When do Endogenous Portfolios Matter for HANK?

Adrien Auclert* Matthew Rognlie† Ludwig Straub‡ Tomáš Ťapák§

March 2024

Abstract

Most of the literature studying heterogeneous-agent New Keynesian models takes household portfolios as exogenously given. What changes when agents are allowed to hedge aggregate risk? We develop a simple sequence-space method to solve for endogenous portfolios, impulse responses, and second-order risk premia in heterogeneous-agent models. Applying our method to a simple HANK model, we show that the effects of monetary shocks and balanced-budget fiscal shocks are unchanged, but that unrestricted portfolio choice can significantly attenuate the response to deficit-financed fiscal shocks. The associated hedging portfolios appear counterfactual. Imposing more realistic short-sale constraints, or adding additional aggregate shocks to the model so that markets are incomplete with respect to these shocks, usually restores standard outcomes.

*Stanford University, CEPR and NBER. Email: aauclert@stanford.edu.
†Northwestern University and NBER. Email: matthew.rognlie@northwestern.edu.
‡Harvard University, CEPR and NBER. Email: ludwigstraub@fas.harvard.edu.
§Stanford University. Email: tapak@stanford.edu.
This research is supported by the National Science Foundation grant SES-2042691. We thank Fernando Alvarez, Anmol Bhandari, Mike Golosov, Lars Hansen, and Rohan Kekre for helpful comments.

1 Introduction

Over the past decade, macroeconomists have turned to heterogeneous-agent models to revisit the aggregate effects of monetary and fiscal policy and study their distributional implications (see e.g. Kaplan, Moll and Violante 2018, Auclert, Rognlie and Straub 2023a, and many others.) The vast majority of this Heterogeneous-Agent New Keynesian literature studies settings where agents hold asset portfolios that are exogenously fixed.[1] This assumption is natural given that these models are either solved with first-order perturbation methods, or by assuming a perfect-foresight economy hit by unanticipated "MIT" shocks: in either case, the portfolio choice is indeterminate, and so the model is consistent with any distribution of household portfolios. However, agents that perceive aggregate risk and that can invest their wealth in different types of assets have a well-defined optimal portfolio in the neighborhood of the steady state. It is reasonable to assume that allowing agents to hold these endogenous, "zeroth-order" porfolios (Devereux and Sutherland 2011) may mute some of the redistribution channels highlighted in the existing literature, and perhaps overturn some of its main conclusions.

In this paper, we quantify the extent to which endogenous portfolios can matter. We ask how the core results of the HANK literature change once agents are allowed to hedge aggregate risk by holding optimal rather than exogenous portfolios. We find that the effect of hedging can be quite large, although this requires allowing agents to have access to many types of assets (so that markets are essentially complete with respect to aggregate risk), and letting them take potentially large short positions in those assets.

We begin by proposing a new sequence-space method for solving for the zeroth-order portfolios in general equilibrium heterogeneous-agent models, extending the results in Auclert, Bardóczy, Rognlie and Straub (2021). The idea is to study the portfolio choice at a date -1 when shocks can only realize at date 0. We show, using a second-order perturbation of the optimality conditions of this problem, that when there are enough assets, so that markets are complete with respect to aggregate risk, the following risk-sharing condition must hold to first order for all realizations of aggregate shocks:

$$\frac{\mathbb{E}\left[u'\left(c_{0,i}\right)\right]}{\mathbb{E}\left[u'\left(c_{ss,i}\right)\right]} = \lambda_0 \qquad \forall i \tag{1}$$

Here, $\mathbb{E}\left[\cdot\right]$ takes expectations over the realizations of idiosyncratic risk, ss denotes the steady state with no aggregate risk, and the index 0 represents any possible realization of aggregate shocks.

Equation (1) says that agents' marginal utilities must be expected, at the time of the portfolio choice, to all increase or decline in proportion conditional on any shock. This equation can be used to test for the optimality of exogenous portfolios, and, when this test fails, to solve for the portfolios and the λ_0 that imply (1) while clearing all asset markets.[2] We show that this method boils

[1] While the literature has emphasized the endogenous choice of asset holdings in accounts of differing liquidity (following Kaplan and Violante 2014), it has taken the mix of assets held in these accounts as a given.

[2] In fact, in this complete-markets case, it is possible and more natural to circumvent solving for portfolios altogether, and instead directly solve for the optimal exposure of agents to shocks.

down to a simple modification of sequence-space Jacobians to account for the effect of holding optimal rather than exogenous portfolios. We also show that the computation of these modified sequence-space Jacobians involves the same objects as those used for exogenous portfolios. The computation is therefore immediately implementable on any model solved with the sequence-space Jacobian method, so that the consequences of endogenous portfolios can be examined in a straightforward way.

We apply our method to a simple HANK model in the spirit of Auclert et al. (2023a). Agents face idiosyncratic income risk and borrow and save in a single account subject to a borrowing constraint. Production is linear in labor, there are sticky wages with equal rationing in the labor market, and product markets are monopolistically competitive with flexible prices. Agents have log utility over consumption and can trade two types of assets: real risk-free government bonds, and shares that are claims to firms' profits. The central bank sets the real interest rate exogenously. Fiscal policy uses a proportional tax on both labor and capital, and chooses the time paths of government spending and tax revenue subject to an intertemporal budget constraint. We calibrate the model so that there are no government bonds in the initial steady state. Further, in our baseline exogenous-portfolio calibration, we assume that all agents hold 100% stock portfolios. We initially consider fiscal and monetary policy shocks in isolation, so that there is always a single shock. Given that there are two assets, markets are complete with respect to aggregate risk, and we can use our core method to jointly solve for the portfolios and the impulse responses to these shocks.

When the single shock is a balanced-budget government spending shock, we show that the impulse responses under endogenous portfolios are the same as under exogenous portfolios. The reason is that with these shocks, and a constant real interest rate, there is a unit government spending multiplier on output, with no effect on consumption for any agent (Haavelmo 1945, Auclert et al. 2023a). Therefore, condition (1) is satisfied conditional on this shock, with $\lambda_0 = 1$. In fact, since post-tax dividends are unaffected by the increased production, the stock return equals the bond return even in the period of the shock. This implies that, even though they perceive the aggregate risk, agents here are still indifferent between all portfolios.

When the single shock is a monetary policy shock, we show that again, the impulse responses under endogenous portfolios are the same as under exogenous portfolios. Here, however, portfolios are determinate, since the stock return falls on impact relative to the bond return when monetary policy tightens. Nevertheless, it turns out that in our baseline model, a 100% stock portfolio allocation is optimal for all agents. The reason is that, under our assumptions about preferences (in particular log utility), technology (equal rationing), and with 100% stock portfolios, in equilibrium the Euler equation $dc_{it} = -c_{it} \sum_{s \geq 0} \frac{dr_{t+s}}{1+r}$ is satisfied for every agent i—a result initially derived by Werning (2015). Hence, uniform 100% stock portfolios effectively imply that all agents are equally exposed to monetary policy shocks, so that these portfolios achieve the complete markets allocation. Alternative portfolio choices would be suboptimal for agents: given the equilibrium risk premium on stocks, increasing or reducing exposure to the stock market would lower their expected utility.

Finally, when the single shock is a deficit-financed tax cut, the impulse responses under endogenous portfolios differ significantly from those under exogenous portfolios. When portfolios are exogenous, deficit-financed tax cuts have large and persistent effects on economic activity in HANK, echoing earlier results in the literature (e.g. Auclert et al. 2023a, Bilbiie 2021, Angeletos, Lian and Wolf 2023, Auclert, Rognlie and Straub 2023b). For instance, in our baseline calibration, the impact transfer multiplier is 0.2, and the effect persists for about five years, with a cumulative multiplier of around 0.77. By contrast, under endogenous portfolios, the impact transfer multiplier is only 0.08 (60% less than under exogenous portfolios), though the effect persists for roughly the same amount of time, so that the cumulative multiplier is still 0.53 (30% less than under exogenous portoflios).

The intuition for these results is as follows. With 100% stock portfolios, deficit-financed fiscal transfers disproportionately raise the consumption of poor agents, and therefore disproportionately lower their expected marginal utility, violating condition (1). Since the stock market is booming, optimal portfolios reduce the stock market exposure of poor agents and raise the exposure of rich agents—and since rich agents have lower marginal propensity to consume out of capital gains than poor agents, these portfolios reduce the aggregate transfer multiplier. The effect of this additional redistribution is large on impact, explaining the large reduction in the transfer multiplier. But it is also relatively short-lived, so that the overall persistence of the output effect is less affected by endogenous portfolios.

Our core results described so far assume complete markets with respect to aggregate risk—there must be at least one more asset than there are shocks in the model—and that agents face no portfolio constraints, only a constraint on the total value of assets. Intuitively, these assumptions let agents maximally hedge aggregate shocks, so they deliver an upper bound on the impact of endogenous portfolios relative to exogenous portfolios. In two extensions, we relax these assumptions in turn, and evaluate the quantitative effect of portfolio constraints and incomplete markets in HANK.

We first consider cases with incomplete markets, where there are more shocks than assets available. This extension is important, because in practice the set of shocks may be large relative to the typical assets available to households. We show that the solution in this case can be computed using a projection of the complete-market exposures on the column space of the return matrix. This has two consequences. First, impulse responses to different shocks are coupled, since the presence of one shock influences the portfolios and therefore the impulse response to other shocks.[3] Second, the solution involves solving a simple fixed point problem, where the return matrix depends on impulse responses and vice-versa. We apply this method to solve for the effect of deficit financed shocks in our baseline model, assuming that monetary policy shocks are also present. We find that the effects are nearly the same as with exogenous portfolios. This is because monetary policy shocks are easier to hedge than deficit-financed shocks as they move returns much more,

[3] The sequence-space solution therefore requires solving jointly for the effect of all shocks, which can require solving a very large system when the number of shocks is large. We will discuss an iterative approach to facilitate this.

so optimal portfolios under incomplete markets are close to the 100% exogenous stock portfolios.

We next introduce asset-specific constraints, rather than a simple constraint on total net worth. This extension is also important, because the complete-market portfolios that hedge deficit-financed shocks in our baseline calibration are highly implausible: they imply that agents near the borrowing constraints take extreme short positions in the stock market—hundreds of thousands of times their net worth—to invest in the bond market. We show that solving for this requires another, more complex fixed-point problem, where the return matrix and the multipliers on constraints must be solved jointly with impulse responses. Implementing this procedure in our baseline model with only deficit-financed shocks, and constraining agents to hold no less than -100% and no more than 200% of their net worth in equities—i.e. banning large short and leveraged long positions—we find that the impact transfer multiplier goes back to 0.16 and the cumulative to 0.61, very close to the exogenous-portfolio result.

We conclude that allowing for hedging of aggregate shocks can, in some cases, attenuate the importance of the effects highlighted in the HANK literature with exogenous portfolios. However, allowing for realistic short-sale constraints or for a large number of shocks relative to the number of assets tends to restore more typical outcomes. For instance, we study asset portfolios when agents can hold nominal in addition to real bonds, or long-term in addition to short-term bonds. These assets do provide additional hedging opportunities by adding independent variation in returns. However, once we scale up the number of shocks by allowing for deficit-financed shocks to have different degrees of persistence, the baseline transfer multiplier moves back to its value with exogenous portfolios.

Our results have implications for asset pricing. The λ_0 in equation (1) has the interpretation of a cross-sectional stochastic discount factor. We show that in fact, even though it can be computed using the first-order solution with optimal portfolios, λ_0 delivers the relative risk premia for different assets up to second order.[4] We show, in our applications, that the risk premia are very small: they are zero for the balanced-budget government spending shock, equal to the usual consumption-CAPM formula for the monetary-policy shock, and given by a virtual-consumption-CAPM formula for the deficit-financed fiscal shock. In our calibration, we find risk premia on stocks vs. bonds of 25 annual basis points for monetary policy and essentially zero for fiscal policy. This is not surprising given that agents have standard CRRA utility with risk aversion equal to 1 in our baseline calibration, so that our baseline model is subject to a standard equity premium puzzle. The asset pricing literature has pointed out that cyclical dynamics in idiosyncratic risk can raise the equity premium to empirically reasonable levels (Mankiw 1986, Constantinides and Duffie 1996, Storesletten, Telmer and Yaron 2004, Storesletten, Telmer and Yaron 2007). While our baseline model assumes away these effects, our method should easily be applicable to models that accommodate them, enabling the study of heterogeneity in models with empirically realistic risk premia.

[4] Solving for the absolute level of risk premia, on the other hand, requires a second order solution, since this is affected by precautionary savings with respect to aggregate risk.

