi^w_{k,t+1}}{\partial x_{kt}} = -(1 + \pi^w_{k,t+1}),$$

the overall first order condition for (A.76) with respect to x_{kt} delivers

$$0 = \epsilon N_{kt} \left[v'(N_{kt}) - \frac{\epsilon - 1}{\epsilon} \frac{1-\tau_t}{\mu_t} \frac{W_{kt}}{P_t} u'(\widetilde{C}_{kt}) \right] - \psi \pi^w_{kt} (1 + \pi^w_{kt}) + \bar{\beta} \psi \mathbb{E}_t \left[\pi^w_{k,t+1} (1 + \pi^w_{k,t+1}) \right] \tag{A.77}$$

In a symmetric equilibrium, $W_{kt} = W_t$, $N_{kt} = N_t$, and $\widetilde{C}_{kt} = C_t$. Substituting this into (A.77) and dividing by ψ delivers the wage Phillips curve

$$\pi^w_t (1 + \pi^w_t) = \frac{\epsilon}{\psi} N_t \left(v'(N_t) - \frac{\epsilon - 1}{\epsilon} \frac{1-\tau_t}{\mu_t} w_t u'(C_t) \right) + \bar{\beta} \mathbb{E}_t \left[\pi^w_{t+1} (1 + \pi^w_{t+1}) \right] \tag{A.78}$$

where $\kappa = \frac{\epsilon}{\psi}$, and $\pi^w_t = W_t/W_{t-1} - 1$ is the rate of wage inflation.

Finally, recall that given competitive flexible-price production $Y_t = N_t$, the real wage satisfies $w_t = 1$ at all dates, so in particular $\pi^w_t = \pi_t$. Substituting these relationships into the wage Phillips curve delivers

$$\pi_t (1 + \pi_t) = \kappa N_t \left(v'(N_t) - \frac{\epsilon - 1}{\epsilon} \frac{1-\tau_t}{\mu_t} u'(C_t) \right) + \bar{\beta} \mathbb{E}_t \left[\pi_{t+1} (1 + \pi_{t+1}) \right]$$

which is equation (47) in the main text.

F.1.4 Summary of equations

The quantitative HANK model is characterized by the following set of equations:

- The implementation of the Bellman equation (A.73) and its corresponding law of motion and aggregation equation, i.e. equations (A.41)–(A.50), delivering

$$\begin{aligned} v_t &= \mathcal{V}(Z_t, r_{t-1}, \mathbb{E}_t v_{t+1}) && \text{(A.79)} \\ \boldsymbol{\phi}_t &= \Lambda(Z_t, r_{t-1}, \mathbb{E}_t v_{t+1}, \boldsymbol{\phi}_{t-1}) && \text{(A.80)} \\ A_t &= \mathcal{A}(Z_t, r_{t-1}, \mathbb{E}_t v_{t+1}, \boldsymbol{\phi}_{t-1}) && \text{(A.81)} \\ C_t &= \mathcal{C}(Z_t, r_{t-1}, \mathbb{E}_t v_{t+1}, \boldsymbol{\phi}_{t-1}) && \text{(A.82)} \end{aligned}$$

- The government budget constraint and the definition of post-tax income:

$$\begin{aligned} B_t + \tau_t Y_t &= (1 + r_{t-1}) B_{t-1} + Tr + G_t && \text{(A.83)} \\ Z_t &= (1 - \tau_t)\, Y_t && \text{(A.84)} \end{aligned}$$

- The monetary rule (44) and Fisher equation (45), which combined, deliver

$$\begin{aligned} X_t^1 &= \frac{1}{1 + \pi_t} && \text{(A.85)} \\ 1 + r_t &= (1 + r)\,(1 + \pi_t)^{\phi}\, \mathbb{E}_t \left[X_{t+1}^1 \right] && \text{(A.86)} \end{aligned}$$

  where X_t^1 is an auxiliary variable we define so that expectations enter via $\mathbb{E}_t \left[X_{t+1}^1 \right]$.

- The wage Phillips curve (47), substituting the production function $N_t = Y_t$, yielding

$$\begin{aligned} X_t^2 &= \pi_t (1 + \pi_t) && \text{(A.87)} \\ X_t^2 &= \kappa Y_t \left(v'(Y_t) - \frac{\epsilon - 1}{\epsilon} \frac{1 - \tau_t}{\mu_t} u'(C_t) \right) + \tilde{\beta} \mathbb{E}_t \left[X_{t+1}^2 \right] && \text{(A.88)} \end{aligned}$$

  where X_t^2 is a second auxiliary variable.

