Sticky Expectations, Fiscal Transfers, Inflation, and Unemployment in HANK

Noah Kwicklis[1]

November 8, 2025[2]

Link to the latest version.

Abstract

I use a new numerical technique to solve and estimate a medium-scale sticky-expectation heterogeneous agent New Keynesian (HANK) model in order to study the effects of the post-COVID fiscal stimulus. Sticky information frictions triple the one-year transfer multiplier of a stimulus check intervention from 0.10 to 0.30. The model predicts that the U.S. fiscal transfer stimulus in 2020 and 2021 reduced the COVID recession's cumulative real GDP per capita losses by 25%, compared to a scenario where the transfers did not take place. Despite being highly stimulative for output, the model suggests that transitory fiscal transfers have relatively modest impacts on inflation and unemployment.

[1] Job market candidate at UCLA. I thank Lee Ohanian, Pierre-Olivier Weill, João Guerreiro, Andrew Atkeson, and Saki Bigio for their feedback and comments.

[2] First draft: July 20, 2025.

1 Introduction

Deviations from full information rational expectations (FIRE) are qualitatively and quantitatively important for understanding business cycles and are often necessary to reconcile heterogeneous agent New Keynesian (HANK) models with macroeconomic data. In particular, sticky expectations, in which households only update their forecasts of the macroeconomy to full information according to a Poisson process, are key to matching the effects of monetary policy in HANK to empirical estimates. However, we still know relatively little about the effects of fiscal policy on unemployment and inflation in such models. This gap is especially salient given the massive fiscal stimulus undertaken by the U.S. federal government during the COVID-19 pandemic and the job market recovery and surge in prices that followed. Using an estimated sticky-expectation medium-scale HANK model disciplined by both time series and microeconomic evidence, I find that the post-COVID stimulus reduced the output losses during the recession relative to trend by 25%, but likely had only modest impacts on inflation and unemployment.

The model is built with quantitative realism in mind, with search-and-matching frictions in the labor market, inequality and incomplete asset markets, marginal propensities to consume that match survey data, slow-adjusting Taylor and debt rules, and a standard set of Smets and Wouters (2007) shocks and frictions. However, sticky expectations introducing an additional layer of complexity to solving and estimating HANK models. As such, I use a new continuous-time technique developed in Kwicklis (2025b) to convert the model to a sticky expectations system after solving it under FIRE conditions in state space. Sticky expectations then give the model's dynamics enough realism to estimate the model's structural parameters using narratively identified monetary policy shocks. The state space solution, in turn, allows me to use standard Kalman filtering and smoothing techniques to decompose macroeconomic time series into historical shocks. I then study the implications of the estimated model for the effect of transitory deficit-financed fiscal transfers – essentially stimulus checks – on output, inflation, and unemployment.

In the sticky expectations environment, I find that exogenous fiscal transfers that mean revert quickly are about a third more stimulative on impact than they would be in a similar full information version of the model. After a year, the total transfer multiplier from the fiscal stimulus is 0.30, three times higher than in the FIRE setting and despite the fact that the one-time transfer I model is not targeted. In the model, households only gradually realize that aggregate income has risen thanks to the stimulus policy, but are also slow to understand the policy's implications for future debt stabilization. The second effect dominates, making the transfers more powerful.

Despite the highly stimulative nature of transfers, I also find that they do not substantially boost employment or inflation. This is after the government disburses stimulus checks to individuals, constrained households spend them quickly, leading to an acute consumption-led increase in aggregate demand. This surge in economic activity is sharp and transitory, while firm hiring and repricing decisions in standard search-and-matching and sticky price models are longer-horizon investments that take time and resources. As such, the firms in the model are too slow to capitalize significantly on abrupt surges in demand with higher prices or vacancy posts and instead increase labor on its intensive

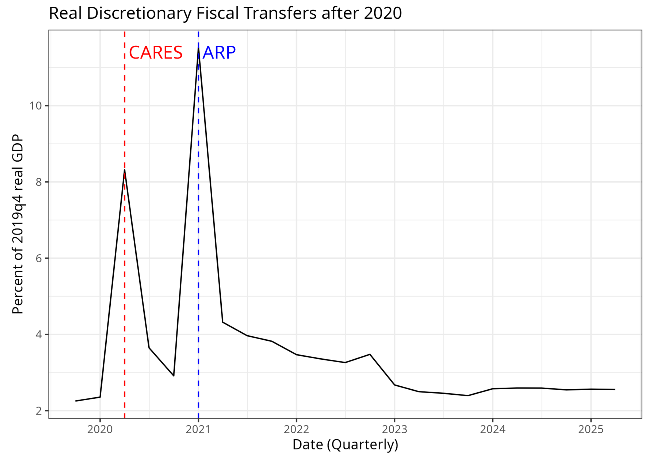

Figure 1: Quarterly discretionary fiscal transfers after 2020 as a percentage of 2019q4 GDP. Data includes all government social benefits to persons (personal current transfer receipts) excluding payments for Social Security, Medicare/Medicaid, veteran's benefits, and unemployment insurance. The series is deflated using the GDP deflator. CARES Act transfers were disbursed at the start of 2020q2 (marked in red), while American Rescue Plan (ARP) Act funds were disbursed near the end of 2021q1 (marked in blue).

margin (hours worked per worker). The empirical estimates from the identified monetary shock also imply a very flat Phillips curve, further dampening the responsiveness of prices to fiscal policy.

I then evaluate the relative effects of monetary and fiscal transfer policy in the historical time series of output and inflation. I focus on the extraordinary fiscal stimulus enacted to boost the U.S. economy after the recession that began in 2020: after the passage of the CARES Act and the American Rescue Plan, discretionary[3] federal social benefit transfers to households as a percentage of pre-crisis quarterly GDP jumped sharply twice, to over 8% in 2020q2 and to 11.5% in 2021q1 (see Figure 1). I estimate the persistence and variance parameters of a set of nine shocks in the model via Sequential Monte Carlo (SMC), using postwar U.S. data and a measurement equation reminiscent of Smets and Wouters (2007) but with the addition of fiscal transfers to GDP and unemployment rate data. The historical shock decomposition suggests that had such fiscal transfers not occurred, the recession's GDP losses relative to trend would have been 33% larger – or equivalently, that the stimulus reduced the losses from the recession by by a quarter. The model finds little role for fiscal stimulus in lowering unemployment or driving up inflation.

I briefly survey related work on HANK models and departures from FIRE in the literature review, along with work related to fiscal stimulus during and after the pandemic. In Section 2, I introduce the medium-scale HANK model with sticky expectations and unemployment frictions that I use to study fiscal transfers. In section 3, I describe the model's steady-state calibration, the estimation of key structural parameters using impulse response function matching, and the Bayesian estimation of the remaining parameters using Sequential Monte Carlo. Section 4 presents the results for the effects of fiscal transfers with and without sticky expectations and the implications of the estimated model for the COVID era. Section 5 concludes.

[3] Excluding social security, Medicare/Medicaid, veterans' benefits, and unemployment insurance.