The assumption of exogenous portfolios is made by a majority of papers in the heterogeneous-agent literature, including almost all of the HANK literature.[5] In the international macroeconomics literature, which has long studied heterogeneous agents in the form of different countries, the question of optimal portfolios has been studied in the context of cross-country risk sharing and international diversification (e.g. Backus and Smith 1993, Baxter and Jermann 1997). Devereux and Sutherland (2011) and Tille and van Wincoop (2010) provide solution methods to obtain portfolios in multi-country models, which can be implemented in state-space solvers such as Dynare. Our method builds on the ideas developed in these papers, but is aimed at models with large amounts of heterogeneity, where state-space methods are generally intractable.

Bhandari, Bourany, Evans and Golosov (2023) develop the theory of second-order perturbations for heterogeneous-agent models in the state space, with an application to portfolio choice among many other topics. The two methods should be equivalent and deliver the same solution. The distinctive feature of our approach is that we develop it in the sequence space, and that we do not need to solve for the second-order perturbation solution to obtain the zeroth-order portfolios or the impulse responses with these portfolios. Instead, we leverage the implications of a very general second-order perturbation problem with heterogeneous agents to develop a sequence-space method that is fast and efficient.

The paper proceeds as follows. Section 2 provides a simple second-order perturbation theory for a heterogeneous-agent problem. It shows how complete markets deliver equation (1) and discusses risk premia and the case with incomplete markets. Section 3 introduces our simple HANK model, and shows how to apply the results from section 2 to this setting to obtain impulse responses with endogenous portfolios using a simple correction to the sequence-space Jacobian. Section 4 studies the effects of monetary and fiscal policy shocks in the baseline model, under exogenous and endogenous portfolios. Section 5 considers extensions to portfolio constraints, incomplete markets, and additional types of assets and shocks. Section 7 concludes.

2 Heterogeneous-agent portfolios and risk premia

This section introduces a general, static portfolio problem with heterogeneous agents, and derives the restrictions imposed by a second-order perturbation for the relationship between marginal utilities and returns, asset portfolios, and risk premia. The results developed here are related to those obtained in the static incomplete markets general equilibrium literature (Magill and Quinzii 2008), as well as to second-order perturbation theory in DSGE models (Schmitt-Grohé and Uribe 2004, Devereux and Sutherland 2011, Tille and van Wincoop 2010). Relative to this literature, our emphasis is on the restrictions imposed by optimality with an arbitrary number of agents. We also derive a solution approach which, to the best of our knowledge, is new. Later, we will apply this static problem to date -1 of a heterogeneous-agent model when shocks realize only at date 0, and

[5] In the heterogeneous-agent literature more broadly, the Krusell-Smith method is sometimes used to solve for portfolios. Krusell and Smith (1997) provide an early example of this approach. An alternative is global solution methods, as in Guvenen (2009).

argue that this delivers the correct zeroth-order portfolios.

2.1 Setting and perturbation

Our model has heterogeneous households indexed by i. Each household i has wealth a_i that it can allocate between $K+1$ assets a_i^k, for $k = 0, \ldots, K$. Each asset k has total supply A^k, price p^k, and yields stochastic payoffs $x^k(\boldsymbol{\epsilon})$, where $\boldsymbol{\epsilon} \equiv (\epsilon_1, \ldots, \epsilon_Z)'$ is a vector of Z shocks.

We consider a perturbation that scales the variances of these shocks by a common factor σ. We therefore write $\epsilon_z = \sigma \bar{\epsilon}_z$, and assume that the primitive $\bar{\epsilon}_z$ are independent and symmetrically distributed, with mean 0 and variance $\overline{\sigma}_z^2$. Hence, we have $\mathbb{E}[\boldsymbol{\epsilon}] = \mathbf{0}$ and $\mathbb{E}[\boldsymbol{\epsilon}\boldsymbol{\epsilon}'] = \sigma^2 \boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma} = \text{diag}\left(\overline{\sigma}_1^2, \ldots \overline{\sigma}_Z^2\right)$ is the $Z \times Z$ matrix of primitive shock variances.

The objective of household i is to maximize the expected value function W_i, which depends on incoming wealth next period $\sum_{k=0}^K x^k(\boldsymbol{\epsilon}) a_i^k$, and also directly on the realization of shocks $\boldsymbol{\epsilon}$. Hence, i's portfolio choice problem is:

$$\begin{aligned} \max_{\{a_i^k\}} \quad & \mathbb{E}\left[W_i\left(\sum_{k=0}^K x^k(\boldsymbol{\epsilon}) a_i^k, \boldsymbol{\epsilon}\right)\right] \\ \text{s.t.} \quad & \sum_{k=0}^K p^k a_i^k = a_i \end{aligned} \tag{2}$$

Denoting the Lagrange multiplier on i's budget constraint by γ_i, this problem has the following classic first-order conditions:

$$\mathbb{E}\left[x^k(\boldsymbol{\epsilon}) W_i'\left(\sum_{k=0}^K x^k(\boldsymbol{\epsilon}) a_i^k, \boldsymbol{\epsilon}\right)\right] = \gamma_i p^k \quad \forall i, k \tag{3}$$

which must hold for every i and for every k. Writing di for the distribution of agents i, market clearing in all asset markets imposes:

$$\int a_i^k di = A^k \quad \forall k \tag{4}$$

Given primitives a_i and W_i, as well as the parameter σ, equilibrium is a set of prices for each asset p^k and Lagrange multipliers for each agent γ_i, such that the optimality conditions (3) are satisfied for each (i, k) pair, and all asset markets clear, i.e. (4) holds for all k.

We now work out the implications of these equations for a perturbation in σ up to the second order. We write $p^k(\sigma)$, $\gamma_i(\sigma)$ for the solution at a given σ and study their second-order Taylor expansion around $\sigma = 0$. We note that, given that the distribution of $\boldsymbol{\epsilon}$ is symmetric, these must be even functions of σ : $p^k(-\sigma) = p^k(\sigma)$ and $\gamma_i(-\sigma) = \gamma_i(\sigma)$. This implies, in particular, that $\frac{d\gamma_i}{d\sigma} = \frac{dp^k}{d\sigma} = 0$, a result that we will use several times below.

Zero-th and first-order perturbation. Applying (3) at $\sigma = 0$, we find $\gamma_i / W_i' = x^k / p^k$ for all i and k, where p^k stands for $p^k(0)$, γ_i for $\gamma_i(0)$, x^k for $x^k(\mathbf{0})$, and W_i' for $W_i'\left(\sum_{k=0}^K x^k a_i^k, \mathbf{0}\right)$. Hence, the returns on all assets must equal a common constant R, and this is also the rate entering the Euler equation of all agents:

$$\gamma_i / W_i' = x^k / p^k = R \tag{5}$$

In particular, $\sum_{k=0}^K x^k a_i^k$ is also just $R \sum_{k=0}^K p^k a_i^k = R a_i$. Equation (5) gives the usual result that, with no aggregate uncertainty, all assets must have equal returns.

Next, differentiating (3) with respect to σ gives us

$$\mathbb{E}\left[\frac{dx^k}{d\sigma} W_i' + x^k \frac{dW_i'}{d\sigma}\right] = \frac{d\gamma_i}{d\sigma} p^k + \gamma_i \frac{dp^k}{d\sigma} \tag{6}$$

Given the definition $x^k(\boldsymbol{\epsilon}) = x^k(\sigma\bar{\epsilon}_1, \ldots, \sigma\bar{\epsilon}_Z)$, and $W_i\left(\sum_{k=0}^K x^k(\sigma\bar{\boldsymbol{\epsilon}}) a_i^k, \sigma\bar{\boldsymbol{\epsilon}}\right)$, we have that

$$\frac{dx^k}{d\sigma} = \sum_{z=1}^{Z} \frac{\partial x^k}{\partial \epsilon_z} \bar{\epsilon}_z \quad \text{and} \quad \frac{dW_i'}{d\sigma} = \sum_{z=1}^{Z} \frac{dW_i'}{d\epsilon_z} \bar{\epsilon}_z \tag{7}$$

where we have defined the total derivative of W_i' with respect to ϵ_z as

$$\frac{dW_i'}{d\epsilon_z} \equiv W_i'' \sum_{k=0}^{K} \frac{\partial x^k}{\partial \epsilon_z} a_i^k + \frac{\partial W_i'}{\partial \epsilon_z} \tag{8}$$

reflecting the direct dependence of the value function on the shock ϵ_z, as well as its indirect dependence through the effect of asset returns. Since $\mathbb{E}[\bar{\epsilon}_z] = 0$, applying (7), we see that the left-hand side of (6) is zero. The right-hand side of (6) is also zero, given our symmetry result above, so equation (6) holds regardless of portfolios.

Second-order perturbation. Now, differentiating (6) with respect to σ gives us:

$$\mathbb{E}\left[\frac{d^2x^k}{d\sigma^2}\right] W_i' + 2\mathbb{E}\left[\frac{dx^k}{d\sigma}\frac{dW_i'}{d\sigma}\right] + x^k \mathbb{E}\left[\frac{d^2 W_i'}{d\sigma^2}\right] = \frac{d^2\gamma_i}{d\sigma^2} p^k + 2\frac{d\gamma_i}{d\sigma}\frac{dp^k}{d\sigma} + \gamma_i \frac{d^2 p^k}{d\sigma^2}$$

After dividing by $x^k W_i' = \gamma_i p^k$ from (5) on both sides, we obtain:

$$\mathbb{E}\left[\frac{d^2x^k/x^k}{d\sigma^2}\right] + 2\mathbb{E}\left[\frac{dx^k/x^k}{d\sigma}\frac{dW_i'/W_i'}{d\sigma}\right] + \mathbb{E}\left[\frac{d^2 W_i'/W_i'}{d\sigma^2}\right] = \frac{d^2\gamma_i/\gamma_i}{d\sigma^2} + \frac{d^2 p^k/p^k}{d\sigma^2}$$

We can rewrite this as simply

$$\mathbb{E}\left[\frac{dx^k/x^k}{d\sigma}\frac{dW_i'/W_i'}{d\sigma}\right] = \alpha_i + \beta^k \tag{9}$$

where α_i, which depends only on the household i, and β^k, which depends only on the asset k, are defined as

$$\alpha_i \equiv \frac{1}{2}\left(\frac{d^2\gamma_i/\gamma_i}{d\sigma^2} - \mathbb{E}\left[\frac{d^2 W_i'/W_i'}{d\sigma^2}\right]\right)$$
$$\beta^k \equiv \frac{1}{2}\left(\frac{d^2 p^k/p^k}{d\sigma^2} - \mathbb{E}\left[\frac{d^2 x^k/x^k}{d\sigma^2}\right]\right)$$

Using again (7), and the fact that $\mathbb{E}\left[\overline{\boldsymbol{\epsilon}}\overline{\boldsymbol{\epsilon}}'\right] = \boldsymbol{\Sigma}$, we can rewrite (9) as:

$$\sum_{z=1}^{Z} \frac{\partial x^k/x^k}{\partial \epsilon_z} \frac{dW_i'/W_i'}{d\epsilon_z} \overline{\sigma}_z^2 = \alpha_i + \beta^k \qquad \forall i, k \tag{10}$$

We note that this applies to the product of two first derivatives, and therefore, intuitively, places restrictions on the relationship between the impulse response of returns and marginal utilities. Finally, using (10) for asset k relative to asset 0 (where we note that 0 could correspond to any reference asset in the economy), we obtain:

$$\sum_{z=1}^{Z} \left(\frac{\partial x^k/x^k}{\partial \epsilon_z} - \frac{\partial x^0/x^0}{\partial \epsilon_z}\right) \frac{dW_i'/W_i'}{d\epsilon_z} \overline{\sigma}_z^2 = \beta^k - \beta^0 \equiv b^k \qquad \forall i \tag{11}$$

Equation (10) says that all households equalize their average sensitivity to shocks z, interacted with the relative returns on asset k, to a k-specific term b^k. We will soon see that this term has the interpretation of a relative risk premium on asset k. Stacking $\mathbf{b} \equiv \left(b^1, \ldots, b^K\right)'$ as a $K \times 1$ vector of relative risk premia, $\boldsymbol{\lambda}_i \equiv \left(\frac{dW_i'/W_i'}{d\epsilon_1}, \ldots, \frac{dW_i'/W_i'}{d\epsilon_Z}\right)'$ as a $Z \times 1$ vector of sensitivities of marginal utility to each shock, and defining the $Z \times K$ matrix $\mathbf{X}$ with elements equal to the relative returns of each asset to each shock $X_{zk} \equiv \frac{\partial x^k/x^k}{\partial \epsilon_z} - \frac{\partial x^0/x^0}{\partial \epsilon_z}$, equation (11) becomes:

$$\mathbf{X}'\boldsymbol{\Sigma}\boldsymbol{\lambda}_i = \mathbf{b} \quad \forall i \tag{12}$$

2.2 Complete markets

Suppose that $K = Z$; in other words, that the number of assets equals the number of shocks plus 1. Then $\mathbf{X}$ is a square matrix. Additionally, suppose the following assumption is satisfied:

Assumption 1 (Spanning)**.** *The rows of* $\mathbf{X}$ *are linearly independent.*

Assumption 1 says that the relative returns across assets vary sufficiently across shocks. Under these two assumptions, the $Z \times Z$ matrix $\mathbf{X}'\boldsymbol{\Sigma}$ is invertible. Condition (12) can therefore be rewritten:

$$\boldsymbol{\lambda}_i = (\mathbf{X}')^{-1}\boldsymbol{\Sigma}^{-1}\mathbf{b} \qquad (\equiv \boldsymbol{\lambda})$$

This leads us to our first main result.