- The stochastic processes for bonds, government spending and markups, yielding

$$\begin{aligned} B_t &= B + \rho_B (B_{t-1} - B) + \epsilon_t^B && \text{(A.89)} \\ G_t &= G + \rho_G (G_{t-1} - G) + \epsilon_t^G && \text{(A.90)} \\ \mu_t &= 1 + \rho_p (\mu_{t-1} - 1) + \epsilon_t^p && \text{(A.91)} \end{aligned}$$

- The market clearing conditions for goods and assets

$$A_t = B_t, \qquad C_t + G_t = Y_t \tag{A.92}$$

Letting $X_t = \left(v_t, \boldsymbol{\phi}_t, A_t, C_t, Y_t, \tau_t, Z_t, r_t, \pi_t, X_t^1, X_t^2, B_t, G_t, \mu_t\right)$, the F system is determined by the fourteen equations (A.79)–(A.92), which fit the framework in (1). Note that by Walras's law, the asset and goods market clearing conditions in (A.92) are equivalent. We use the former in practice.

F.1.5 Reduced set of perfect-foresight equations

Following section 4.1, we can reduce the problem of finding the perfect-foresight solution by solving for the smaller set of unknown variables $U_t = \{Y_t, \pi_t\}_{t\geq 0}$ from any date 0. Given a nonlinear sequence of current and past shocks $\left\{\epsilon^G_{-k}, \epsilon^B_{-k}, \epsilon^p_{-k}\right\}_{k=0..L}$, we can solve for the time path of $\{B_t\}_{t=-L}^T$ by initializing $B_{-L} = B_{ss}$ and then solving recursively, for $k = -L, \ldots, 0$,

$$B_{-k} = B_{ss} + \rho_B \left(B_{-k-1} - B_{ss}\right) + \epsilon^B_{-k}$$

and then for $t > 0$

$$B_t = B_{ss} + \rho_B \left(B_{t-1} - B_{ss}\right)$$

We proceed similarly for G_t and μ_t.

Given this and knowledge of $\{Y_t\}_{t\geq 0}$, we obtain the tax rate

$$\tau_t = \frac{(1 + r_{t-1})B_{t-1} + Tr + G_t - B_t}{Y_t}$$

and given knowledge of $\{\pi_t\}_{t\geq 0}$, we obtain the real interest rate

$$1 + r_t = (1 + r)\left(1 + \pi_t\right)^{\phi} \frac{1}{1 + \pi_{t+1}}$$

we then evaluate the distribution at date 0 by initializing , $\boldsymbol{\phi}_{-L} = \boldsymbol{\phi}_{ss}$ and then solving for $k = -L, \ldots, 0$,

$$\begin{aligned} v_{-k} &= \mathcal{V}(Z_{-k}, r_{-k}, \mathbb{E}_{-k} v_{-k+1}) \\ \boldsymbol{\phi}_{-k} &= \Lambda(Z_{-k}, r_{-k}, \mathbb{E}_{-k} v_{-k+1}, \boldsymbol{\phi}_{-k-1}) \end{aligned}$$

where $\mathbb{E}_{-k} v_{-k+1}$ assumes the sequence of innovations $\{\epsilon_{-L}, \epsilon_{-L+1}, \ldots, \epsilon_{-k}, 0, 0, \ldots\}$.

Finally, we evaluate the residuals

$$H_t^1 \equiv A_t\left(\{Y_s - \tau_s Y_s, r_s\}, \boldsymbol{\phi}_{-1}\right) - B_t \tag{A.93}$$

and

$$H_t^2 = \pi_t\left(1 + \pi_t\right) - \kappa Y_t \left(v'\left(Y_t\right) - \frac{\epsilon - 1}{\epsilon} \frac{1 - \tau_t}{\mu_t} u'\left(Y_t - G_t\right)\right) - \tilde{\beta} \pi_{t+1}\left(1 + \pi_{t+1}\right) \tag{A.94}$$

overall, these mappings define a nonlinear function

$$H_t\left(\{U_s\}_{s\geq 0}, \left\{\epsilon^G_{-k}, \epsilon^B_{-k}, \epsilon^p_{-k}\right\}_{k=0..L}\right), \quad t \geq 0 \tag{A.95}$$

given any guess for unknowns $\{U_s\} = \{Y_s, \pi_s\}_{s\geq 0}$.