1.1 Literature Review

i. Models with Information frictions and HANK

Numerous papers have explored how information frictions and rational inattention are key to explaining the slow movement of macroeconomic aggregates following changes in macroeconomic news relative to the speed of adjustment achieved in full information rational expectations models. In particular, Campbell and Deaton (1989) documents the "excess smoothness" of aggregate consumption's movement in response to news about permanent income – even as Campbell and Mankiw (1989) finds that aggregate consumption is excessively sensitive to *current* income relative to the permanent income hypothesis of Friedman (1957). Mankiw and Reis (2002) describe a model in which price-setters update their information about the macroeconomic only periodically, and in so doing generates realistic inertia in the movement of inflation. Reis (2006) suggests that such periodic updating may be the result of rational inattention if macroeconomic information is costly to obtain or process, and then shows that it can explain aggregate consumption's excess smoothness – along with its excess sensitivity to current income, if costly processing shortens some households' planning horizons. Finally, using survey data evidence, Carroll (2003) shows that households are slow to update their expectations about macroeconomic aggregates, and this phenomenon is well captured by a Calvo-style (effectively, Poisson) process in which households only update their forecasts of the future periodically – a framework now known as "sticky expectations."

In a seminal paper that merges HANK with the sticky expectation structure of Carroll (2003), Carroll et al. (2020) use a Krusell and Smith (1998) approach to solve a simple state-space HANK model by tracking the entire distribution of infrequently-updating household expectations. They then demonstrate that their simulated model replicates the empirically observed sluggish response of household consumption to macroeconomic events, providing a key friction that explains consumption's "excess smoothness" puzzle. Simultaneously, the high marginal propensities to consume of liquidity-constrained households in the incomplete markets environment (as documented earlier in Kaplan, Moll, and Violante (2018), Auclert, Rognlie, and Straub (2024), and other HANK papers) explain the excess sensitivity of consumption to current income and aggregate demand. Carroll et al. (2020) then studies the stimulus checks issued after the 2008 Global Financial Crisis as a laboratory to test the household-level empirical implications of HANK with information frictions. My analysis is complementary: in contrast, I study the *macroeconomic* implications of stimulus check policy in the HANK environment in an estimated, medium-scale model.

In a related branch of the literature, Auclert, Rognlie, and Straub (2020) also combine a HANK model with sticky expectations. The authors instead focus on reconciling the high marginal propensities to consume documented in Fagereng, Holm, and Natvik (2021) with empirical evidence pertaining to the effects of monetary policy related to the vector autoregression (VAR) and local projections (LP) literatures, a step that I follow in my paper. They show that information frictions are critical to reconcile the high MPCs of HANK models required to match microeconomic survey data with the empirical estimates of a slow-moving "hump" shaped response to a cut in interest rates. In this way, my work continues this logic to examine the implications of such models for fiscal policy, given that monetary policy changes are relatively well-studied macroeconomic experiments that can be used to identify key structural parameters.

ii. Fiscal transfers, output, and inflation

When considering the interaction between information frictions and fiscal stimulus, Gabaix (2020) builds a simple behavioral New Keynesian model with cognitive discounting, in which agents attenuate their forecasts of the future toward the steady-state at a rate geometric in the forecasting horizon. The relative down-weighting of future macroeconomic forecasts leads the representative agent to violate Ricardian equivalence, leading deficit-financed fiscal transfers to become stimulative, as their expectations of the future taxation to stabilize the debt. In numerical simulations, I show that a similar phenomena occurs in my estimated HANK model, leading transfers stimulus to be significantly more stimulative with sticky expectations.

Recent fiscal policy has been analyzed in policy frameworks that involve fiscal dominance. Bianchi, Faccini, and Melosi (2023) estimate a rational expectations two-agent New Keynesian (TANK) model in a Fiscal Theory of the Price Level (FTPL) paradigm with active fiscal policy and passive monetary policy in the sense of Leeper (1991) and find that fiscal deficits could have accounted for the post-COVID inflation in their framework. Smets and Wouters (2024) estimate a similar RANK model and find transfer shocks under an active fiscal policy could have accounted for a significant fraction – but not a majority – of inflation in that same era while still boosting real GDP significantly. Kwicklis (2025a) explores the tradeoff between output and inflation in a more stylized rational expectations HANK setting when fiscal policy is active and finds that the ratios of positive output gaps to inflation broadly match the experience of the post-COVID era. I do not model such fiscally dominant regimes in this paper, and instead employ a more conventional active monetary, passive fiscal framework. As a consequence, I show that fiscal transfers in an estimated HANK model with realistically high MPCs can be powerfully stimulative, but only slightly inflationary, if one is not in a fiscally-led policy regime.

My paper is also related to several recent, less model-based analyses. Baker et al. (2023) study the stimulus payments from the 2020 CARES Act using high-frequency banking data and estimate the average recipient spent about 25-30 cents on the dollar in the first quarter of receiving their transfer, a number broadly in line with the quarterly version of the MPCs estimated in Fagereng, Holm, and Natvik (2021) used to calibrate my HANK model. Consistent with HANK theory, they find that households with few liquid assets were the most responsive to the payments.

Orchard, Ramey, and Wieland (2025) argue that the high transfer multipliers associated with heterogeneous agent models produce implausible counterfactuals for no-stimulus scenarios following the 2008 Global Financial Crisis. The authors argue this is because the MPC estimates used to calibrate them are upwardly biased; many popular two-way fixed effect differences-in-differences MPC estimates inadvertently mix treatment and control groups when the control group receives a delayed treatment, while the social security numbers used to distribute stimulus checks are not entirely random. Notably, though, the lottery data MPC estimates of Fagereng, Holm, and Natvik (2021) does not have this econometric problem – although Orchard, Ramey, and Wieland (2025) note that the former must impute consumption for its estimates using wealth and income data. Even so, as a high-frequency event study that observes the transfers directly the aforementioned Baker et al. (2023) does not have "forbidden comparison" concerns and observes consumption expenditures directly.

iii. Solving HANK models with Sticky Expectations

As alluded to in the introduction, this paper is an empirical application of the solution method employed in Kwicklis (2025b), which shows that augmenting a FIRE system with the average expectation of the economy's dynamics can be used to convert the solution to represent a sticky expectation one. This approach naturally builds off of previous work on solving FIRE state-space HANK models, such as Bayer and Luetticke (2020), which in turn is a powerful refinement of Reiter (2009). I also draw from the work of Achdou et al. (2021) and Ahn et al. (2018), which explains steady-state perturbation solutions in continuous time HANK models more broadly. For converting the continuous time model to the discretely sampled data frequencies, I rely on the methods employed by Christensen, Neri, and Parra-Alvarez (2024), which provides a guide for properly integrating continuous time equations to discretized measurements of stocks and flows.

2 The Medium-Scale HANK Model

In this section, I outline the medium scale model that I estimate using the original 7 Smets and Wouters (2007) series, plus data on unemployment and transfers as a share of GDP. Most households participate in an incomplete markets setting, but the profits from managing the physical capital stock are rebated to a representative financier agent that follows a standard Euler equation. My framework is therefore conceptually similar to two agent New Keynesian (TANK) models – except instead of a hand-to-mouth mass of households, the economy has an entire heterogeneous agent block whose decisions dominate the consumption dynamics, as the financiers receive only 6% of aggregate income and have strong smoothing motives.

Time $t \geq 0$ is continuous. Markets are incomplete for an ex-post heterogeneous continuum of households, who pay and receive taxes and transfers to and from the government and trade a single non-contingent liquid asset (government bonds) to smooth consumption and insure themselves from idiosyncratic shocks to productivity and employment. Employed households earn wages in the labor market, while unemployed households receive unemployment benefits that are a fraction of their steady-state wage rate. All households work the same number of hours. Each household's expectations about macroeconomic variables are sticky, such that they only update their forecasts of macroeconomic variables periodically.