Proposition 1. *Suppose that $K = Z$ and assumption holds. Then for each shock z, there exists a λ_z such that*

$$\frac{dW_i'/W_i'}{d\epsilon_z} = \lambda_z \qquad \forall i \tag{13}$$

To connect this to a typical setting with heterogeneous agents, suppose that agent i has utility function u_i over consumption c_i in the next period, where c_i equals asset income $\sum_{k=0}^{K} x^k(\boldsymbol{\epsilon})\, a_i^k$ plus labor income $y_i(e')$, which depends on a realization of idiosyncratic risk e'. In this case, we have $W_i\left(\sum_{k=0}^{K} x^k(\boldsymbol{\epsilon})\, a_i^k, \boldsymbol{\epsilon}\right) = \mathbb{E}\left[u_i\left(\sum_{k=0}^{K} x^k(\boldsymbol{\epsilon})\, a_i^k + y(e'), \boldsymbol{\epsilon}\right)\right]$, and therefore (13) reads

$$\frac{d\mathbb{E}\left[u_i'(c_i)\right]/\mathbb{E}\left[u_i'(c_i)\right]}{d\epsilon_z} = \lambda_z \qquad \forall i$$

which corresponds to equation (1) in the introduction.

Proposition 1 provides us with a simple test of portfolio optimality in a setting where $K = Z$. To understand the test, note that standard first-order methods allow us relatively easily to solve for steady-state x^k, W^i, as well as $\frac{\partial x^k}{\partial \epsilon_z}$ and $\frac{dW_i'}{d\epsilon_z}$ for given shocks z, conditional on given incoming portfolios $\{a_i^k\}$ for all agents. With these objects, one can form the matrix of relative returns $\mathbf{X}$ to test if the spanning assumption 1 is satisfied, and then test whether $\frac{dW_i'/W_i'}{d\epsilon_z}$ are equalized across agents i for all shocks z. If so, proposition 1 tells us that the portfolios are optimal.

Proposition 1 also implies a method for solving for optimal portfolios directly. This works as follows. Recalling that $W_i\left(\sum_{k=0}^{K} x^k(\boldsymbol{\epsilon})\, a_i^k, \boldsymbol{\epsilon}\right)$, the derivative with respect to ϵ_z has two terms: the direct dependence on ϵ_z and the indirect dependence through returns. More specifically, we have from (8) that:

$$\begin{aligned} \frac{dW_i'/W_i}{d\epsilon_z} &= \frac{W_i''}{W_i'} \sum_{k=0}^{K} \frac{\partial x^k}{\partial \epsilon_z} a_i^k + \frac{\partial W_i'/W_i'}{\partial \epsilon_z} \\ &= R\frac{W_i''}{W_i'} \sum_{k=0}^{K} \frac{\partial x^k/x^k}{\partial \epsilon_z} p^k a_i^k + \frac{\partial W_i'/W_i'}{\partial \epsilon_z} \\ &\equiv R\frac{W_i''}{W_i'} \frac{ds_i}{d\epsilon_z} + \frac{\partial W_i'/W_i'}{\partial \epsilon_z} \end{aligned}$$

where the second equality follows from $p^k = x^k/R$ from (5), and the third defines $\frac{ds_i}{d\epsilon_z} \equiv \sum_{k=0}^{K} \frac{\partial x^k/x^k}{\partial \epsilon_z} p^k a_i^k$ as the sensitivity of i's payoff to z given the asset portfolio $p^k a_i^k$. When assumption 1 is satisfied, proposition 1 shows that the left-hand side is equalized to λ_z for each i. We can use this to directly solve for the optimal sensitivity of agent i's payoff to the shock z:

$$\frac{ds_i}{d\epsilon_z} = \frac{W_i'}{RW_i''}\left(\lambda_z - \frac{\partial W_i'/W_i'}{\partial \epsilon_z}\right) \tag{14}$$

Integrating across agents using the definition of $\frac{ds_i}{d\epsilon_z}$, and imposing market clearing (4), we have the additional following equations that determine the λ_z:

$$\sum_{k=0}^{K} \frac{\partial x^k / x^k}{\partial \epsilon_z} p^k A^k = \left(\int \frac{W_i'}{RW_i''} di \right) \lambda_z - \int \frac{W_i'}{RW_i''} \frac{\partial W_i' / W_i'}{\partial \epsilon_z} di \quad \forall z \tag{15}$$

Equations (14) and (15) are sufficient to solve for the allocation with optimal portfolios, using only information from the steady-state (x^k, p^k, R, W_i, W_i', W_i'') and the first-order perturbation to shock z ($\frac{\partial x^k}{\partial \epsilon_z}$, $\frac{\partial W_i'}{\partial \epsilon_z}$). This is because the relevant information for the first-order solution is actually only the sensitivities $\frac{ds_i}{d\epsilon_z}$, rather than the portfolios themselves. However, once these sensivities are known for each agent i, it is straightforward to back out the portfolios that sustain them. Specifically, given that

$$\frac{ds_i}{d\epsilon_z} \equiv \left(\sum_{k=0}^{K} X_{zk} \frac{p^k a_i^k}{a_i} + \frac{\partial x^0 / x^0}{\partial \epsilon_z} R \right) a_i \quad \forall z$$

if we define $\omega_i^k \equiv \frac{p^k a_i^k}{a_i}$ is the share of agent i's portfolio in asset k, we have that $\sum_{k=0}^{K} X_{zk} \omega_i^k a_i = t_{iz}$ for each z, where

$$t_{iz} \equiv \frac{ds_i}{d\epsilon_z} - R a_i \frac{\partial x^0 / x^0}{\partial \epsilon_z} \tag{16}$$

is the optimal sensitivity of i's payoff to z net of the sensitivity of the portfolio when all wealth is invested in asset 0. In matrix form, this can be written as:

$$\mathbf{X} \boldsymbol{\omega}_i a_i = \mathbf{t}_i \tag{17}$$

where $\boldsymbol{\omega}_i = \left(\omega_i^0, \ldots, \omega_i^K\right)'$ is the vector giving the portfolio of agent i, and $\mathbf{t}_i = \left(t_{i0}, \ldots, t_{iZ}\right)'$ is the vector of t's, which can be interpreted as transfers made by agents to each other over and above what they get from investing in asset 0. Since $\mathbf{X}$ is invertible, this implies that, provided $a_i \neq 0$, the porfolios are given by:

$$\boldsymbol{\omega}_i = \mathbf{X}^{-1} \frac{\mathbf{t}_i}{a_i} \tag{18}$$

Note that the amounts invested in each asset $p^k a_i^k$ are well-defined even when agents have zero wealth, $a_i = 0$. Given this, we expect portfolio shares to diverge for agents with wealth close to 0, a point that will be apparent in our application of section 4.

Case with $K > Z$. When $K > Z$, then assuming that the spanning condition 1 is still satisfied, markets remain complete with respect to aggregate shocks: Proposition 1 continues to hold, and one can still use (14)–(15) to solve for optimal sensitivities. However, portfolio choice is undetermined, as any portfolio $\boldsymbol{\omega}_i$ satisfying $\mathbf{X}\boldsymbol{\omega}_i = \mathbf{t}_i$ is optimal for agent i (there are $K - Z$ dimensions of indeterminacy.) One can resolve this indeterminacy by forcing $K - Z$ portfolio shares to be fixed, or through some other device.

2.3 Risk premia

Earlier we previewed an interpretation of the terms b^k as relative risk premia on assets. In this section, we see why this interpretation is justified. Define $R^k(\sigma) \equiv \mathbb{E}\left[x^k(\sigma\bar{\epsilon})\right]/p^k(\sigma)$ as the expected return on asset k. Note that $R^k(0) = R$, $\frac{dR^k(\sigma)/R}{d\sigma} \equiv 0$ by the arguments developed in the first-order perturbation, and that, by definition of β^k, we have

$$\frac{d^2 R^k/R}{d\sigma^2} = \mathbb{E}\left[\frac{d^2 x^k/x^k}{d\sigma^2}\right] - \frac{d^2 p^k/p^k}{d\sigma^2} = -2\beta^k$$

Hence, the second order expansion of R^k in σ is given by

$$R^k(\sigma) \approx R - R\beta^k\sigma^2$$

In particular, the relative risk premium on asset k vs. asset 0 has second-order expansion

$$\frac{R^k(\sigma) - R^0(\sigma)}{R} \approx -\left(\beta^k - \beta^0\right)\sigma^2 = -b^k\sigma^2$$

Combining equation (11) with equation (13), we obtain the following result:

Proposition 2. *Under complete markets, the risk premia on asset k relative to asset 0 satisfy, to second order in* σ,

$$\frac{R^k(\sigma) - R^0(\sigma)}{R} \approx -\sum_{z=1}^{Z} X_{zk}\lambda_z\bar{\sigma}_z^2\sigma^2 \tag{19}$$

Proposition 2 shows that λ_z has the interpretation of a relative stochastic discount factor. The textbook formula for asset returns (eg Campbell 2003, Cochrane 2005) says that the expected excess return on any asset relative to the risk-free rate is

$$\frac{\mathbb{E}\left[R^k\right] - R^f}{R^f} = -\text{Cov}\left(m, R^k\right)$$

where m is the stochastic discount factor. This implies in particular that, for any two assets k and 0, the relative return is

$$\frac{\mathbb{E}\left[R^k\right] - \mathbb{E}\left[R^0\right]}{R^f} = -\text{Cov}\left(m, R^k - R^0\right)$$

Equation (19) has the same content when replacing m by λ, but it only works in the limit of $\sigma^2 \to 0$, and the denominator involves the common steady-state return rather than the risk-free rate. The level of the risk-free rate is $R^f(\sigma) - R \approx -R\beta^f\sigma^2$ to second order, which requires knowledge of the level of β^f (capturing the effect of precautionary savings). This, in turn, requires a full second-order solution to the problem. In contrast, proposition 2 shows that we can obtain relative risk premia and λ using only the information from a first-order perturbation.

2.4 Incomplete markets

Suppose now that $K < Z$. The optimality condition is still given by (12), but in general this does not imply a common $\boldsymbol{\lambda}_i$ for all households. We know from (17), however, that the effective transfers $\mathbf{t}_i$ must lie in the column space of the relative return matrix $\mathbf{X}$.

Proposition 3. *Suppose that $K < Z$ and that the columns of $\mathbf{X}$ are linearly independent. Let $\mathbf{t}_i^{CM}$ be the transfer vector implied by complete markets, obtained from solving (14)–(15) and forming (16). Let*

$$\mathbf{t}_i = \mathbf{X}(\mathbf{X}'\boldsymbol{\Sigma}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}\mathbf{t}_i^{CM} \tag{20}$$

be the weighted projection of $\mathbf{t}_i^{CM}$ on the column space of $\mathbf{X}$. Then, $\mathbf{t}_i$ gives the sensitivities of agent i payoffs at the optimal portfolios $\boldsymbol{\omega}_i$, and these portfolios are given by $\mathbf{X}\boldsymbol{\omega}_i a_i = \mathbf{t}_i$. Moreover, the risk premia under incomplete markets b^k are the same as in the complete markets allocation.

Proof. See appendix. □

Proposition 3 suggests a very simple approach to solve for any given incomplete markets allocation, which we will implement in practice in section 3.2.

2.5 Portfolio constraints

Suppose that we augment our original problem (2) with additional linear portfolio constraints $l = 1, \ldots, L$

$$\sum_{k=0}^{K} \theta_l^k p^k a_i^k \leq 0$$

where θ_l^k is the loading of constraint l on asset k. For instance, if asset 1 is equity, then for a constraint preventing any short positions in equity, we can write $\theta_l^1 = -1$ and $\theta_l^k = 0$ for $k \neq 1$. If asset 0 is bonds, then for a constraint limiting borrowing to a certain fraction α of equity, we can write $\theta_l^0 = -1$ and $\theta_l^1 = -\alpha$.