If we want to recover the whole X_t, we form $Z_t = (1-\tau_t)Y_t$, and then get $v_t, \boldsymbol{\phi}_t, A_t, C_t, X^1_t, X^2_t$ from their definitions in (A.79)–(A.82), (A.85), (A.87).

The one-shot principle requires getting the Jacobian $\mathbf{H_U}$ of the targets (A.93)-(A.94) at the steady state $(\mathbf{U}_{ss}, 0)$. Then, for given sequence $\left\{\epsilon^G_{-k}, \epsilon^B_{-k}, \epsilon^p_{-k}\right\}_{k=0..L}$, to obtain accuracy of order k for unknowns, we form the sequence $\mathbf{U}^{[0]} = \mathbf{U}_{ss}$,

$$\mathbf{U}^{[k]} = \mathbf{U}^{[k-1]} - \mathbf{H}_{\mathbf{U}}^{-1}\mathbf{H}\left(\mathbf{U}^{[k-1]}, \{\epsilon_{-k}\}_{k=0..L}\right)$$

following (35), which requires only k nonlinear evaluations of the $\mathbf{H}$ function in (A.95).

F.1.6 Deterministic steady state

We maintain the same calibration of the household part of the model described in section 2.2 and appendix A.5. We assume that the Taylor rule coefficient is $\phi = 1.5$. We set the slope of the Phillips curve to $\kappa = 0.01$ and assume isoelastic disutility of labor with a Frisch elasticity of 0.5, so $v'(N) = \zeta N^2$. We assume a steady state inverse wage markdown of $\frac{\epsilon}{\epsilon-1} = 1.1$.

We consider the deterministic steady state with zero inflation $\pi = 0$, implying that ζ solves

$$\zeta N^2 = \frac{\epsilon - 1}{\epsilon}\frac{1}{\mu}(1-\tau)C^{-\gamma}$$

at our deterministic steady state values of $\frac{\epsilon}{\epsilon-1} = 1.1$, $\mu = 1$, $\tau = 0.33$, $C = Y - G = 0.83$, $\gamma = 1$, and $N = Y = 1$ (see table 1), implying $\zeta = 0.73384447$.

F.1.7 Perturbed expectations steady state

To obtain a perturbed expectations steady state, we note that expectations enter the F system via $\mathbb{E}_t v_{t+1}$, $\mathbb{E}_t X^1_{t+1}$ and $\mathbb{E}_t X^2_{t+1}$. Hence, we obtain the size dependence v_{00}, X^1_{00} and X^2_{00}. We then consider the nonlinear steady state function

$$H(U, \Delta)$$

which gives the residuals in conditions (A.93) and (A.94) from using unknowns $U = (Y, \pi)$ when Δ gives the perturbation to v_{t+1}, X^1_{t+1} and X^2_{t+1}. Finally, to correctly solve to first order we use equation (40),

$$U_{\sigma\sigma} = -H_U^{-1}\frac{\partial H\left(U_{ss}, \left(v_{00}, X^1_{00}, X^2_{00}\right)\delta\right)}{\partial \delta}$$

We follow a similar principle for perturbed expectations transitions.

F.1.8 GMM estimation

We use our higher-order characterization of first and second moments in equation (18) and proposition 6 to estimate our model parameters using the generalized method of moments (GMM). Let $\boldsymbol{\theta}$ denote the vector of parameters that we are estimating, $\hat{\boldsymbol{m}}$ denote a vector of empirical moments, and $\boldsymbol{m}(\boldsymbol{\theta})$ be the model counterpart of these moments, obtained from solving the model at $\boldsymbol{\theta}$. The GMM estimator solves

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} (\boldsymbol{m}(\boldsymbol{\theta}) - \hat{\boldsymbol{m}})' W (\boldsymbol{m}(\boldsymbol{\theta}) - \hat{\boldsymbol{m}}) \tag{A.96}$$

for a given positive-definite weighting matrix W.[A-35]