A frictional labor market connects unemployed job seekers to vacancies posted by labor agencies using an aggregate matching function. Wages are set via a flexible rule that slowly adjusts to labor market tightness, hours worked per worker, past wages, and inflation. Labor agencies then rent labor to monopolistically competitive intermediate firms with nominal rigidities, who produce a differentiated output good by combining effective hours worked and rented capital. These intermediate outputs are combined by final goods sector, which sell the resulting composites to households for consumption.

The capital stock is owned and managed by a competitive mutual fund sector, which pays adjustment costs for changes to the path of investment and utilization costs for the intensity of capital usage. These mutual funds are in turn owned by a representative block of financier households, who have access to complete markets and smooth consumption according to an Euler equation. Since financiers can participate in all markets, a single rate of return prevails in the economy. In addition, I assume that the financiers are attentive and operate with full information and rational expectations.

The government consists of a treasury and a central bank. The treasury pursues passive fiscal policy in the sense of Leeper (1991), but pays down its debt only very slowly. Similarly, the central bank pursues active monetary policy by adjusting nominal interest rates more than one-for-one with inflation according to a smoothed Taylor rule.

In what follows, I detail the model as it would be solved under a rational expectations solution, with the understanding that the linearized FIRE solution can be converted into a sticky expectations one using the methodology of Sections 2 and 3.

2.1 Households

A continuum of households is ex-post heterogeneous in terms of their liquid (risk-free) liquid real assets a, their idiosyncratic labor productivity z, and their employment state $\ell \in \{E, U\}$, where E stands for an employed labor market state and U stands for an unemployed labor market state. Idiosyncratic labor market productivity evolves according to a log Ornstein-Uhlenbeck process driven by a Brownian motion W_t:

$$d \log z_t = \theta_z \log z_t + dW_t.$$

Post-tax labor market income is denoted y_t, which depends on the household's labor market state and labor-augmenting productivity:

$$y_t(a, z, l) = \begin{cases} (1-\tau) z w_t h_t & \text{if } l = E, \\ \kappa_u (1-\tau) z w_{NSS} & \text{if } l = U \end{cases}$$

such that unemployment benefits are set via κ_u to replace 40% of household labor income in steady-state. Lastly, households earn a rate of return of $\tilde{r}_t$ on their bonds, which is equal to

$$\tilde{r}_t = r_t + \zeta_{r,t}$$

where r_t is the real interest rate related to the central bank's policy rate i_t via the Fisher relation $r_t = i_t - \pi_t$. $\zeta_{r,t}$ is analogous to the bond premium shock of Smets and Wouters (2007).

In this framework, households maximize their expected constant relative risk averse (CRRA) utility with a risk aversion coefficient of γ. A household with beliefs indexed by i will have a value function that takes the form

$$\begin{aligned} V_t(a, z, \ell) = \max_{(c_s^i)s \geq t} \mathbb{E}_t^i &\left[\int_t^\infty e^{-\rho(s-t)} \frac{(c_s^i)^{1-\gamma} - 1}{1-\gamma} ds \right] \\ \text{s.t. } \frac{\mathbb{E}_t^i[da_t]}{dt} &= y_t(z, l) + \tilde{r}_t a_t - c_t^i + T(a, z) \\ d \log z_t &= \theta_z \log z_t + dW_t. \end{aligned}$$

Using the methodology of Sections 2 and 3, it suffices to write the households' full information rational expectations (FIRE) problem, with the understanding that it can be later converted into the sticky expectation framework. Households *believe* that they are solving the following Hamilton Jacobi Bellman (HJB) equation:

$$
\begin{aligned}
\rho V_t(a,z,l) = \max_c & \left\{ \frac{c^{1-\gamma}-1}{1-\gamma} + \partial_a V_t(a,z,l)(\tilde{r}_t a + y_t(a,z,l) - c) + T_t(a,z) \right\} \\
& + \partial_z V_t(a,z,l) z \left[\frac{1}{2}\sigma_z^2 - \theta_z \log(z) \right] + \frac{1}{2}\sigma_z^2 z^2 \partial_z^2 V_t(a,z,l) \\
& + \lambda_{ll',t}[V_t(a,z,l') - V_t(a,z,l)] + \partial_t V_t(a,z,l) \\
\text{s.t. } & a \geq 0.
\end{aligned} \tag{1}
$$

In the FIRE setting, the distribution of households over the idiosyncratic state space will evolve according to the Kolmogorov Forward Equation

$$
\partial_t \mu_t(a,z,l) = -\frac{\partial}{\partial a}\left[\frac{da_t}{dt}\mu_t(a,z)\right] - \frac{\partial}{\partial z}\left[\frac{\mathbb{E}_t[dz_t]}{dt}\mu_t(a,z)\right] - (\lambda_{l,l',t}\mu_t(a,z,l) - \lambda_{l',l,t}\mu(a,z,l')). \tag{2}
$$

2.2 Final and Intermediate Goods Firms

As is standard in the New Keynesian literature, a competitive final goods sector with firms indexed by $j \in [0,1]$ produces output with a constant elasticity of substitution ε:

$$
Y_t = \left[\int y_t(j)^{\frac{1-\varepsilon}{\varepsilon}} dj\right]^{\frac{\varepsilon}{1-\varepsilon}}
$$

where monopolistically competitive intermediate goods firms hire capital K_t^f from a mutual fund and and labor L_t from a search-and-matching sector to produce output $y_t(j)$, such that firm j produces:

$$
y_t(j) = e^{\zeta_t^{TFP}}[K_t^f(j)]^\alpha L_t(j)^{(1-\alpha)}.
$$

Intermediate good firms choose their prices while internalizing the demand for their output and price adjustment costs as in Rotemberg (1982). To generate price indexing in continuous time, I assume that firms only pay to adjustment costs that generate a deviation of their price growth from the previous backward-looking trend of aggregate inflation. That is, firms pay costs

$$
\Theta\left(\frac{\dot{p}_t(j)}{p_t(j)} - \bar{\pi}_t\right) = \frac{\theta_\pi}{2}(\pi_t(j) - \bar{\pi}_t)^2 Y_t
$$

where $\bar{\pi}_t$ is a moving exponential average of past aggregate inflation with decay rate κ_π:

$$
\frac{d\bar{\pi}_t}{dt} = -\kappa_\pi(\bar{\pi}_t - \pi_t). \tag{3}
$$

The recursive problem for profit maximizing firms is then the HJB

$$r_t J_t^F(p,\bar{\pi}_t) = \max_{\pi_t(j)} \Bigg\{ \left(\frac{p_t(j)}{P_t} - m_t\right)\left(\frac{p_t(j)}{p_t}\right)^{-\varepsilon} Y_t - \frac{\theta_\pi}{2}(\pi_t - \bar{\pi}_t)^2 Y_t$$
$$+ \partial_p J_t^F(p_t(j),\bar{\pi})\pi_t(j)p_t(j) + \partial_{\bar{\pi}} J_t^F(p_t(j),\bar{\pi}_t)\kappa_\pi(\pi_t - \bar{\pi}_t) + \partial_t J_t(p_t(j),\bar{\pi}_t)\Bigg\}.$$

In the model appendix, I show that the symmetric equilibrium yields a backward-indexed Phillips Curve:

$$\frac{\mathbb{E}_t[d\pi_t]}{dt} = r(\pi_t - \bar{\pi}_t) + \frac{\varepsilon}{\theta_\pi}\left[\frac{\varepsilon - 1}{\varepsilon} - \mathrm{mc}_t\right] - \kappa_\pi(\bar{\pi} - \pi) + \zeta_{\pi,t} \tag{4}$$

where mc_t is the marginal cost of production $\mathrm{mc}_t(j) = \exp(-\zeta_t^{\mathrm{tfp}})\left(\frac{r_t^k}{\alpha}\right)^\alpha\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}$. Lastly, $\zeta_{\pi,t}$ is a price markup shock, again in the style of Smets and Wouters (2007).