Letting η_{il} be the Lagrange multiplier on constraint l for household i, our first-order condition (3) gains a new term $\sum_l \eta_{il}\theta_l^k p^k$ on the right. Since households are indifferent between portfolios to first order, the Lagrange multipliers are zero to first order as well. The second-order condition (10), however, gains a term $\sum_l \frac{d^2\eta_{il}/\gamma_i}{d\sigma^2}\theta_l^k$ on the right. Once we subtract $k = 0$ from both sides, this term becomes

$$\sum_l \frac{d^2\chi_{il}/\gamma_i}{d\sigma^2}(\theta_l^k - \theta_l^0) \tag{21}$$

on the right of equation (11) . Stacking objects as before, and now defining the matrix $\boldsymbol{\Theta}$ by $\Theta_{lk} \equiv \theta_l^k - \theta_l^0$ and the vector $\boldsymbol{\eta}_i$ by $\eta_{il} \equiv \frac{d^2\chi_{il}/\gamma_i}{d\sigma^2}$, equation (12) is

$$\mathbf{X}'\boldsymbol{\Sigma}\boldsymbol{\lambda}_i = \mathbf{b} + \boldsymbol{\Theta}'\boldsymbol{\eta}_i \quad \forall i \tag{22}$$

The new term $\boldsymbol{\Theta}'\boldsymbol{\eta}_i$ reflects the shadow value of constraints. By complementary slackness, we will have $\eta_{il} > 0$ only if constraint l is binding for household i. Together with the stacked constraint $\boldsymbol{\Theta}\boldsymbol{\omega}_i a_i \leq 0$, these conditions characterize portfolios.

The nonlinearity introduced by possibly-binding portfolio constraints makes solving this problem more difficult than in the previous cases. A simple strategy is to proceed iteratively, first calculating the portfolios as before (ignoring $\boldsymbol{\eta}_i$), then enforcing constraints for portfolios that violate them and re-solving (22) to clear markets given the remaining degrees of freedom, then rechecking whether constraints bind, and so on. We discuss embedding such an iterative strategy in general equilibrium in section 3.2.

2.6 Connection with second-order perturbation of a dynamic problem

The arguments developed above carry over to a fully dynamic setting, where there are recurring shocks. In this setting, W_i and x^k can depend directly on σ as well, reflecting the fact that the volatility of future shocks can affect the value function and returns. Note, however, that the first derivatives of these with respect to σ are still zero, by the symmetry argument discussed in section 2.1. Also, we must allow for a shifter to the value function $\eta = \sigma\bar{\eta}$ that captures the effect of past shocks. The appendix shows that the equations derived above remain the correct equations to characterize zeroth-order portfolios and risk premia in this case.

3 Exogenous vs endogenous portfolios in a simple HANK model

This section sets up a heterogeneous-agent New Keynesian model in the spirit of the literature. We keep the model simple on purpose to simplify the mechanisms at play and facilitate a discussion of our method for solving the model with endogenous portfolios. However, we argue that the results we obtain from the addition of endogenous portfolios are likely to apply to other models in the literature. This is because of two features that are widely shared in the literature. First, the aggregate effects of monetary policy are not too different from those of a representative-agent model, though with very different transmission channels (Werning 2015, Kaplan et al. 2018). In fact, in our baseline calibration the result of Werning (2015) applies exactly and the effects of monetary policy are identical to those of a representative-agent model. Second, deficit-financed fiscal policy has very powerful effects on economic activity, due to the intertemporal Keynesian cross effect studied in Auclert et al. (2023a).

3.1 The simple HANK model

We consider a set of households indexed by i, facing idiosyncratic and aggregate risk. Idiosyncratic risk is to their efficiency in labor e_{it}, which follows a discrete Markov chain with transition matrix Π. Aggregate risk is to monetary and fiscal policy.

In each period t, households face a real wage w_t per unit of efficient labor, a proportional tax rate τ_t on their labor earnings, and work n_{it} hours, determined by labor demand as described below. They attempt to smooth consumption over time, and have access to two assets: stocks s_{it}, which have price p_t and pay a dividend d_t each period, and real short-term bonds b_{it}, which pay a return r_{t-1} determined as of the previous period. Households attempt to smooth consumption over time in the face of risk, subject to a constraint that their net worth $p_t s_{it} + b_{it}$ cannot be negative. Hence, they solve the following program:

$$\begin{aligned} \max \ & \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_{it}) \\ & c_{it} + p_t s_{it} + b_{it} \leq (p_t + d_t) s_{it-1} + (1 + r_{t-1}) b_{it-1} + e_{it} (1 - \tau_t) w_t n_{it} \\ & p_t s_{it} + b_{it} \geq 0 \end{aligned}$$

Production is from labor with unitary productivity, $Y_t = N_t$. There is monopolistic competition in the goods market. Households consume the goods produced by a continuum of firms j, with c_{it} a CES aggregate of intermediate goods with elasticity of substitution $\frac{\mu}{\mu-1}$. Prices are perfectly flexible, so firms set their prices at markup μ over the nominal wage, resulting in a real wage of $w_t = \frac{1}{\mu}$. Firms' dividends are also taxed at rate τ_t. Hence, their aggregate dividends are $d_t = (1 - \tau_t) \left(1 - \frac{1}{\mu}\right) Y_t$. We normalize the mass of firm shares outstanding to 1.

Fiscal policy sets the tax rate τ_t, spends G_t on goods, and has debt outstanding B_t, with

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

we let $T_t \equiv \tau_t Y_t$ denote the total tax revenue of the government. We assume that fiscal policy is specified in terms of plans for government bonds B_t and spending G_t: bonds follow the moving average (MA) process $B_t = B + \sum_{s \geq 0} \widehat{B}_s \epsilon^B_{t-s}$, and government spending follows the MA process $G_t = G + \sum_{s \geq 0} \widehat{G}_s \epsilon^G_{t-s}$, for some coefficients $\left\{\widehat{B}_s, \widehat{G}_s\right\}$ and some set of iid innovations $\left\{\epsilon^B_t\right\}$, $\left\{\epsilon^G_t\right\}$. Given the stochastic process for r_t, these processes imply a process for tax revenue T_t.

We assume that nominal wages are sticky. As described in Auclert et al. (2023a), this implies that labor is rationed, and we pick the equal allocation rule $n_{it} = N_t = Y_t$. Unions reset wages, implying a Phillips curve for price and wage inflation π_t, but this is not material to solve for quantities in our baseline setting.

Monetary policy sets the real interest rate r_t, using a rule for the nominal rate $1 + i_t = \mathbb{E}_t \left[\frac{1+r_t}{1+\pi_{t+1}}\right]$. This rule is subject to stochastic shocks: we have $r_t = r + \sum_{s \geq 0} \widehat{r}_s \epsilon^r_{t-s}$, for some coefficients $\{\widehat{r}_s\}$ and some set of i.i.d innovations $\{\epsilon^r_t\}$. We assume that the innovations $\left\{\epsilon^B_t, \epsilon^G_t, \epsilon^r_t\right\}$ are independent.

In equilibrium, markets for goods, stocks and bonds clear at all dates:

$$Y_t = G_t + \int c_{it} di \qquad \int s_{it} di = 1 \qquad \int b_{it} di = B_t \tag{23}$$

Parameter		Value	Parameter		Value
β	Discount factor (quarterly)	0.9569	Y	Output	1
r	Real interest rate (quarterly)	1%	G	Government spending	0
μ	Markup	1.02	T	Steady-state taxes	0
ρ_e	Autocorrelation of earnings	0.95	B	Bond supply	0
σ_e	Cross-sectional sd of log earnings	0.92	p	Stock market value	1.96

Table 1: Baseline calibration

Calibration of the model without aggregate risk. Our calibration of the model without aggregate risk is summarized in table 1. We assume that utility is $u(c) = \log c$. We normalize output and labor to $Y = N = 1$, and set the markup to $\mu = 1.02$. We assume no government spending or bonds, so that there are also no taxes ($\tau = 0$). We set $r = 1\%$ quarterly. Hence, the stock market value is $\frac{1}{r}\left(1 - \frac{1}{\mu}\right) = 1.96$ times quarterly GDP. In steady state, agents are indifferent between holding stocks and bonds, and only their total asset position $a_{it} = ps_{it} + b_{it}$ is determined. We discretize the income process so that e_{it} follows an AR(1) in logs with quarterly autocorrelation of 0.95 and cross-sectional standard deviation of logs of 0.92, which is a standard high-income-risk calibration in the literature. Finally, we find the $\beta = 0.9569$ that delivers our target for r.

Aggregate shocks. We consider simple AR(1) shocks to government spending, bonds, and monetary policy. We set a common persistence of these shocks of $\rho = 0.9$ quarterly. We then set $\widehat{G}_s = \widehat{B}_s = 0.01 \cdot \rho^s$, and $\widehat{r}_s = 0.0025 \cdot \rho^s$, so that a unit standard deviation shock corresponds respectively to a change in government spending of 1% of GDP, a change in bonds of 1% of quarterly GDP, and a change in the real interest rate of 1 percent at an annualized rate.

Exogenous vs endogenous portfolios. In a first-order perturbation of the model in the standard deviation of aggregate shocks $\{\epsilon_t^B, \epsilon_t^G, \epsilon_t^r\}$, portfolio choice is undetermined. We resolve this indeterminacy by assuming that all agents hold 100% stock portfolios. This choice is natural given our calibration, where there are no government bonds in the aggregate, and obviously satisfies the asset market clearing conditions (23). We can then solve the model using traditional first-order perturbation methods, such as the sequence-space Jacobian method (Auclert et al. 2021).

Agents that perceive aggregate risk have a well defined portfolio choice problem in the neighborhood of the steady state. These zeroth order portfolios must satisfy the second-order perturbation of the portfolio choice problem discussed in section 2. Since portfolios depend on impulse responses and impulse responses depend on portfolios, this in principle involves a fixed-point problem. We now show how to circumvent this fixed point problem in the complete markets case, with a simple modification of the sequence-space Jacobian method.

3.2 Sequence-space Jacobians with endogenous portfolios

We begin by reviewing the sequence-space Jacobian method with exogenous portfolios. This uses certainty equivalence to reduce the problem to one of finding impulse responses without aggregate risk after date 0. We then show how to modify this method to solve for portfolios and impulse responses under endogenous portfolios.

Exogenous portfolios. By certainty equivalence, we can assume that agents have perfect foresight with respect to aggregates after the initial realization of aggregate shocks at date 0. Specifically, a realization $\boldsymbol{\epsilon} \equiv \left(\epsilon^G, \epsilon^B, \epsilon^r\right)$ of innovations to fiscal and monetary policy implies a known time path $\{G_t, B_t, r_t\}_{t\geq 0}$ equal to $\left\{G + \epsilon^G \widehat{G}_t, B + \epsilon^B \widehat{B}_t, r + \epsilon^r \widehat{r}_t\right\}_{t\geq 0}$.

Given an initial distribution of agents $\mathcal{D}$ over state variables $(s_{i,-1}, b_{i,-1}, e_{i0})$, we discuss how to solve the perfect-foresight equilibrium and then linearize in $\left\{\widehat{G}_t, \widehat{B}_t, \widehat{r}_t\right\}$.

The time path for $\{G_t, B_t, r_t\}_{t\geq 0}$ implies the time path $T_t = (1 + r_{t-1}) B_{t-1} + G_t - B_t$ for taxes. We know $w_t = \frac{1}{\mu}$ at all times, and using our rationing rule $n_{it} = N_t = Y_t$, household labor income per unit of skill is $w_t (1 - \tau_t) N_t = \frac{1}{\mu} (Y_t - T_t)$. Moreover, any household that is not at a borrowing constraint enforces $1 + r_t = \frac{p_{t+1} + d_{t+1}}{p_t}$ and is indifferent between holding bonds and stocks. This implies in particular that the asset price is always the present discounted value of future dividends, $p_t = \sum_{s=0}^{\infty} \left(\prod_{u=0}^{s} \frac{1}{1+r_{t+u}}\right) d_{t+s}$.

For $t \geq 0$, given indifference between bonds and stocks after date 0, only the asset position $a_{it} = p_t s_{it} + b_{it}$ is pinned down. At date 0, the incoming portfolios matter $(s_{i,-1}, b_{i,-1})$ matter because the return on stocks is $\frac{p_0 + d_0}{p}$, which may differ from the steady state real interest rate r. Hence, the household problem is:

$$\max_{c_{it}, a_{it}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_{it})$$
$$c_{it} + a_{it} \leq (1 + r_{t-1}) a_{it-1} + e_{it} \left(\frac{Y_t - T_t}{\mu}\right); \quad a_{it} \geq 0; \quad \text{all } t > 0 \tag{24}$$
$$c_{i0} + a_{i0} \leq (p_0 + d_0) s_{i,-1} + (1 + r) b_{i,-1} + e_{it} \left(\frac{Y_0 - T_0}{\mu}\right); \quad a_{i0} \geq 0$$

We note that individual decisions at all times are only a function of the time paths $\{r_t\}$, $\{Y_t - T_t\}$, and the initial value $p_0 + d_0$. Then, given the initial distribution $\mathcal{D}$, the time path of the distribution of agents over the relevant states (here (a_{it}, e_{it})) is known. Therefore, aggregate assets $\int a_{it} di$ are given by a certain sequence-space function $\mathcal{A}_t \left(\{r_t\}, \left\{\frac{Y_t - T_t}{\mu}\right\}; p_0 + d_0, \mathcal{D}\right)$.