In our quantitative HANK model application, we estimate the autocorrelation and standard deviation of the model's shock processes to match the observed variance and first autocovariance of the output gap and inflation, as well as their contemporaneous covariance (see table 4). Here, our vector of empirical moments is

$$\hat{\boldsymbol{m}} = \left(\hat{\mathbb{E}}[Y_t^2],\ \hat{\mathbb{E}}[Y_t Y_{t-1}],\ \hat{\mathbb{E}}[\pi_t^2],\ \hat{\mathbb{E}}[\pi_t \pi_{t-1}],\ \hat{\mathbb{E}}[Y_t \pi_t] \right)', \tag{A.97}$$

where Y_t and π_t denote the demeaned time series of the output gap and inflation, respectively, while $\boldsymbol{m}(\boldsymbol{\theta})$ denotes the model counterpart of these moments. We use the optimal weighting matrix $W = \hat{S}^{-1}$, where $\hat{S}$ denotes the Newey and West (1987) heteroskedasticity and autocorrelation consistent (HAC) estimator of the long-run variance matrix $T \cdot \mathrm{Var}(\hat{\boldsymbol{m}})$, where T denotes the sample size.[A-36]

We compute the output gap as the difference between real and potential GDP, divided by potential GDP, and use the PCE as our measure of the price level. We use quarterly data from 1966Q1 up to 2025Q4 obtained from the Bureau of Economic Analysis (GDP and PCE) and the Congressional Budget Office (potential GDP).

We calculate standard errors for our parameters using the standard expression for the asymptotic variance of our estimator given the optimal weighting matrix (Hansen 1982):

$$\mathrm{Var}(\hat{\boldsymbol{\theta}}) = \left(G' \hat{S}^{-1} G \right)^{-1} / T, \tag{A.98}$$

where $G = \partial \boldsymbol{m}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}'$, the sensitivity matrix of moments to parameters, is evaluated at $\hat{\boldsymbol{\theta}}$.

F.1.9 Ergodic distribution of the second-order pruned solution

Here we show how to numerically compute the entire ergodic density of a second-order MA process as in (11), assuming normally distributed errors $\boldsymbol{\epsilon}$. The idea is to evaluate its characteristic

[A-35] More generally, the moment conditions may not be additively separable into a purely data-dependent piece and a purely $\boldsymbol{\theta}$-dependent one. In that case, a two-step procedure would be required to obtain the optimal weighting matrix W that minimizes the asymptotic variance of $\hat{\boldsymbol{\theta}}$.

[A-36] In our implementation of the HAC estimator, we use a Bartlett kernel with weights $1 - \ell/(L+1)$ for $\ell = 1, \ldots, L$ and maximum lag of $L = 4$. Our time series has length $T = 240$ (1966Q1–2025Q4).

function $\phi_{X_t}(\tau) \equiv \mathbb{E}\left[e^{i\tau X_t}\right]$, which has a closed-form expression, and then apply Fourier inversion.

Start by approximating $X_t^{(2)}$ by a quadratic form up to a truncation horizon L:

$$X_t^{(2)} \approx X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2 + \mathbf{d}'\boldsymbol{\epsilon} + \frac{1}{2}\boldsymbol{\epsilon}'\mathbf{D}\boldsymbol{\epsilon}, \tag{A.99}$$

where $\boldsymbol{\epsilon} = (\epsilon_t, \epsilon_{t-1}, \ldots, \epsilon_{t-(L-1)})'$, $\mathbf{d} = (\mathcal{X}_0, \ldots, \mathcal{X}_{L-1})'$, and $\mathbf{D} = [\mathcal{X}_{jk}]_{0\le j,k\le L-1}$.

The characteristic function of X_t can be expressed in closed form (for instance, see Imhof, 1961):

$$\phi_{X_t}(\tau) = \det\left(\mathbf{I} - i\tau\sigma^2\mathbf{D}\right)^{-1/2} \exp\left(i\tau\left(X + \frac{1}{2}\mathcal{X}_{\sigma\sigma}\sigma^2\right) - \frac{\sigma^2\tau^2}{2}\mathbf{d}'\left(\mathbf{I} - i\tau\sigma^2\mathbf{D}\right)^{-1}\mathbf{d}\right). \tag{A.100}$$

From (A.100), the probability density function $f_{X_t}(x)$ is obtained via the Fourier inversion formula: $f_{X_t}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tau x}\phi_{X_t}(\tau)\,d\tau$.[A-37] We use this expression to calculate the density reported in figure 11(b), as well as in the menu cost model in figure 14(b).