2.3 The labor market

Search in the labor market is facilitated with a standard Cobb-Douglas constant returns to scale matching function, as in much of the literature based on Diamond (1982) and Mortensen and Pissarides (1994). Writing $\iota < 1$ as the elasticity of matches with respect to the unemployment rate, the job finding rate $\lambda_{UE,t}$ and the vacancy filling rate $\lambda_{v,t}$ are respectively

$$\lambda_{UE,t} = \Psi\theta_t^{1-\iota}, \quad \lambda_{v,t} = \Psi\theta_t^{-\iota} \tag{5}$$

where Ψ is the efficiency of the matching function and $\theta_t \equiv \frac{v_t}{u_t}$ is labor market tightness, the ratio of the mass of vacancies v_t to the unemployment rate u_t.

Hiring in the labor market is also managed by identical labor agencies indexed by $j \in [0,1]$. Each agency hires such that their number of employees $N_t(j)$ evolves according to

$$\dot{N}_t(j) = \lambda_{v,t}v_t(j) - \lambda_{EU}N_t(j)$$

Worker hours are differentiated with a CES demand system by intermediate firms, in which ε_ℓ is the elasticity of substitution between workers. If hiring firm j internalizes a downward-sloping demand curve for its laborers' hours $h_t(j)$, then a CES demand system implies

$$h_t(j) = \left(\frac{r_t^l(j)}{r_t^l}\right)^{-\varepsilon_\ell} h_t.$$

From there, I assume a labor market hiring structure that is similar to Bardóczy and Guerreiro (2024), which in turn is based off of Christiano, Eichenbaum, and Trabandt (2016): a representative labor agency firm pays costs η_v for creating a vacancy, and the cost η_l for actually filling the vacancy (essentially, the on-boarding costs of hiring). Besides the imperfect substitutability of workers' hours, I make one notable addition to dampen firms' ability to regulate the hours of existing workers instead of adding new hires: I add a quadratic adjustment cost to deviations in hours worked from the non-stochastic steady state (normalized to 1). These costs stand in for the negotiating costs that keep firms from effortlessly changing the number of labor hours a job demands, forcing it to hire or fire rather than instantly restructure employment agreements. The recursive problem for profit-maximizing labor agencies is therefore

$$
\begin{aligned}
r_t J_t(N_t(j)) = \max_{v_t(j), r_t^l(j)} \Bigg\{ & (r_t^l(j) - w_t) \left(\frac{r_t^l(j)}{r_t^l} \right)^{-\varepsilon_l} N_t(j) Z_t h_t - \frac{\phi_h}{2} (h_t(j) - 1)^2 N_t(j) Z_t \\
& - \eta_v v_t(j) - \eta_l \lambda_{v,t} v_t(j) + J_t'(N_t(j))[\lambda_{v,t} v_t(j) - \lambda_{EU} N_t(j)] + \partial_t J_t(N_t(j)) \Bigg\}.
\end{aligned}
$$

In the Appendix A.1, I show that the vacancy supply problem in a symmetric equilibrium yields the dynamic equation

$$
\frac{\mathbb{E}_t[d\theta_t]}{dt} \frac{1}{\theta_t} = \frac{1}{\iota}(r_t + \lambda_{EU}) \left[1 + \frac{\eta_l}{\eta_v} \lambda_{v,t} \right] - \frac{1}{\iota}(r_t^l - w_t) h_t \left(\frac{\lambda_{v,t}}{\eta_v} \right) \tag{6}
$$

while the rental rate of labor will satisfy

$$
r_t^l = \frac{\varepsilon_l}{\varepsilon_l - 1}(w_t + \phi_h(h_t - h^*)). \tag{7}
$$

Similar to Bardóczy and Guerreiro (2024), I assume real wage growth follows an ad-hoc specification:

$$
\frac{dw_t}{dt} \frac{1}{w_t} = (1 - \rho_w) \left(\phi_{w,\theta} \Delta\theta_t + \phi_{w,h} \Delta h_t - \Delta w_t \right) - \rho_w \pi_t + \zeta_{w,t} \tag{8}
$$

where Δ variables represent deviations from the non-stochastic steady-state, ρ_w governs the smoothness of wage adjustment and the degree of erosion from inflation, and the $\phi_{w,h}, \phi_{w,\theta}$ coefficients regulate the passthrough of hours worked and labor market tightness into real wage growth. Lastly, $\zeta_{w,t}$ is an additive wage markup shock.

2.4 Capital Mutual Fund with Adjustment Costs

A representative competitive mutual fund sector manages the aggregate capital stock of the economy. To keep the setting similar to Smets and Wouters (2007) and previous papers in the literature, this firm pays adjustment costs for altering the rate of aggregate investment I_t relative to a backward-looking moving average $\overline{I}_t$, with the costs denoted $\Phi(I_t, \overline{I}_t)$. It additionally pays costs $\psi_t(u_t^k)$ for utilizing capital, where χ_t is the capital utilization rate. ζ_t^I denotes a shock to the rate of transformation of investment into capital. As such, the representative mutual fund solves

$$
\begin{aligned}
\Pi_0(I_0, K_0) = \max_{(I_t, u_t)_{t \geq 0}} \ & \mathbb{E}_0 \int_0^\infty e^{-\int_0^t r_\tau d\tau} \Big[(r_t^k \chi_t K_t - I_t - \psi(\chi_t) K_t \Big] dt \\
\text{s.t.} \ & \frac{dK_t}{dt} = e^{\zeta_t^I} [1 - \Phi(I_t, \bar{I}_t)] I_t - \delta K_t \\
& \frac{d\bar{I}_t}{dt} = (I_t - \overline{I}_t)
\end{aligned}
$$

For tractability, I assume that the adjustment costs are quadratic and that the elasticity of utilization costs relative to steady-state utilization (normalized to 1) is $(1-\nu)/\nu$. As shown in Appendix A.2, the resulting dynamic equations are then

$$\begin{aligned} r_t^k =& \psi'(\chi_t), \\ 1 =& q_t^k e^{\zeta_t^I} \left[1 - \frac{\phi_I}{2} \left(\frac{x_t}{\overline{x}_t} - 1 \right)^2 - \phi_I \left(\frac{I_t}{\overline{I}_t} - 1 \right) \frac{I_t}{\overline{I}_t} \right], \\ \frac{\mathbb{E}[dq_t^k]}{dt} =& (r_t + \delta) q_t^k - r_t^k u_t^k + \psi(u_t^k). \end{aligned} \tag{9}$$

Here, q_t^k denotes Tobin's marginal Q.

2.5 Financier Agents

Although the capital stock is managed by a mutual fund, that mutual fund receives a small share of GDP in profit, which must be received as income by another actor in the economy. This presents a challenge for single-asset HANK models like mine: the present value of these profits are substantial in the non-stochastic steady-state. If claims to these capital profits are traded as assets by the heterogeneous agents, then this substantially raises the amount of overall wealth available to households in the economy. As shown in Auclert, Rognlie, and Straub (2024), the households are then better able to insure themselves against adverse labor market outcomes and hitting their borrowing constraints, which significantly lowers average MPCs to the point that they no longer match the empirical estimates in Fagereng, Holm, and Natvik (2021).