Given $\{G_t, B_t, r_t\}_{t\geq 0}$ (so $\{T_t\}_{t\geq 0}$) and the initial distribution $\mathcal{D}$, equilibrium under perfect foresight is a sequence $\{Y_t, p_t\}$ for output and stock market values which solves:

$$\mathcal{A}_t\left(\{r_s\}, \left\{\frac{Y_s - T_s}{\mu}\right\}, p_0 + \left(1 - \frac{1}{\mu}\right)(Y_0 - T_0), \mathcal{D}\right) = p_t + B_t \quad \forall t \tag{25}$$

$$p_t = \sum_{s=1}^{\infty} \left(\prod_{u=0}^{s} \frac{1}{1 + r_{t+u}}\right) \left(1 - \frac{1}{\mu}\right)(Y_s - T_s) \quad \forall t \tag{26}$$

We note that we can substitute in (26) into (25) to obtain a single equation

$$H_t\left(\{Y_t\}, \{G_t, B_t, r_t\}, \mathcal{D}\right) = 0 \tag{27}$$

where H_t gives net asset demand $\mathcal{A}_t - (p_t + B_t)$ given the unknown sequence $\{Y_t\}$ and exogenous sequences $\{G_t, B_t, r_t\}$. Stacking $\mathbf{U} = \{Y_t\}$ and $\mathbf{Z} \equiv \{G_t, B_t, r_t\}$ as vectors, we rewrite (27) as:

$$\mathbf{H}\left(\mathbf{U}, \mathbf{Z}, \mathcal{D}\right) = 0 \tag{28}$$

This has the generic form of a nonlinear sequence-space system.

Sequence-space Jacobians with exogenous portfolios. With exogenous portfolios, shocks do not affect the distribution $\mathcal{D}$ in (27). Hence, linearizing (27) delivers

$$\mathbf{H_U} d\mathbf{U} + \mathbf{H_Z} d\mathbf{Z} = 0$$

and assuming that $\mathbf{H_U}$ is invertible, we find the traditional, exogenous-portfolio sequence-space solution $d\mathbf{U} = -\mathbf{H_U}^{-1}\mathbf{H_Z} d\mathbf{Z}$. This involves calculating the sequence-space Jacobians $\mathbf{H_U}$, $\mathbf{H_Z}$ in the standard way (see Auclert et al. 2021).

Sequence-space Jacobians with exogenous portfolios. With endogenous portfolios, but complete markets, section 2.2 shows that we can now think of the distribution as effectively moving with the shocks: starting from the exogenous portfolio distribution, now equation (16) gives the optimal sensitivity of an agent's payoff to any given shock, which because the t_i's net out is like a change in the initial distribution. Differentiating (28) now yields

$$\mathbf{H_U} d\mathbf{U} + \mathbf{H_Z} d\mathbf{Z} + \mathbf{H}_{\mathcal{D}} d\mathcal{D} = 0 \tag{29}$$

The distributional change is that implied by equations (14)–(16), which, if we write $\boldsymbol{\lambda}$ for the $Z \times 1$ vector of λ_z's, reads

$$d\mathcal{D} = \mathbf{D}_\lambda d\boldsymbol{\lambda} + \mathbf{D_U} d\mathbf{U} + \mathbf{D}_Z d\mathbf{Z} \tag{30}$$

Moreover, $d\lambda$ can be solved for using the market clearing equations (15), this implies

$$d\boldsymbol{\lambda} = \boldsymbol{\lambda}'_{\mathbf{U}} d\mathbf{U} + \mathbf{D}_Z d\mathbf{Z} \tag{31}$$

Putting (29)–(31) together, we have:

$$\left(\mathbf{H}_U + \underbrace{\mathbf{H}_{\mathcal{D}} \mathbf{d}_\lambda \boldsymbol{\lambda}_U' + \mathbf{H}_{\mathcal{D}} \mathbf{D}_U}_{\equiv \mathbf{H}_U^{corr}} \right) d\mathbf{U} + \left(\mathbf{H}_Z + \underbrace{\mathbf{H}_{\mathcal{D}} \mathbf{d}_\lambda \boldsymbol{\lambda}_Z' + \mathbf{H}_{\mathcal{D}} \mathbf{D}_Z}_{\equiv \mathbf{H}_Z^{corr}} \right) d\mathbf{Z} = 0 \tag{32}$$

The sequence-space solution is then just $d\mathbf{U} = -\left(\mathbf{H}_U + \mathbf{H}_U^{corr}\right)^{-1}\left(\mathbf{H}_Z + \mathbf{H}_Z^{corr}\right) d\mathbf{Z}$.

Equation (32) shows that, in the complete-markets case, one can simultaneously solve for endogenous portfolios and impulse responses by simply adding a correction to sequence-space Jacobians for the effects of endogenous portfolios. This involves minimal extra work relative to the standard exogenous portfolio method, and as we show next, requires essentially the same objects as those involved in calculating the Jacobians $\mathbf{H}_U$, $\mathbf{H}_Z$.

Calculating corrected Jacobians in practice. Equation (32) shows that, to calculate impulse responses to endogenous portfolios, we need to calculate, for each input X to the H function, a) the Jacobians $\mathbf{H}_X$ of the exogenous-portfolio model for a default set of portfolios, b) the Jacobians $\mathbf{H}_X^{corr}$, reflecting the complete-market correction, and then c) add the two together. The corrected Jacobians $\mathbf{H}_X + \mathbf{H}_X^{corr}$ can then be used in place of the uncorrected Jacobians to form the solution $d\mathbf{U}$, $d\mathbf{Z}$ in equation (32). Once the solution is known, the equation for $d\mathcal{D}$ can be evaluated in (30), from which we can back out the supporting portfolios using equation (18), and the equation for $d\boldsymbol{\lambda}$ (31) can be evaluated, from which it is immediate to calculate second-order relative risk premia using equation (19).

We now discuss how to obtain the $\mathbf{H}_X^{corr}$ matrix. This is a Jacobian matrix whose columns are made up of the impulse responses of the target dH_t to individual shocks dX_s at different dates s, running through two effects. First, there is the direct effect ($\mathbf{H}_{\mathcal{D}}\mathbf{D}_X$) of the shock on the target H_t via the adjustment of transfers to the shock in equation (14). Second, there is the indirect effect ($\mathbf{H}_{\mathcal{D}}\mathbf{d}_\lambda\boldsymbol{\lambda}_X'$) of the shock on the targets via the loading of λ on the shock in equation (15), and the response of transfers to λ, again in (14). Remembering that the distribution matters for H_t through the aggregate asset function $\mathcal{A}_t$ in (25), this can be done concretely as follows.

Letting $Z_t \equiv \frac{Y_t - T_t}{\mu}$ denote aggregate post-tax labor income, define the value function corresponding to the sequence problem in (24) as a function of post-return assets a^p:

$$V_t^p(e', a^p) = \max_{a'} u(a^p + e' Z_t - a') + \mathbb{E}_0[V_{t+1}^p(e'', (1+r_t)a')|e']$$

Note that if we write $c \equiv a^p + e'Z_t - a'$ above, then $V_t^{p\prime} = u'(c)$. We then write

$$W_t(e, a^p) \equiv \mathbb{E}[V_t^p(e', a^p)|e]$$

taking expectations over the e' that will be defined in period t. It follows that

$$W_t'(e, a^p) = \mathbb{E}[u'(c(e', a^p))|e] \tag{33}$$
$$W_t''(e, a^p) = \mathbb{E}[u''(c(e', a^p)) \cdot mpc(e', a^p)|e] \tag{34}$$

where $mpc(e, a^p) \equiv c'(e', a^p)$.

At this point, for shocks at date 0, $W_0(e, a^p)$ maps to the $\bar{W}$ we defined in section 2, with the shocks ϵ now being subsumed in the subscript 0 (as opposed to the steady-state W), individuals i being indexed by their state (e, a^p), the default portfolio having return r_0, and $1 + r \equiv R$.

Precalculation. We need steady-state objects $W'(e, a)$, $W''(e, a)$, and $D^{beg}(e, a)$, where the latter is the "very-beginning-of-period" distribution after a was chosen last period, but prior to the draw of a new e' this period. We will specify these, and the other quantities we will use, over the pre-return grid for convenience, since that is the grid we will generally use in calculation. For $W'(e, a)$ and $W''(e, a)$, we simply evaluate (33) and (34) for the steady-state consumption and MPC at (e, a). D^{beg} can be calculated by simply applying the asset policy update to steady-state D.

Then, it is useful to aggregate $-\frac{W'(e,a)}{RW''(e,a)} \cdot D^{beg}(e, a)$ across all (e, a) to obtain $\Lambda \equiv \int -\frac{W_i'}{RW_i''} di$, the denominator in (15), in the process saving $\frac{W'(e,a)}{RW''(e,a)}$, which gives the sensitivity of $dT(e, a)$ to $d\lambda$ (for any shock $d\epsilon_z$, which we suppress in the notation here).

Extra backward step for fake news algorithm. We now can make just a slight modification to the fake news algorithm. When calculating any Jacobian in response to a shock at date s, we will naturally obtain as part of the calculation $dc_0(e', a)$, implying by (33) a shock

$$dW_0'(e, a) = \mathbb{E}[dc_0(e', a) \cdot u''(c(e', a))|e]$$

to dW_0'. We divide this by $-RW''(e, a)$ pointwise to obtain the "partial equilibrium" transfers $dT^{pe}(e, a)$. Then, we aggregate the partial equilibrium transfers and divide by Λ to obtain $d\lambda$, the proportional change in marginal value W_0' associated with the shock. Finally, we add $\frac{W'(e,a)}{RW''(e,a)} d\lambda$ to $dT^{pe}(e, a)$ to obtain the true transfers $dT(e, a)$, per equation (14).

At this point, starting at $D^{beg}(e, a)$, we construct a perturbation $dD_0^{beg}(e, a)$ from the transfers $dT(e, a)$, for simplicity sending a mass $\frac{dT(e,a_i)}{a_{i+1}-a_i}$ from gridpoint a_i to gridpoint a_{i+1}. This gives us $d\mathcal{D}$ in equation (30).

Constructing the complete markets correction H_X^{corr}. To obtain the overall Jacobian, we need to form $\mathbf{H}_{\mathcal{D}} d\mathcal{D}$ in (29). This answers the question "by how much is the path of aggregate assets affected if the initial distribution is perturbed by $d\mathcal{D}$". As explained in Auclert et al. (2021), the answer to this question is given by the product of the expectations vector for assets, $\mathcal{E}_t$, with the perturbation $d\mathcal{D}$. Since we constructed one such perturbation for each shock to inputs dX_s, we find that the complete market correction matrix H^{corr}_{ts} has

$$H^{corr}_{X,ts} = \mathcal{E}_t' \mathcal{D}_s \tag{35}$$

Equation (35) resembles the fake news matrix of the exogenous incomplete markets algorithm of Auclert et al. (2021), except that this starts in row 0 with $\mathcal{E}_t$ rather than $\mathcal{E}_{t-1}$, and it is actually the full correction to the Jacobian rather than a fake news matrix.

Incomplete markets. In proposition 3, we found the simple relationship $\mathbf{t}_i = \mathbf{X}(\mathbf{X}'\mathbf{\Sigma}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}\mathbf{t}_i^{CM}$: with incomplete markets, the effective transfer vector $\mathbf{t}_i$ to household i is simply the projection $\mathbf{P} \equiv \mathbf{X}(\mathbf{X}'\mathbf{\Sigma}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}$ of the complete-market transfer $\mathbf{t}_i^{CM}$ onto the column space of the relative return matrix $\mathbf{X}$.

Since the correction terms in (32) are linear in transfers, this projection also applies directly to the Jacobian corrections. Importantly, we can no longer solve shock-by-shock, but must instead solve for the impulse responses to all shocks jointly. This gives us a block sequence-space system, with $Z \times Z$ blocks total, and where block z, z' maps sequences in response to shock z' to sequences in response to shock z. The correction Jacobians in the z, z' block are then the corresponding entries of the projection matrix, $P_{z,z'}$, times the complete-markets correction terms ($\mathbf{H}^{corr}_U$ and $\mathbf{H}^{corr}_Z$). The non-correction Jacobians, $\mathbf{H}_U$ and $\mathbf{H}_Z$, appear only in diagonal blocks z, z, since spillovers between different shocks only occur through portfolio choice.