F.2 Quantitative menu cost model details

F.2.1 Detrended model

Recall that $m^\star$ denotes the constant desired log markup over marginal cost. Following Auclert et al. (2024c), we incorporate this markup into the definition of the price gap. If P_{it} is firm i's log nominal price and a_{it} its log productivity, its flexible log price is $m^\star + MC_t - a_{it}$. We define $x_{it} = P_{it} + a_{it} - m^\star$, so that its deviation from the flexible log price is $x_{it} - MC_t$.

We first detrend the model to make it stationary. Denote by $y = x - \mu t$ the detrended price gap, and by $w_t = W_t - \mu t$, $mc_t = MC_t - \mu t$ and $p_t = P_t - \mu t$ the detrended log nominal wage, marginal cost and actual price, respectively. We therefore have

$$mc_t = \chi w_t + (1-\chi)p_t \tag{A.101}$$

Note that if there is no adjustment, the price gap y evolves as $y' = y + \epsilon - \mu$. The Bellman equation is therefore given by

$$\begin{aligned} v_t(y) &= \tfrac{1}{2}(y - mc_t)^2 + \beta\mathbb{E}_t\left[\int u_{t+1}(y - \mu + \epsilon)f(\epsilon)\,d\epsilon\right] \\ u_t(y) &\equiv \lambda v_t(y_t^*) + (1-\lambda)\min\{v_t(y), v_t(y_t^*) + \xi\} \end{aligned}$$

where $v_t(y)$ denotes the loss from an incoming price gap of y at t, and $u_t(y)$ is the expected loss

[A-37] While this approach allows us to compute univariate marginal densities, evaluating the joint likelihood of an observation vector $(X_1, \ldots, X_T)$ for maximum likelihood estimation of second-order models is much more challenging. Evaluating a sample of size T requires integrating a T-dimensional characteristic function, which induces a severe curse of dimensionality even for modest sample sizes.

before the adjustment decision but after the idiosyncratic cost shock. Let

$$y_t^* = \arg\min_y v_t(y)$$

denote the reset price gap, and define the upper and lower boundaries as

$$v_t(\ell_t) = v_t(r_t) = v_t(y_t^*) + \xi$$

so that the inaction region is $[\ell_t, r_t]$.

Let $g_t(y)$ be the density after date-t idiosyncratic shocks but before date-t adjustment, with $\int_{\mathbb{R}} g_t(y)\,dy = 1$. With this convention, the frequency of price adjustment is

$$\text{freq}_t = 1 - (1-\lambda)\int_{\ell_t}^{r_t} g_t(z)\,dz \tag{A.102}$$

Under linear aggregation, the mean detrended price gap q_t and the actual detrended log price are given by

$$\begin{aligned} q_t &= (1-\lambda)\int_{\ell_t}^{r_t} z g_t(z)\,dz + \text{freq}_t y_t^*, \\ p_t &= m^\star + q_t. \end{aligned} \tag{A.103}$$

Finally, the evolution of the density is given by

$$g_{t+1}(y) \;=\; (1-\lambda)\int_{\ell_t}^{r_t} g_t(z) f(y - z + \mu)\,dz + \text{freq}_t f(y - y_t^* + \mu) \tag{A.104}$$

Note that, since $\int_{\mathbb{R}} f(\epsilon)\,d\epsilon = 1$ and $\int_{\mathbb{R}} \epsilon f(\epsilon)\,d\epsilon = 0$, we have, using $\epsilon = y - z + \mu$, that $\int_{\mathbb{R}} y f(y - z + \mu)\,dy = z - \mu$. Integrating the density equation therefore implies that the mean pre-adjustment price gap evolves as

$$\begin{aligned} \int_{\mathbb{R}} g_{t+1}(y) y\,dy &= (1-\lambda)\int_{\ell_t}^{r_t} g_t(z)\,(z-\mu)\;dz + \text{freq}_t\,(y_t^* - \mu) \\ &= q_t - \mu\left((1-\lambda)\int_{\ell_t}^{r_t} g_t(z)\,dz + \text{freq}_t\right) \\ &= q_t - \mu = p_t - m^\star - \mu. \end{aligned} \tag{A.105}$$