In order to match both the large amounts of capital and the high sensitivity of the average household to current income, I introduce a new representative "financier" agent, which does not supply labor, owns the mutual fund that owns the capital stock, and receives the profits from the fund's activities. Only the financier can hold claims to capital; the heterogeneous agent households cannot, and can instead only hold the bonds issued by the government as assets. I assume that financiers do not hold bonds in the non-stochastic steady-state, but they can participate in all asset markets, such that no arbitrage equates the ex-ante real rate of return across the economy.

Financier households (aggregated to a representative agent) closely resemble the representative consumers found in standard complete-markets representative agent New Keynesian (RANK) models like Smets and Wouters (2007): they are infinite horizon utility optimizers who choose their consumption c_t^{FIN} to maximize their constant relative risk-averse (CRRA) preferences, with a same degree of relative risk aversion as the heterogeneous agent households. They discount the future at a rate of ρ^{FIN} that matches the steady-state real interest rate r. Additionally, they value consumption relative to a β share of an external habit term $\overline{c}_t^{\text{FIN}}$ that is a moving average of the past consumption of other financiers.

The setup and solution of the financier agents' decision problem is in Appendix A.3. As such, they choose their consumption in accordance with an Euler equation with habit formation:

$$\frac{\mathbb{E}_t[dc_t^{\text{FIN}}]}{dt} = \frac{1}{\gamma}(r_t - \rho^{\text{FIN}})(c_t^{FIN} - \beta\overline{c}_t^{\text{FIN}}) - \beta^{\text{FIN}}(\overline{c}_t^{\text{FIN}} - c_t^{\text{FIN}}) \tag{10}$$

where the external habit $\overline{c}_t^{\text{FIN}}$ accumulates according to

$$\frac{d\overline{c}_t^{\text{FIN}}}{dt} = (c_t^{\text{FIN}} - \overline{c}_t^{\text{FIN}}). \tag{11}$$

As well-insured households with habit formation, financiers have very small marginal propensity to consume. Combined with the fact that the profits they consume are small relative to the size of the economy to begin with (less than 5% of GDP in the steady-state), this implies that they contribute very little to aggregate demand dynamics. This enables me to keep the path of aggregate consumption largely determined by the realistically calibrated heterogeneous agent households, while still matching the large ratio of assets to income seen in the real data.

2.6 Government Policy

2.6.1 Monetary Policy

Monetary policy is managed by a central bank according to a standard delayed Taylor rule. Nominal interest rates i evolve according to

$$di_t = \rho_i(r_{nss} + \phi_\pi \pi_t + \zeta_{MP,t} - i_t)dt + d\delta_{MP,t}. \tag{12}$$

For flexibility, nominal interest rates are affected both gradually through $\zeta_{MP,t}$ and immediately through $d\delta_{MP,t}$, the initial condition of $\zeta_{MP,t}$; the latter allows the interest rate to jump on impact. Monetary policy's responsiveness to inflation is moderated by the Taylor rule parameter ϕ_π. The smoothing parameter ρ_i governs the baseline speed at which interest rates catch up to what a static Taylor rule would proscribe.

As a baseline, I consider active monetary policy in the sense of Leeper (1991) by setting $\phi_\pi = 1.5$, a standard value in the literature. To consider the model economy under an alternative "passive" monetary policy configuration, I consider an interest rate peg that sets $\phi_\pi = 0$.

2.6.2 Fiscal Policy

Fiscal policy is managed by the Treasury, which is also assumed to follow a rule to stabilize the real value of government debt – or not, depending on the policy configuration. The government can also exogenously induce transfer "shocks" ζ_T, effectively unforeseen stimulus payments sent to the heterogeneous household block.

The total government debt position B_t evolves according to

$$\frac{dB_t}{dt} = [-(\text{Tax}_t - G_t) + r_t B_t]dt \tag{13}$$

where Tax_t represents net tax revenue to the government and G_t represents real government expenditures. Government expenditures G are equal to their steady-state level times a mean-reverting shock $\zeta_{G,t}$:

$$G_t = G_{NSS} e^{-\zeta_{G,t}}. \tag{14}$$

Net taxes consists of labor income taxes, minus unemployment insurance claims, minus total transfers to the public T_t:

$$\text{Tax}_t = \underbrace{\tau w_t L_t}_{\substack{\text{Income}\\ \text{taxes}}} - \underbrace{\kappa_u w_{nss} L_{nss}}_{\text{UI claims}} - \underbrace{\sum_\ell \int \int T_t(a,z)\mu_t(a,z,\ell) da\ dz}_{\text{Total transfers}} \tag{15}$$

Non-UI transfers are rebated to households in the joint wealth and income space according to

$$T_t(a,z) = T_{nss} - \kappa_{Fiscal} \frac{z}{Z_t}(B_t - B_{nss}) + 4Y_{nss} \times \zeta_T. \tag{16}$$

Equation (16) implies that the government automatically adjusts the net lump-sum transfers that it pays out to households according to a rule that stabilizes the government's debt asymptotically so long as κ_{fiscal} sufficiently exceeds the government's interest rate r_t. To reduce the effect of the transfers on the agents, however, the reduction in transfers (effectively, the debt stabilization tax) is proportional to the household's steady-state labor income. Steady-state transfers are set to balance the budget in the absence of macroeconomic shocks.

Transfer shocks $\zeta_{T,t}$ to all (non-financier) households are scaled by four times steady-state output. As such, a shock of $\zeta_t = 0.01$ is equivalent to transfers worth 1% of steady-state real GDP within the time increment.

2.7 Market Clearing

For the labor market to clear, total labor demand must equal effective aggregate labor supply:

$$L_t = \int_0^1 L_t(j)dj = \int \int z_t h_t \mu_t(a,z,\ell = E) da\ dz \tag{17}$$

Note that h_t is not indexed by the idiosyncratic variables, as all employed households work the same number of hours to satisfy labor demand. Similarly, the unemployment rate u_t must aggregate from the heterogeneous agents' distribution:

$$u_t = \int \int \mu(a,z,\ell = U) da\ dz$$

The capital market similarly clears the capital held by the mutual fund times that capital utilization rate equals the total physical capital K_t^f used by the intermediate firm production sector:

$$K_t^f = \int_0^1 K_t^f(j)dj = \chi_t K_t. \tag{18}$$

For the goods market to clear, total aggregate expenditures must equal aggregate income:

$$Y_t = (C_t + c_t^{\text{FIN}}) + I_t + G_t. \tag{19}$$

As in Hagedorn, Manovskii, and Mitman (2019), I assume that the frictional capital utilization and Rotemberg adjustment costs are "virtual," and do not contribute to the resource constraint or aggregate expenditures. Equivalently, one could assume that adjustment the costs for some firms are pure profit for others and therefore amount to a transfer that is net zero out when rebated back to households or financiers.

By Walras' law, the bond market will also clear, completing the mathematical description of the model.

3 Calibration and Estimation

In this section, estimation comes in three parts. First, I calibrate the model to be in line with the existing literature in the non-stochastic steady-state. A key feature is that the model matches the profile of marginal propensities to consume of Fagereng, Holm, and Natvik (2021). In the second step, I estimate the structural parameters of the model by matching the IRFs of the model to local projections estimates of a narratively identified monetary policy shock. This ensures that the model roughly matches the empirical dynamics of output, consumption, investment, inflation, unemployment, hours worked, and wage growth, for at least the effects of monetary policy. In the third step, I estimate the remaining parameters related to the stochastic shocks using a Bayesian method known as Sequential Monte Carlo (SMC).