Given a projection matrix $\mathbf{P}$, this block sequence-space system is linear and can be directly solved for the impulse responses to each shock. These impulse responses, however, determine the relative return matrix $\mathbf{X}$ and thus $\mathbf{P}$, so that the overall equilibrium is a nonlinear fixed point. A simple and seemingly effective strategy to solve this is fixed-point iteration: start with a $\mathbf{P}$, solve the block sequence-space system given this $\mathbf{P}$ to obtain impulse responses, update $\mathbf{X}$ and $\mathbf{P}$, and repeat until convergence.[6]

Speeding up for many shocks. When there are many shocks, the block sequence-space system can become large, so that a direct linear solution is slow or even infeasible. Rather than a direct linear solution, one can try an iterative approach for this inner problem as well: solve for equilibrium sequences ignoring off-diagonal blocks, calculate the error including the off-diagonal blocks, solve again using only the diagonal blocks to offset this error, and so on. Since the block system is diagonally dominant—the non-correction terms, appearing only the diagonal blocks, are generally larger, especially for $t, s \gg 0$—this approach is quite effective in practice, and can cut the time required to solve the system by a factor of 100 or more when there are many (7+) shocks.[7]

[6] Simple initial guesses for $\mathbf{P}$ are either the identity, corresponding to the complete-market solution, or the $\mathbf{P}$ and $\mathbf{X}$ implied by the solution with baseline portfolios.

[7] For further efficiency, one can adapt this iterative idea to use a standard iterative method like GMRES, using the diagonal blocks as preconditioners for the overall system. It can also potentially be beneficial to ignore the diagonal correction terms, or replace them with their complete-market counterparts, so that the diagonal blocks can be inverted once and reused even when $\mathbf{P}$ changes.

Portfolio constraints. In section 2.5, we discuss how to solve for portfolios subject to constraints, conditional on relative returns $\mathbf{X}$ and the direct effects $\frac{\partial W_i'/W_i'}{\partial \epsilon_z}$ of each shock on marginal utility. In general equilibrium, both are endogenous, and are affected by the portfolios themselves. We can find the fixed point iteratively, starting with guesses for $\mathbf{X}$ and $\frac{\partial W_i'/W_i'}{\partial \epsilon_z}$ (for instance, from the exogenous-portfolio solution), then solving for portfolios subject to constraints and obtaining the implied general equilibrium $\mathbf{X}$ and $\frac{\partial W_i'/W_i'}{\partial \epsilon_z}$, and so on.[8]

4 Revisiting the effects of fiscal and monetary policy shocks

We now use the method developed in the previous two sections to ask how much endogenous portfolios matter for HANK models. We begin by considering each of our three shocks $\left(\epsilon^G, \epsilon^B, \epsilon^r\right)$ in isolation. Technically, we set the variance of the other two shocks to zero, so that there is a single aggregate shock in the model. Given that agents have access to a stock and a bond, we have two assets and one shock, and therefore markets are complete with respect to aggregate risk whenever the spanning condition is satisfied. In the notation of section 2, we have $K = Z = 1$, and we just need to verify assumption 1 to apply proposition 1.

4.1 Balanced-budget fiscal shocks

We start with pure shocks to government spending G. Given that $\overline{\sigma}^B = \overline{\sigma}^r = 0$, we have $r_t = r$ and $B_t = B = 0$, and therefore $T_t = G_t$. Hence, these are balanced-budget government spending shocks.

Figure 1 visualizes the impulse response to our AR(1) government spending shock $\widehat{G}_s$. We find that we show that the impulse responses under endogenous portfolios are the same as under exogenous portfolios. The reason is that with these shocks, and $r_t = r$, in this model there is a unit government spending multiplier on output, with no effect on consumption for any agent (Haavelmo 1945, Auclert et al. 2023a). Here, this can be verified by seeing that $Y_t = Y + G_t = Y + T_t$, $p_t = p$, $d_t = d$, $B_t = 0$ solves equations (25)–(26). Since r, $Y - T$, and $p + d$ are constant, all the inputs into the household problem in (24) are constant. Since $\frac{p_0 + d_0}{p} = 1 + r$, the spanning condition (assumption 1) is actually violated, so households are fully indifferent between all portfolios. We summarize these results in the following proposition.

Proposition 4. *A shock to ϵ^G, with $\overline{\sigma}^B = \overline{\sigma}^r = 0$, implies $dc_{it} = 0$ for all agents i. Exogenous and endogenous portfolios deliver the same solution, even if additional assets are available to complete markets. Portfolios are undetermined.*

Proposition 4 shows that the unitary multiplier result from the HANK literature is unaffected by endogenous portfolios.

[8] Since solving for portfolios is itself iterative, we can achieve greater efficiency by not iterating to convergence every time this inner loop is run, but instead iterating it only a few times.

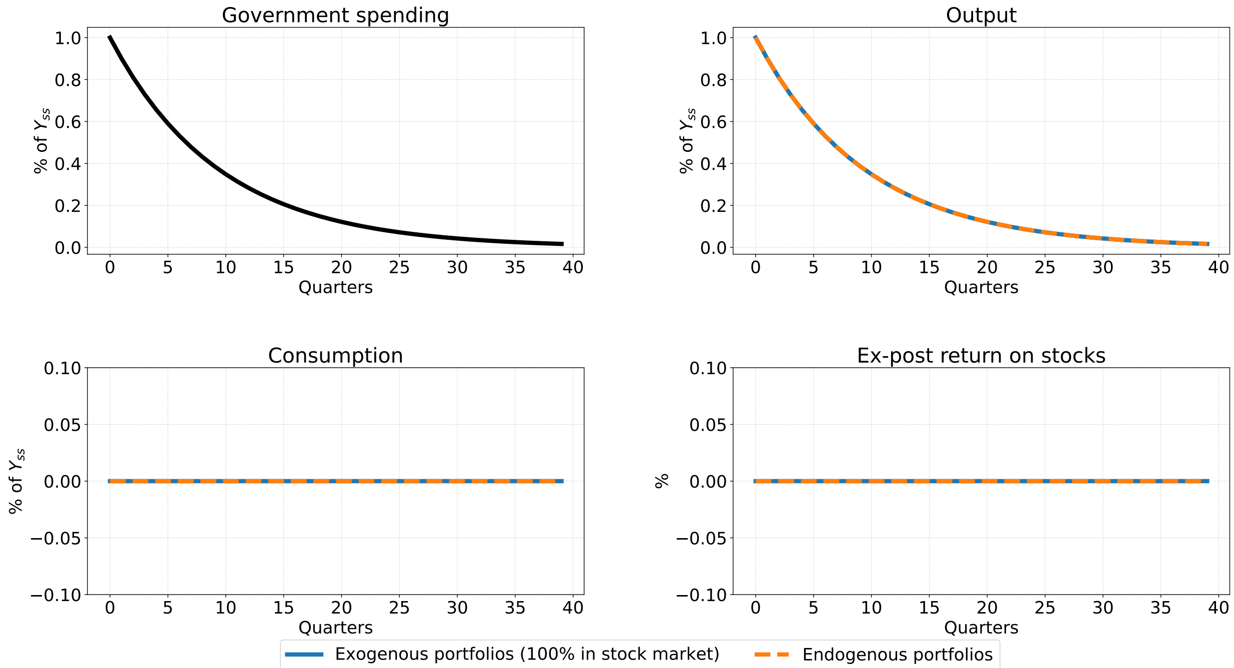

Figure 1: Impulse responses to balanced budget government spending shock

4.2 Monetary policy shocks

We now look at the case of pure monetary policy shocks. Here, portfolios are determinate since the monetary policy shock moves the stock market. Figure 2 visualizes the impulse responses to our shock $\widehat{r_s}$. As expected, the monetary policy shock lowers output and consumption, as well as the impact return on stocks as the stock price falls. However, here again, the impulse responses under endogenous portfolios are the same as under exogenous portfolios.

The result for this result is that both impulse responses are the same as the impulse response from a representative agent model, given by the Euler equation $dC_t = -C\sum_{s\geq 0} \frac{dr_{t+s}}{1+r}$. This is because under our assumptions about preferences (in particular log utility), technology (equal rationing), and 100% stock portfolios, in equilibrium the Euler equation is satisfied for every agent i—a result initially derived by Werning (2015). We summarize this result in the following proposition.

Proposition 5. *A shock to ϵ^r, with $\overline{\sigma}^G = \overline{\sigma}^B = 0$, implies $dc_{it} = -c_{it}\sum_{s\geq 0} \frac{dr_{t+s}}{1+r}$ for all agents i. Optimal portfolios are 100% stocks. Hence, exogenous and endogenous portfolios deliver the same solution.*

Intuitively, given the risk premium on stocks, any agent finds that increasing or reducing exposure to the stock market would lower their expected utility. Proposition 5 shows that the fact that the equivalence of HANK and RANK for monetary shocks is unaffected by endogenous portfolios.

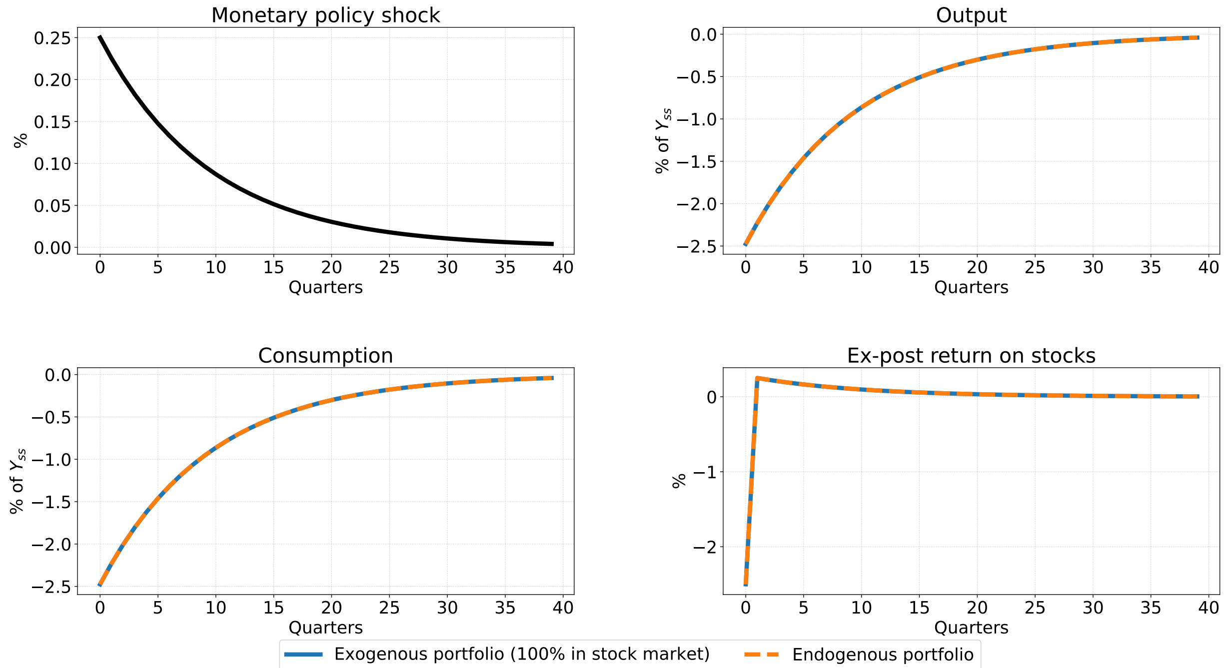

Figure 2: Impulse responses to monetary policy shock

4.3 Deficit-financed fiscal shocks

Finally, we consider the case of pure deficit shocks, $\hat{B}_t$. Given that $G = 0$, these shocks induce a path for taxes $T_t = (1+r)\,\hat{B}_t - \hat{B}_{t-1}$, and given that $\hat{B}_t \propto \rho^t$, this implies a one-time increse in transfers followed by a series of small tax increases as we go back to the steady state (figure 3, top left panel). Now, the impulse responses under endogenous portfolios differ significantly from those under exogenous portfolios. When portfolios are exogenous, deficit-financed tax cuts have large and persistent effects on economic activity in HANK, echoing earlier results in the literature (e.g. Auclert et al. 2023a, Bilbiie 2021, Angeletos et al. 2023, Auclert et al. 2023b). Here, the impact transfer multiplier is 0.2, and the effect persists for about five years, with a cumulative multiplier of around 0.77. By contrast, under endogenous portfolios, the impact transfer multiplier is only 0.08, though the effect persists for roughly the same amount of time, so that the cumulative multiplier is still 0.53.