Deterministic steady state. In the steady state, the equations above hold without t's. We have

$$\begin{aligned} q &= (1-\lambda)\int_{\ell}^{r} z g(z)\,dz + \text{freq}\cdot y^* \\ p &= m^\star + q \\ mc &= \chi w + (1-\chi)p \end{aligned}$$

where w is the exogenous detrended nominal wage. Actual inflation is $\pi = \mu$. Given (A.105), we also have $q = \mu + \int_{\mathbb{R}} g(y) y \, dy$. Combining the two expressions for q gives

$$\int_{-\infty}^{\ell} (y^* - y) g(y) \, dy + \lambda \int_{\ell}^{r} (y^* - y) g(y) \, dy + \int_{r}^{\infty} (y^* - y) g(y) \, dy = \mu$$

which decomposes inflation into nominal log price changes from y to y^* in all three parts of the distribution of gaps. We note that $q - mc$ is the average deviation from the flexible log price, so that the log price–marginal-cost markup is $p - mc = m^\star + q - mc$. The corresponding relations between w, mc and p are

$$p = w + \frac{1}{\chi} (m^\star + q - mc) \qquad mc = w + \frac{1 - \chi}{\chi} (m^\star + q - mc) \,. \tag{A.106}$$

Hence the log price–wage markup is $(m^\star + q - mc)/\chi$.

F.2.2 Numerical implementation

We represent the value function $v_t(y)$ and the pre-adjustment density $g_t(y)$ by their values on a common uniform grid with $N = 1000$ points, and interpolate both objects using degree-five, not-a-knot splines, which have four continuous derivatives.

Backward iteration. Denote by $\widehat{v}_{t+1}$ the interpolating spline for the value function at $t+1$. We obtain the associated policies y^*_{t+1} by solving $\widehat{v}'_{t+1}(y^*_{t+1}) = 0$, and ℓ_{t+1} and r_{t+1} by solving $\widehat{v}_{t+1}(y) = \widehat{v}_{t+1}(y^*_{t+1}) + \xi$ using a root-finding method. Since shock innovations have density $f(\epsilon) = \phi(\epsilon/\sigma_\epsilon)/\sigma_\epsilon$ where ϕ is the standard normal density, we then calculate $\mathcal{E}_{t+1}(y) \equiv \mathbb{E}_\epsilon [u_{t+1}(y - \mu + \epsilon)]$ as

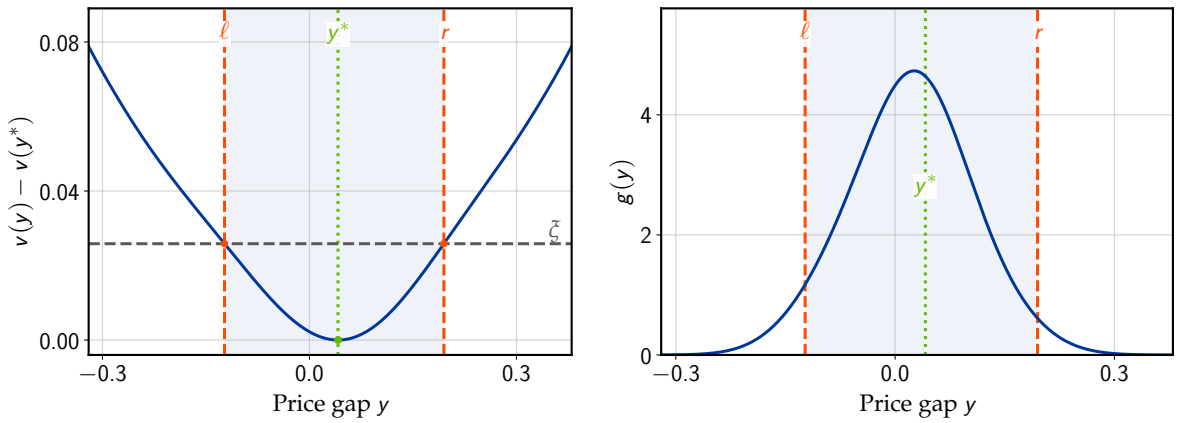
$$\begin{aligned} \mathcal{E}_{t+1}(y) = {} & \lambda \widehat{v}_{t+1}(y^*_{t+1}) + (1-\lambda) \int_{\ell_{t+1}}^{r_{t+1}} \widehat{v}_{t+1}(z) f(z - y + \mu) \, dz \\ & + (1-\lambda) \left(1 - \Phi\left(\frac{r_{t+1} - y + \mu}{\sigma_\epsilon} \right) + \Phi\left(\frac{\ell_{t+1} - y + \mu}{\sigma_\epsilon} \right) \right) (\widehat{v}_{t+1}(y^*_{t+1}) + \xi) \end{aligned} \tag{A.107}$$