3.1 Calibration and the Non-Stochastic Steady-State

Table 1 details the parameters of the model that are directly calibrated, along with the source for the parameters or the targeted values of the calibration. Most of these pertain to terms that directly influence the non-stochastic steady-state. The model is calibrated for a quarterly frequency.

In the heterogeneous household block, ρ is set so that the asset market clears when $r = 2\%$ annually. The idiosyncratic income process mean reversion term θ_z and variance term σ_z^2 are calibrated to match micro data on the autocovariance of wages detailed in Floden and Lindé (2001), as is done in McKay, Nakamura, and Steinsson (2016) and Kwicklis (2025a).

The labor market parameters are targeted so that the steady-state labor market tightness is $\theta_{nss} = 1$. Under this assumption, the parameter governing the efficiency of the Cobb-Douglas matching function in the labor market may be read as exactly the inverse of the average unemployment duration. By virtue of matches occurring according to a Poisson process, unemployment spells in the model are exponentially distributed. I pick an average duration of 1.15 quarters (14.95 weeks), roughly the average duration of the 1990s according to data from the Current Population Survey. For the match function's elasticity with respect to the unemployment rate, I use the empirical estimates of Pissarides (1986). I pick the separation rate to target a 4% level of unemployment in the absence of aggregate shocks; the average worker stays at a job for 27.6 quarters.

I assume that total vacancy posting costs are equal to 11% of the average worker's quarterly wage bill. These vacancy costs are split 51%-49% between the regular posting cost η_v and the vacancy filling cost η_l.

For the quadratic hours adjustment costs, I set ϕ_h to be 0.5. This is roughly the value that generates a transitory 5 basis point reduction in the trough of the unemployment rate following a 25 basis point reduction in the policy rate, a value roughly consistent with the empirical impulse response functions of the next section.

Table 1: Calibrated HANK Model Parameters

Parameter	Symbol	Value	Source or Target
Households			
Internally Calibrated:			
Quarterly Time Discounting	ρ	0.049	$r = 2\%$ Annually
Idiosyncratic Income Shock Variance	σ_z^2	0.017	Floden and Lindé (2001)
Idiosyncratic Shock Mean Reversion	θ_z	0.034	Floden and Lindé (2001)
Elasticity of Intermediate Subst.	ε	10	10% profit share
Assumed from Literature:			
Relative Risk Aversion	γ	2.0	McKay et al (2016)
Labor Market			
Matching function efficiency	Ψ	0.86	Avg. unemp. duration of $\approx$ 15 weeks
Elasticity of matching w.r.t. u	ι	0.70	Pissarides (1986)
UI replacement rate	κ_u	0.40	UI benefits pay 40% of wages
Flat vacancy posting cost	η_v	0.10	5.7% of quarterly wage bill
Intensive vacancy posting cost	η_l	0.1218	5.4% of quarterly wage bill
Exogenous job loss rate	λ_{EU}	0.0362	Target steady-state u of 4%
Labor adjustment cost parameter	ϕ_h	0.5	Rate hike to peak unemp. passthrough of 20%
Capital Market and Production			
Capital depreciation rate	δ	0.02075	8.65% annual depreciation rate
Elasticity of Y w.r.t. K	α	0.33	60% labor share of income
Capital utilization cost	ι	0.45	Elasticity of u_t^k to r_t^k of 0.81
Financier Agents			
Financier time discounting	ρ^{FIN}	0.005	$r = 2\%$ annually
Financier habit weight	β^{FIN}	0.70	Smets and Wouters (2007) prior mean
Government			
Steady state government debt	B_{NSS}	2.37	HANK $iMPC_0 \approx 0.45$
Income Tax Rate	τ	0.30	Approx. US tax wedge (OECD 2025)
Government share of expenditure	0.15	G/Y	
Taylor rule coefficient	ϕ_π	1.5	Active monetary policy
Fiscal adjustment coefficient	κ_{fiscal}	0.05	Passive fiscal policy

The marginal distribution of liquid assets is depicted in the first panel of Figure 2. The unemployed households' asset distribution (in red) is shifted significantly closer to the borrowing constraint compared to the employed population (in black), where 19% of unemployed households in the model have completely exhausted their savings. The total stey-state bond position $B_{NSS} = \int\int\int a\mu_{NSS}(a,z,l)\ da\ dz\ dl$ of the heterogeneous agent block is chosen to bring the average intertemporal MPC (iMPC) out of a one-time increase in liquid wealth to 0.44 after one year, close to the estimates of Fagereng, Holm, and Natvik (2021) reproduced in Auclert, Rognlie, and Straub (2024). As shown in the second panel of Figure 2, the iMPCs decline thereafter, also broadly in line with the empirical estimates. In the last panel of Figure 2, I document how the households' one-year MPCs out of liquid wealth in the cross section decline in liquid wealth position and labor market income.

Outside of the non-stochastic steady-state, I study the conventional active monetary/passive fiscal framework, in the terminology of Leeper (1991). As such, I set the Taylor rule coefficient to $\phi_\pi = 1.5$, a standard value set in the literature. For a stationary linearized solution to exist, the government must then passively adjust its tax schedule to keep real government debt from exploding. As such, I set the fiscal adjustment parameter to $\kappa_{fiscal} = 0.05$, such that government debt above steady-state levels is paid back by raising net taxes only very gradually over time.

3.2 IRF Matching using a Monetary Policy Shock

I integrate the model to a discrete-time quarterly level using the methodology of Christensen, Neri, and Parra-Alvarez (2024). As in Christiano, Eichenbaum, and Evans (2005) and many other papers, I fit many of the model's remaining structural parameters by matching the model's impulse response functions (IRFs) to empirical estimates of the effect of a monetary policy shock. I follow a similar strategy as in Auclert, Rognlie, and Straub (2020): I use an identified C. D. Romer and D. H. Romer (2004) narrative shock to calculate the monthly impulse response function using the local projections method of Jordá (2005). As in Ramey (2016), I control for two lags of the narrative shock, the unemployment rate, the CPI, the GDP deflator, the commodity Producer Price Index, and an index of total U.S. industrial production, and the federal funds rate (FFR). I also include contemporaneous levels of the macroeconomic variables, except for the FFR (and, of course, the Romer and Romer (2004) shock). All real variables are deflated using the GDP deflator.

For the dependent variables in the projection, I use the logs of real GDP as total output, log total real personal consumption expenditures (PCE) as consumption, the log of real investment as I, the log of the PCE price index as p, and the log of real nonfarm hourly compensation and total nonfarm hours worked as w and h. I use the headline unemployment rate to measure u. I measure nominal interest rates using the end-of-month FFR. Like Auclert, Rognlie, and Straub (2020), I log-linearly interpolate variables reported at the quarterly frequency to obtain monthly series when necessary.