To understand the intuition behind these effects, note first that, with 100% stock portfolios, deficit-financed fiscal transfers disproportionately raise the consumption of poor agents, and therefore disproportionately lower their expected marginal utility, violating condition (1). The left panel of figure 4 verifies this by plotting the marginal utility ratio from equation (1), $\lambda_0(a',e) \equiv \frac{\mathbb{E}[u'(c_0(a',e'))|e]}{\mathbb{E}[u'(c_{ss}(a',e'))|e]}$, for agents at different points in the distribution (a', e). We see that agents that choose assets close to the constraint $a' = 0$ tend to have a larger-than-average decline in their marginal utility conditional on a positive deficit shock (which raises transfers), especially when they have low income. Hence, condition (1) is not satisfied.

Since the stock market is booming (it is simple to show that the stock price increase on impact

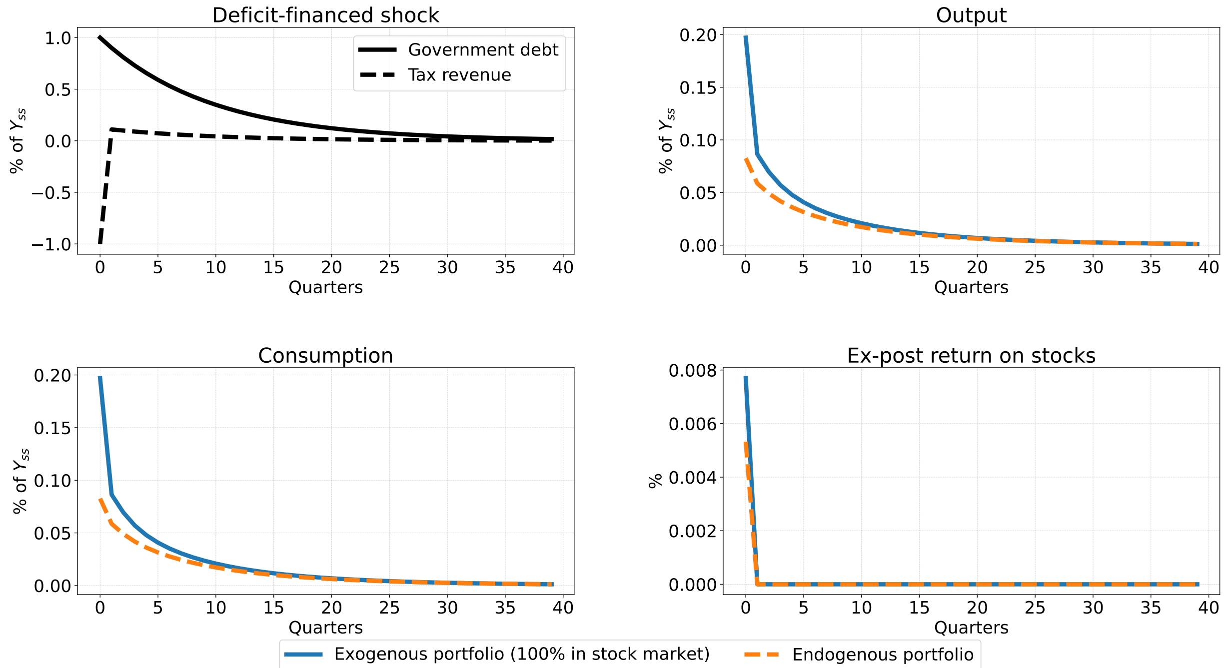

Figure 3: Impulse responses to deficit-financed transfer shock

is proportional to the value of the cumulative transfer multiplier), optimal portfolios reduce the stock market exposure of poor agents and raise the exposure of rich agents. Since rich agents have lower marginal propensity to consume out of capital gains than poor agents, optimal portfolios reduce the aggregate transfer multiplier. The effect of this additional redistribution is large on impact, explaining the large reduction in the transfer multiplier. But it is also relatively short-lived, so that the overall persistence of the output effect is less affected by endogenous portfolios.

The right panel of figure 4 visualizes the optimal portfolios underlying figure 3, calculated using equation (18). Note that the portfolio shares are extreme, especially for agents close to the borrowing constraints: all agents with low asset choices for the next period have short positions in the stock market, including some with thousands of times their net worth. The intuition is that the logic of portfolio choice in (18) is to pick a certain total exposure to the stock market, rather than a certain portfolio share. The optimal exposure of poor agents is negative, so as net worth shrinks towards zero, their portfolio share becomes more and more negative and tends to $-\infty$.

Obviously, those portfolio shares are extreme and very likely counterfactual. Hence, the result from this section should be seen as an upper bound of how much endogenous portfolios can shrink the transfer multiplier in a HANK model. We explicitly consider the consequence of adding realistic restrictions to portfolio choice in section 6, where we will see that, as is intuitive, reasonable portfolio constraints make the endogenous portfolio impulse response match the exogenous portfolio response much more closely.

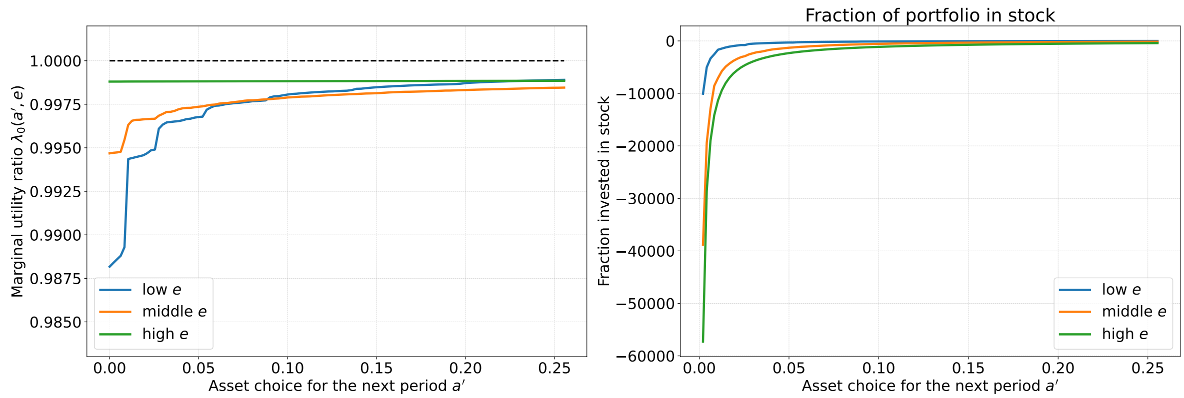

Figure 4: λ test at 100% stock portfolios (left) and optimal portfolios (right)

4.4 Risk premia

Section 2.3 showed that our procedure allows us to recover relative risk premia on assets. Table 2 reports the level of these risk premia, reported as the percentage difference between the stock and the bond return. The risk premium on the monetary policy shock is 6.1×10^{-4}, or 24 basis points annually. This is small because the regular consumption-CAPM formula approximately holds here, and utility is log. Given that an AR(1) 25bp shock to monetary policy raises consumption on impact by $\frac{0.25}{1-\rho}\frac{1}{1+r} \simeq 2.47\%$ and lowers the stock market return by 2.47%, the covariance betwen aggregate consumption and returns is equal to 6.1×10^{-4}, which is the risk premium in the standard formula when risk aversion is 1.

The risk premium on the deficit-financed fiscal shock is 20000 times lower, at 3.6×10^{-8} at a quarterly level, or a few one-hundredths of hundredths basis points. The reason is that the "quantity of risk" here is small, as the deficit-financed fiscal shock has limited effect on both consumption (with a transfer multiplier of 0.2) and returns (with an impact effect of the stock market of only 3bp annually). The latter can be understood since the shock to returns is $\frac{r}{1+r}$ times the cumulative multiplier in our model.[9] The intuition is that a demand boom on its own only affects the stock market via the cumulative multiplier on output. Given this, the consumption-CAPM formula does not hold exactly,[10] but it still provides a very useful approximation for the effect on the risk premium.

[9]The shock to returns is $\frac{\Delta(p+d)}{p+d} = \frac{d\Delta PDV(Y)}{p+d} = \frac{r}{1+r}\Delta PDV(Y)$, ie the annuitization ratio times the cumulative multiplier. (The effect from changing taxation does not affect the stock market since $\Delta PDV(T) = 0$).

[10]Instead, a "C^*-CAPM" formula holds for pricing assets, where C^* corrects aggregate C for the effect of idiosyncratic shocks.

	G shock	r shock	B shock
$\frac{R^{stock}-R^{bond}}{R}$ (quarterly)	0	6.1×10^{-4}	3.86×10^{-8}

Table 2: Risk premia for our baseline shocks

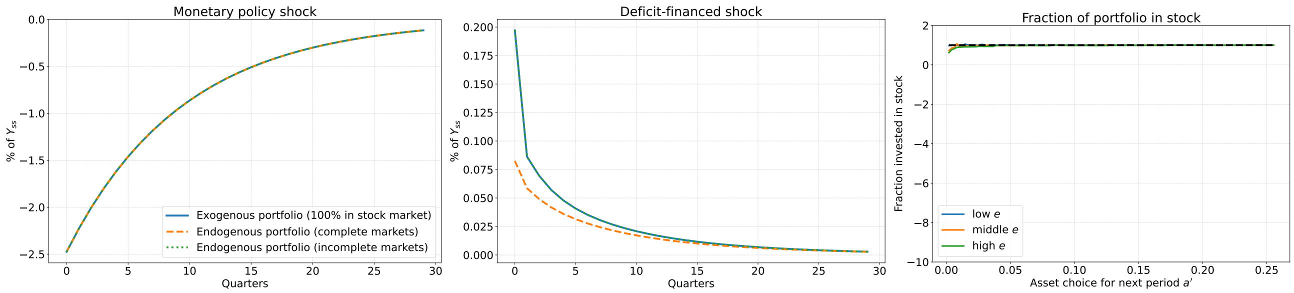

Figure 5: Impulse responses of output and portfolios when both shocks are present

5 Incomplete markets

We found that, with endogenous portfolios, the effects of deficit-financed fiscal shocks can be greatly attenuated relative to exogenous portfolios. However, this was a scenario where all other shocks in the model were turned off. It is natural to consider the case where multiple shocks occur. In our baseline model, the only two effective shocks are the monetary policy shock and the deficit-financed fiscal shock, so that with a bond and a stock, $Z = 2$ but $K = 1$ and markets are incomplete (section 5.1). Adding long-term bonds or nominal bonds restores market completeness (section 5.2); however, there could also be many different types of fiscal shocks (section 5.3).

5.1 Two assets, two shocks

We begin with the case of our full model with both the monetary policy and the deficit-financed fiscal shock present. We solve the model following the incomplete market procedure described in proposition 3 and section 3.2. Figure 5 shows the impulse response to monetary and fiscal shocks in this incomplete markets scenario, as well as the sustaining portfolios. Note that now impulse responses to different shocks are coupled, since the presence of one shock influences the portfolios and therefore the impulse response to other shocks. However, in this particular case, the impulse responses to both shocks are virtually identical to those with exogenous portfolios. The intuition is that the fiscal shock is much harder to hedge than the monetary policy shock, given the limited variation in returns that it induces, and therefore optimal portolios are near 100% stocks. Since they are not exactly equal to 100% stocks for all agents, however, there is a small difference relative to the exogenous portfolio case, but here it is not detectable.

5.2 More assets: adding nominal or long-term bonds

Households in practice do have more assets they can use than just a stock and a bond. In particular, they can use longer-maturity instruments, and can invest in nominal assets.