where Φ is the standard normal cdf. We evaluate the integral in (A.107) over the inaction region $[\ell_{t+1}, r_{t+1}]$ with 40 Gauss-Legendre nodes. Given this value at $t+1$, we then obtain the Bellman update $\widehat{v}_t(y) = \frac{1}{2}(y - mc_t)^2 + \beta \mathbb{E}_t [\mathcal{E}_{t+1}(y)]$ at t, and store these values on the grid. This defines a function

$$\mathbf{v}_t = \mathcal{V}(mc_t, \mathbb{E}_t [\mathcal{E}(\mathbf{v}_{t+1})]) \tag{A.108}$$

where we note that aggregate shocks enter via the expectation of $\mathcal{E}_{t+1}(y)$ in (A.107).

Forward iteration and aggregation. Denote by $\widehat{g}_t$ the interpolating spline for the pre-adjustment density at t, and $\{y^*_t, \ell_t, r_t\}$ the associated policies. We implement (A.102) by calculating the fre-



(a) Value function relative to minimum $v(y) - v(y^*)$

(b) Pre-adjustment density $g(y)$

Figure F.1: Deterministic steady state of the menu cost model

quency of price adjustment freq_t and $g_{t+1}(y)$ as

$$\text{freq}_t \quad = \quad 1 - (1-\lambda) \int_{\ell_t}^{r_t} \widehat{g}_t(z)\, dz$$

by integrating $\widehat{g}_t(z)$ exactly as a piecewise polynomial. We implement (A.104) by calculating

$$g_{t+1}(y) = (1-\lambda) \int_{\ell_t}^{r_t} \widehat{g}_t(z) f(y - z + \mu)\, dz + \text{freq}_t f(y - y_t^* + \mu)$$

using Gauss-Legendre quadrature with 40 nodes for the integral. Finally, we calculate the mean detrended price gap and the actual detrended log price by implementing (A.103) as

$$q_t = (1-\lambda) \int_{\ell_t}^{r_t} z \widehat{g}_t(z)\, dz + \text{freq}_t y_t^*, \qquad p_t = m^\star + q_t.$$

Let $\mathcal{Q}$ denote the aggregation map for q_t. This gives us the functions

$$\mathbf{g}_{t+1} \quad = \quad \mathcal{G}\left(mc_t, \mathbb{E}_t\left[\mathcal{E}\left(\mathbf{v}_{t+1}\right)\right], \mathbf{g}_t\right) \tag{A.109}$$

$$p_t \quad = \quad m^\star + \mathcal{Q}\left(mc_t, \mathbb{E}_t\left[\mathcal{E}\left(\mathbf{v}_{t+1}\right)\right], \mathbf{g}_t\right) \tag{A.110}$$

Deterministic steady state. We iterate on the Bellman equation (A.108) and the distribution update equation (A.109) until convergence to stationary $v(y)$ and $g(y)$. Let $\overline{q}$ denote the solution to these equations when we normalize $\overline{mc} = 0$ and $m^* = 0$. Then, given (A.106), the solution when we normalize $w = 0$ and have an arbitrary m^* instead is

$$p = \frac{m^\star + \overline{q}}{\chi}, \qquad mc = \frac{1-\chi}{\chi}\left(m^\star + \overline{q}\right), \qquad q = \overline{q} + mc = p - m^\star.$$

As described in the main text, we calibrate ξ, λ, and σ_ϵ to match, in the deterministic steady

state, a frequency of price adjustment of freq $= 0.239$, a median absolute size of price changes of 0.085, a share of free adjustments in total adjustment of 0.75, trend inflation at $\mu = 0.8512\%$ and a share of value added in production of $\chi = 0.7$. We also calibrate m^* so that the price-wage markup in the deterministic steady state is $p - w = 10\%$; this implies that the price-marginal cost markup is $p - mc = \chi(p - w) = 7\%$.