Gathering the controls in X_t, labeling the Romer and Romer (2004) shock as RR_t, and denoting

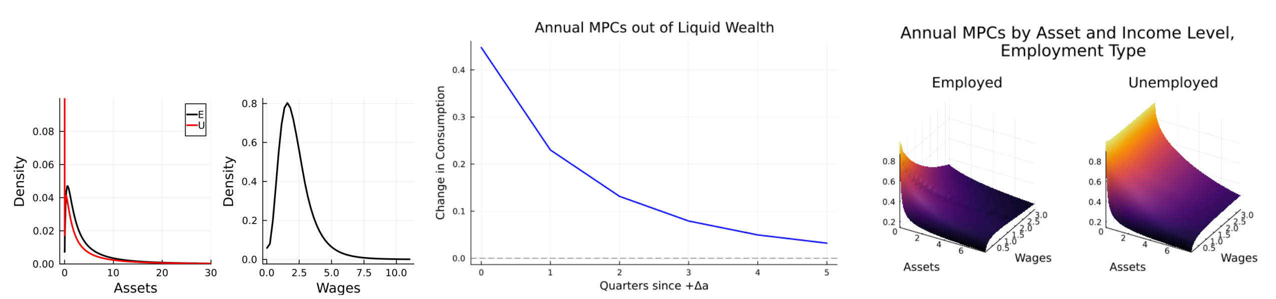

Figure 2: Leftmost panel: the distribution of assets and wages in the non-stochastic steady-state for employed workers (E) and unemployed households (U). Middle panel: the intertemporal MPCs (iMPCs) of households in the model out of liquid wealth after a one-time transfer at time 0 in the non-stochastic steady-state. Rightmost panel: the 1-year iMPCs out of liquid wealth in the non-stochastic steady-state, by liquid asset position, income, and employment type.

Table 2: HANK non-stochastic steady state quantities

Description	Symbol	Value
First year iMPC out of liquid assets		0.443
Debt to Annual Income	$B_{NSS}/(4Y_{NSS})$	0.170
Correlation btw. Income and Assets	$\text{Corr}(a, z)$	0.688
Share of households with $a = 0$	$\int \mu_{NSS}(0, z)dz$	0.020
Asset Gini Coefficient		0.648
Income Gini Coefficient		0.334
Unemployment Rate	u_{nss}	0.040
Capital-to-Income Ratio	$K_{nss}/(4Y_{nss})$	2.883

outcome j as Y_t^j, the local projection for outcome variable j at forecast h thus takes the form

$$Y_{t+h}^j = RR_t\beta_h^j + X_t'\eta_h^j + \varepsilon_{t,h}^j.$$

Like Auclert, Rognlie, and Straub (2020), I use the 1969-1996 time period – the time period covered by the original Romer and Romer (2004) paper – to generate my estimates. $\widehat{\beta}_h^j$ is then my estimate of the IRF of outcome j at horizon h; I collect these estimates into a stacked empirical IRF vector $\widehat{\text{IRF}}$.

After obtaining the empirical estimates of the effect of a monetary shock, I estimate the model's parameters by minimizing a quadratic loss function. Collecting the parameters to be estimated in Θ and writing the model impulse response function as $\text{IRF}(\Theta)$, the estimated parameters solve the problem

$$\widehat{\Theta} = \arg\min_{\Theta} \, (\text{IRF}(\Theta) - \widehat{\text{IRF}})'\widehat{\Sigma}^{-1}(\text{IRF}(\Theta) - \widehat{\text{IRF}}).$$

Here, the $\widehat{\Sigma}$ matrix contains the squared Newey and West (1987) standard errors of the empirical local projections estimates on its main diagonal and is zero elsewhere.

The GMM estimates of the model parameters are reported in Table 3. The local projection impulse response functions are presented in Figure 3; the shaded region represents a 95% confidence interval while the model impulse response functions are depicted in dark orange. All impulse responses are scaled to match a 25 basis point initial drop in the federal funds rate.

Table 3: Estimated Model Parameters to Match IRFs

Parameter	Symbol	Estimate	S.E.
Household update rate	λ	0.266	(0.115)
Monetary policy smoother	ρ_i	0.375	(0.028)
Monetary policy shock persistence	η_{MP}	0.0421	(0.0088)
Capital Adjustment Cost	ϕ_k	8.563	(0.284)
Rotemberg Inflation Cost	$\theta_\pi/100$	37.829	(0.145)
Real wage smoother	ρ_w	0.0912	(0.0363)
Tightness to real wage feedback	$\phi_{w\theta}$	0.00350	(0.00203)
Hours to real wage feedback	ϕ_{wh}	0.813	(0.120)

Notably, the learning rate suggests that roughly a quarter of households have updated to a macroeconomic shock after one quarter, 40% have updated after two quarters, and nearly 65% have updated by the end of the year, a faster pace of learning than that estimated in Auclert, Rognlie, and Straub (2020).

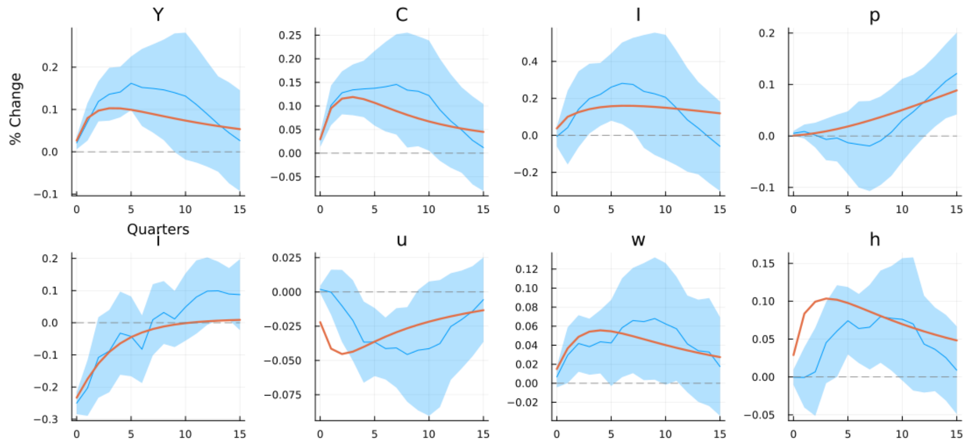

Figure 3: Model and local projection impulse response functions to a 25 basis point reduction in the nominal interest rate

The square root of the average mean square deviation between the orange and blue lines of 3 is 1.30 standard errors. In particular, the model struggles to fit the slow fall in unemployment captured by the local projection – hence why I calibrate ϕ_h to roughly match the magnitude of the drop, if not its timing. The model also generates an enormous amount of nominal rigidity, in order to acquiesce with the slow rise in the price level documented in the empirical estimates (a feature also depicted in the estimates of Auclert, Rognlie, and Straub (2020)). This implies that the effect of future output gaps on inflation movements is slow and muted in general.

3.3 Sequential Monte Carlo Estimation

Having estimated a subset of the structural model parameters using GMM and the empirical response to an identified monetary policy shock, I estimate the persistences and variances of the remaining shocks using a Bayesian technique known as Sequential Monte Carlo (SMC). SMC has recently gained prominence as a technique to estimate macroeconomic models; Herbst and Schorfheide (2014) demonstrates that SMC can better sample from the potentially irregularly shaped and multi-modal posteriors that arise in DSGE settings, while Cai et al. (2020) details how SMC can be used for efficient real-time estimation and forecasting. Acharya et al. (2023) uses the algorithm to estimate a full information medium-scale HANK model.