We modify the model so that households can invest in a long-term real bond (with exponentially decaying coupon) or in a nominal short term bond. With our two shocks, this restores complete markets. The sustaining portfolios are presented in figures 6 and 7 respectively. Those port-

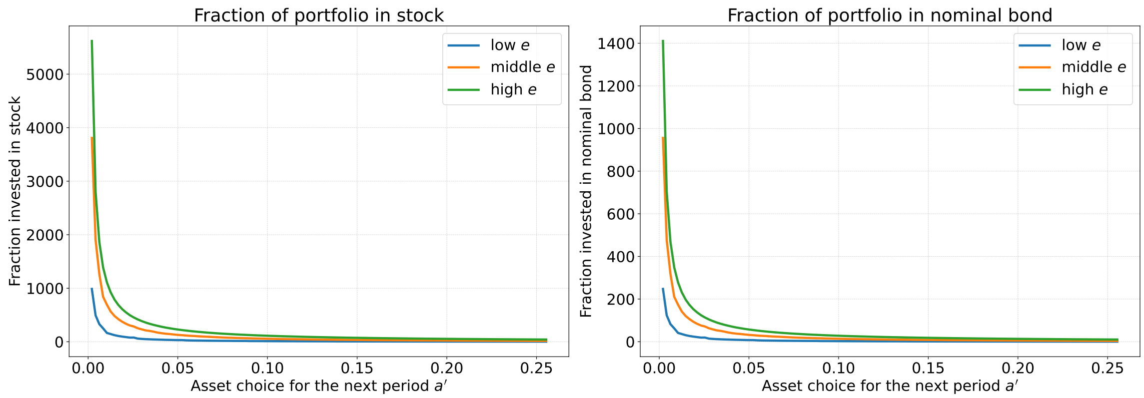

Figure 6: Portfolios with nominal, real one period bonds and stocks

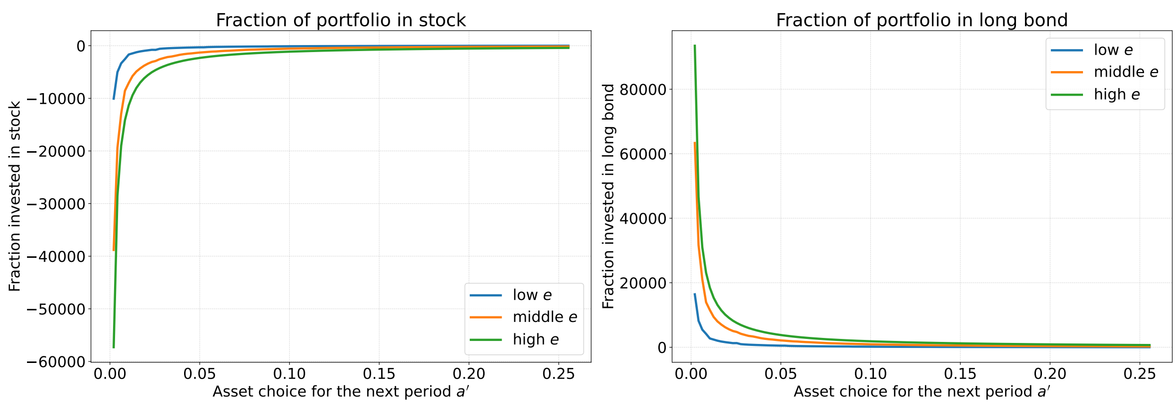

Figure 7: Portfolios with short, long real bonds and stocks

folios are still fairly extreme and implausible.

5.3 More shocks

In practice, households do not face a single type of any shock: for instance, fiscal transfers could be financed with different paths for government debt; and in our setting this corresponds to different shocks. Given this consideration, incomplete markets may be the more practically relevant case. Here we investigate how much this could matter. We look at the impulse response to our baseline $\rho = 0.9$ deficit-financed shocks where there are other deficit-financed shocks that could hit with different persistences. On the left panel of figure 8, we progressively add shocks that have persistence 0.9, 0.91, 0.92 up to 0.95, and visualize the impulse response to our original shock. We see that, in this incomplete markets situation, the impulse response converges back to the original exogenous-portfolio one. This is because stock prices respond more to the more-persistent shocks—and to offset this, households choose less aggressive stock positions relative to the original exogenous portfolio.

Inversely, the right panel shows what happens when we progressively add less persistent shocks, with persistences 0.89, 0.88 up to 0.85. Here, the incomplete markets impulse response

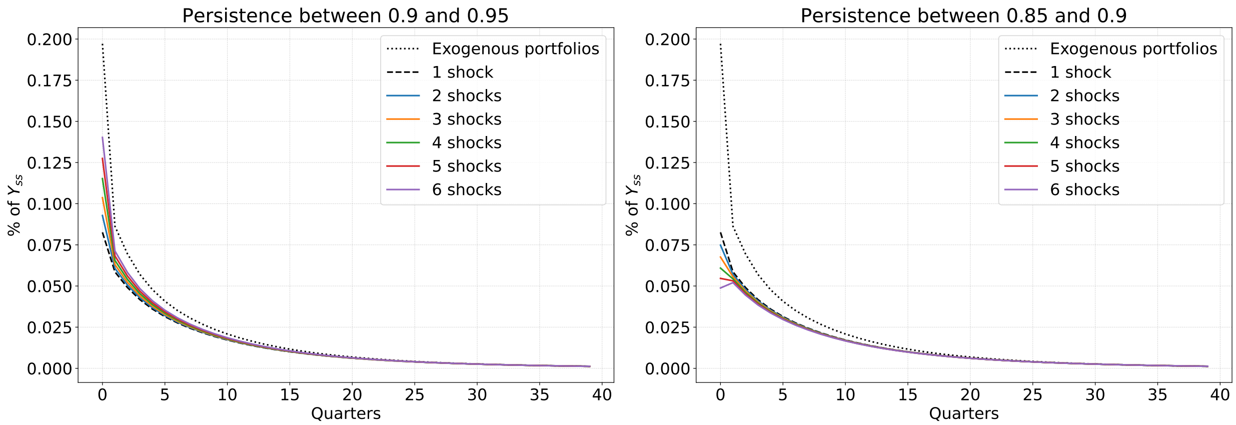

Figure 8: Impulse response to original ρ shock when other N shocks are present

progressively moves even further away from the complete markets one. The lesson is that incomplete markets do not have a monotonic relationship with the exogenous portfolio or complete markets impulse responses.

6 Portfolio restrictions

We return to our baseline deficit-financed transfer shock with $\rho = 0.9$, with no other shocks, but now use our method to impose both short-sale and leverage constraints. Specifically, we constrain stocks to be at least -100% and at most 200% of net worth. The former can be thought of as requiring households who short stocks to hold twice the value of their short equity position in bonds as collateral; the latter can be thought of as a leverage constraint preventing households from borrowing in bonds against more than half the value of their stocks.

The resulting equilibrium path for output is plotted in the left panel of figure 8. We see that it is virtually identical to the result with our baseline, exogenous portfolios.

The underlying portfolios are visualized in the right panel of figure 8. We see that once our restrictions are imposed, most households in equilibrium are either at their short sale or leverage constraints. This is a dramatic compression of the extreme portfolios we saw earlier for the complete-markets case in figure 4, bringing them much closer to the baseline, exogenous portfolios. The compression is most dramatic for the low-asset, high-MPC agents, whose consumption is most responsive to portfolio returns. It is accordingly no surprise that the general equilibrium results are now much closer to the exogenous portfolio case.

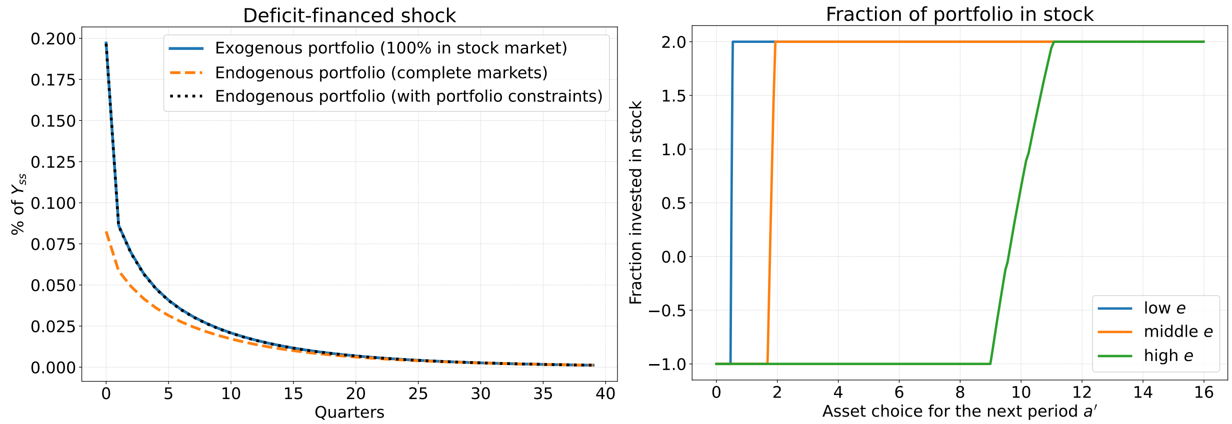

Figure 9: Deficit-financed shock with short-sale constraints

7 Conclusion

How much of a difference do endogenous portfolios make in HANK models? We find that they do not always make a difference. When they do, they tend to lower the impact effect of shocks on consumption. Consider for instance our deficit-financed fiscal shock, which disproportionately boosts the consumption of poor agents, these agents hedge against these shocks by choosing portfolios that underperform when the shock hits, and the overall consumption effect is therefore diminished. However, we found that the optimal portfolios underlying these results tend to be rather extreme, with poor agents taking large short positions in the stock market. When we add reasonable short sales and leverage constraints, the large reduction in the aggregate deficit-financed multiplier goes away. We obtain a similar though more subtle result when considering incomplete markets, when the model has more shocks than assets.

Our results should not be taken to mean that endogenous portfolios can never generate plausible portfolio distributions across agents, or that they can never make a significant difference when they do. Our approach is general enough to directly apply to any model solved with the sequence-space Jacobian method, including extensions to incomplete markets and portfolio constraints. We hope that it can be useful to investigate a wealth of questions where endogenous portfolios are important and relevant, such as asset allocations across countries, across risky asset categories such as stocks and houses, across maturities in mortgage borrowing, and so on.

Similarly, while our baseline model features small equity premia, the asset pricing literature has pointed out that cyclical dynamics in idiosyncratic risk can raise the equity premium to empirically reasonable levels (Mankiw 1986, Constantinides and Duffie 1996, Storesletten et al. 2004, Storesletten et al. 2007). Solving for portfolios and risk premia in models with these features is now feasible, opening up the door to studying the interactions between macroeconomics and finance in heterogeneous-agent settings with realistic risk premia.

References

Angeletos, George-Marios, Chen Lian, and Christian K. Wolf, "Can Deficits Finance Themselves?," Working Paper 31185, National Bureau of Economic Research, April 2023.

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

_, Matthew Rognlie, and Ludwig Straub, "The Intertemporal Keynesian Cross," Working Paper 25020, National Bureau of Economic Research, May 2023.

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

Backus, David K. and Gregor W. Smith, "Consumption and Real Exchange Rates in Dynamic Economies with Non-Traded Goods," *Journal of International Economics*, November 1993, *35* (3–4), 297–316.

Baxter, Marianne and Urban J. Jermann, "The International Diversification Puzzle Is Worse Than You Think," *American Economic Review*, March 1997, *87* (1), 170–180.

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

Bilbiie, Florin O., "Monetary Policy and Heterogeneity: An Analytical Framework," *Manuscript*, August 2021.

Campbell, John Y., "Chapter 13 Consumption-based asset pricing," in "Handbook of the Economics of Finance," Vol. 1, Part B, Elsevier, 2003, pp. 803–887.

Cochrane, John H., *Asset Pricing: Revised Edition*, Princeton University Press, January 2005.

Constantinides, George M. and Darrell Duffie, "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy*, April 1996, *104* (2), 219–240.

Devereux, Michael B. and Alan Sutherland, "Country Portfolios in Open Economy Macro-Models," *Journal of the European Economic Association*, April 2011, *9* (2), 337–369.

Guvenen, Fatih, "A Parsimonious Macroeconomic Model for Asset Pricing," *Econometrica*, November 2009, *77* (6), 1711–1750.

Haavelmo, Trygve, "Multiplier Effects of a Balanced Budget," *Econometrica*, 1945, *13* (4), 311–318.

Kaplan, Greg and Giovanni L. Violante, "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, July 2014, *82* (4), 1199–1239.

— , **Benjamin Moll, and Giovanni L. Violante**, "Monetary Policy According to HANK," *American Economic Review*, March 2018, *108* (3), 697–743.

Krusell, Per and Anthony A. Smith, "Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns," *Macroeconomic Dynamics*, 1997, *1* (02), 387–422.

Magill, M. J. P. and M. Quinzii, *Incomplete Markets, Volume 2: Infinite Horizon Economies*, Edward Elgar Publishing Company, 2008.

Mankiw, N. Gregory, "The Equity Premium and the Concentration of Aggregate Shocks," *Journal of Financial Economics*, September 1986, *17* (1), 211–219.

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*, January 2004, *28* (4), 755–775.

Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron, "Cyclical Dynamics in Idiosyncratic Labor Market Risk," *Journal of Political Economy*, June 2004, *112* (3), 695–717.

— , **Christopher I. Telmer, and Amir Yaron**, "Asset Pricing with Idiosyncratic Risk and Overlapping Generations," *Review of Economic Dynamics*, October 2007, *10* (4), 519–548.

Tille, Cédric and Eric van Wincoop, "International Capital Flows," *Journal of International Economics*, March 2010, *80* (2), 157–175.

Werning, Iván, "Incomplete Markets and Aggregate Demand," Working Paper 21448, National Bureau of Economic Research, August 2015.