This calibration implies $(\xi, \lambda, \sigma_\epsilon, m^*) = (0.025857, 0.17925, 0.058966, 0.07018688)$. Figure F.1 visualizes the steady-state value $v(y)$ relative to $v(y^*)$ and the pre-adjustment density $g(y)$ under the normalization $w = 0$. The adjustment bands are $\ell = -0.12312$ and $r = 0.19484$, with reset point $y^* = 0.04143$. Due to trend inflation, firms reset to a price gap slightly above steady state log marginal cost; this price gap then drifts down at rate μ. The mean post-adjustment gap is $q = 0.02981312$, so that $p = m^\star + q$.

Nonlinear transitions. To solve nonlinear transitions under perfect foresight, we use p_t as our endogenous unknown variable. Our sequence-space residual function is

$$H_t = p_t - m^\star - \mathcal{Q}_t\left(\{\chi w_s + (1-\chi)p_s\}\right)$$

where $\mathcal{Q}_t$ returns the mean detrended price gap from the sequence-space version of (A.108)–(A.110) with initial condition given by the steady state g. To first order, the sequence $d\mathbf{p} = (dp_0, \ldots)$ solving $H(\mathbf{p}, \mathbf{w}) = 0$ for given $d\mathbf{w} = (dw_0, \ldots)$ is given by

$$\left(\mathbf{I} - (1-\chi)\mathbf{\Psi}\right) d\mathbf{p} \quad = \quad \chi \mathbf{\Psi} d\mathbf{w}$$

where $\mathbf{\Psi}$ is the Jacobian of the mean gap to marginal cost, equivalently of the actual price to marginal cost since $dp_t = dq_t$ for fixed $m^\star$. We use the explicit expression for this passthrough matrix derived in Auclert et al. (2024c). We continue to use the inverse Jacobian $\left(\mathbf{I} - (1-\chi)\mathbf{\Psi}\right)^{-1}$ for one-shot corrections at every order.

For second-order derivatives, we use the central symmetric difference formula (42) with step size $h_2 = 0.0002$. For third-order derivatives, we use the seven-point stencil (A.63) with step size $h_3 = 0.002$. We continue to use polarization identities to obtain mixed derivatives. We assess numerical accuracy by using the step-size test (h test) from section E.4, and by varying the grid size $N \in \{500, 1000, 2000, 4000\}$. For the size derivative and four mixed-history derivatives, the default $N = 1000$ differs from a quintic numerical reference with $N = 4000$ and the same $h_3 = 0.002$ by at most 0.00006%. Testing $h_3 \in \{0.001, 0.002, 0.004\}$ at $N = 1000$ gives discrepancies of at most 0.0012% against that reference. At $N = 1000$ and $h_3 = 0.002$, halving h_2 or doubling quadrature nodes from 40 to 80 changes these third-order paths by at most 0.000011% in either case. All discrepancies are maximum absolute differences over the first 20 periods, expressed as percentages of each reference path's peak absolute value; the last two checks use the default-grid path as reference.

Perturbed expectations steady states and transitions. Equations (A.108)–(A.110) show that the primitive system has aggregate shocks entering via expectations of $\mathcal{E}_{t+1}(y)$ in (A.107). For perturbed expectations, we differentiate the function $\mathcal{E}_0(y;\nu)$ along the general-equilibrium path generated by an aggregate wage innovation ν. Let $\mathcal{E}_{00}(y)$ denote the second derivative. We solve for the steady state price $p(\delta)$ obtained by perturbing the expectation in (A.107) by the constant $\delta\mathcal{E}_{00}(y)$, and then obtain $p_{\sigma\sigma} \simeq \frac{p(\delta)-p(-\delta)}{2\delta}$ (since m^* is fixed, this is also $q_{\sigma\sigma}$). Next, let $\mathcal{E}_{00t}(y)$ denote the third derivative with respect to a squared innovation ν and a prior shock occurring t periods ago. We solve for the RSS term at third order by starting transitions from the resulting risky steady state; and the TVP term by calculating the transition generated by the time-varying expectation sequence $\mathcal{E}_{00t}(y)$.