The reader can consult these papers for an in-depth technical discussion, but as a broad overview: SMC combines importance sampling with Metropolis-Hasting Markov Chain Monte Carlo methods and particle swarm techniques to sample from the model's posterior distribution. The user draws a large number of initial values ("particles") from the parameter space, often distributed according to the models' prior distribution, and constructs importance weights to sample from a series of bridge distributions. These bridge distributions are typically a log-linear weighting of the prior and likelihood that strongly weight the prior at initialization and gradually converge upon the model's true posterior. At each iteration, the algorithm conducts a Metropolis-Hastings step to change ("mutate") each of the particles stochastically – but weighted in the direction of increasing probability mass. The user periodically re-samples the particles from their distribution (constructed using their importance weights) in order to prevent any one particle from becoming too dominant. By the time the target distribution converges to the posterior, the particle swarm and its weights should be distributed according to the posterior, allowing the researcher to evaluate moments of the posterior by appealing to the law of large numbers.

I use the variant of the adaptive SMC algortithm presented in Cai et al. (2020) and employed in Acharya et al. (2023), which tunes the rate of bridge distribution convergence according to the effective sample size of the particles. I use standard state space Gaussian methods (e.g., the Kalman filter) to evaluate my likelihood function given the observable data. An advantage of SMC, compared to other Markov Chain Monte Carlo (MCMC) methods is that in addition to being robust to local but suboptimal posterior modes, the algorithm is highly parallelizable. For medium-scale HANK models where each likelihood evaluation can take a second or more, this is a highly useful property.

My measurement equation is similar to the one used in Acharya et al. (2023) and the classic Smets and Wouters (2007): I include i) real GDP per capita growth, ii) real personal consumption expenditures per capita growth, iii) real fixed private investment per capita growth (investment), iv) the change in the GDP deflator for inflation, v) the growth rate of real hourly compensation per worker, vi) the level of average weekly nonfarm hours per worker, and the level of the Federal Funds rate. The details of the construction of each series, taken from the Federal Reserve Economic Data (FRED) repository, are provided in the appendix. Each series in the measurement equation enters the model with its mean removed, as in Acharya et al. (2023) and Bayer, Born, and Luetticke (2024), to represent deviations from a log-linear balanced growth path and to avoid questions of balanced growth preferences in the HANK structure. In the training stage of the model, all data are from the United States over the period of 1965 to 2007, before the Great Financial Crisis, to assess the fit of the model in times of relative normalcy. I include post-crisis and post-pandemic data only to assess the model's fit over the latter periods.

In addition to the seven standard series, I introduce two new ones, which my model is able to describe: viii) the unemployment rate and ix) the level of transfers from the government to households as a share of GDP. I construct this last series as the current dollar value of government social benefits minus spending on Social Security, Medicare, and Medicaid, divided by nominal GDP.

To summarize the nine macroeconomic time series and their connection to theoretical quantities in the model, I write the measurement equation with all variables in terms of deviations from their steady-state:

$$
\text{Meas. Equation}_t = \begin{bmatrix} \text{Real GDP growth}_t \\ \text{Real consumption Growth}_t \\ \text{Real investment Growth}_t \\ \text{Real wage growth}_t \\ \text{Hours worked}_t \\ \text{GDP Deflator Growth}_t \\ \text{FFR}_t \\ \text{Unemployment Rate}_t \\ \text{Transfers to GDP}_t \end{bmatrix} = \begin{bmatrix} \log(Y_t) - \log(Y_{t-1}) \\ \log(C_t) - \log(C_{t-1}) \\ \log(I_t) - \log(I_{t-1}) \\ \log(w_t) - \log(w_{t-1}) \\ \log(h_t) - \log(h_{t-1}) \\ \pi_t \\ i_t \\ u_t \\ (T_t/Y_t) \end{bmatrix} \tag{20}
$$

where here I denote T_t (without arguments) as $T_t = \sum_\ell \int\int T_t(a,z)\mu_t(a,z,\ell)da\ dz$. Note that for flow variables, I integrate the quantities in the continuous time model again to match the data, in the style of Christensen, Neri, and Parra-Alvarez (2024), and then conduct the log differencing.

The result is a system with nine shocks (to monetary policy, TFP, the household bond premium, the productivity of investment, the price markup, the wage rate, the productivity of investment, the labor market matching function efficiency, and fiscal transfers) and nine observables. I assume most of the shocks follow an Ornstein-Uhlenbeck (AR(1)) process, except for the price and wage shocks. For these, I assume that they follow a continuous time autoregressive moving average process (CARIMA) with a (2,1) lag order that sets the second autoregressive component to be negligible, in order to be roughly consistent with the discrete time structure of Smets and Wouters (2007). I map the discretely sampled CARIMA(2,1) processes to their discrete time analogues using the methodology of Thornton and Chambers (2011).

Table 4 summarizes the priors for the shock and variance parameters in the leftmost columns. The priors themselves are as in Smets and Wouters (2007), to be consistent with the rest of the literature.

4 Results

4.1 The Power of Persistence

To understand the implications of the timing of output gaps for unemployment and inflation, I linearize equations (4) and (6), which describe the New Keynesian Phillips curve and the evolution of labor market tightness. Both equations are largely standard elements in DSGE models. I then use them to solve for the dynamics of inflation and unemployment in response to changes in intermediate firms' marginal costs and labor agencies' profits per worker. Appendixes B.2 and B.1 contain the details the analytical derivations. In a partial equilibrium exercise, I then feed exogenous jumps in marginal costs and profits into these solutions, where the driving impulses are composed of simple exponentially decaying functions of varying persistences.

In the first plot of Figure 4, I display the exogenous jumps in per worker profitability that I feed

Table 4: Priors and Posterior estimates from SMC

Parameter	Symbol	Prior			Posterior	
		Dist.	Mean	S. Dev	Mean	S. Dev
AR of ζ_{TFP}	ρ_{TFP}	Beta	0.5	0.11	0.40	0.033
AR of ζ_r	ρ_r	Beta	0.5	0.11	0.91	0.010
AR of ζ_I	ρ_I	Beta	0.5	0.11	0.31	0.0013
AR of ζ_π	ρ_π	Beta	0.5	0.11	0.61	0.089
AR of ζ_w	ρ_w	Beta	0.5	0.11	0.92	0.024
AR of ζ_G	ρ_G	Beta	0.5	0.11	0.26	0.038
AR of ζ_u	ρ_u	Beta	0.5	0.11	0.81	0.028
AR of ζ_T	ρ_T	Beta	0.5	0.11	0.85	0.022
MA of ζ_p	μ_p	Beta	0.5	0.21	0.61	0.093
MA of ζ_w	μ_w	Beta	0.5	0.21	0.45	0.063
S. Dev. of ζ_{MP}	$100\sigma_{MP}$	Inv. Gamma	0.033	2.0	3.8	0.28
S. Dev. of ζ_{TFP}	$100\sigma_{TFP}$	Inv. Gamma	0.033	2.0	0.45	0.033
S. Dev. of ζ_r	$100\sigma_r$	Inv. Gamma	0.033	2.0	4.8	0.49
S. Dev. of ζ_I	$100\sigma_I$	Inv. Gamma	0.033	2.0	4.2	0.37
S. Dev. of ζ_π	$100\sigma_\pi$	Inv. Gamma	0.033	2.0	3.0	0.31
S. Dev. of ζ_w	$100\sigma_w$	Inv. Gamma	0.033	2.0	3.5	0.36
S. Dev. of ζ_g	$100\sigma_g$	Inv. Gamma	0.033	2.0	5.1	1.22
S. Dev. of ζ_u	$100\sigma_u$	Inv. Gamma	0.033	2.0	0.10	0.0091
S. Dev. of ζ_T	$100\sigma_T$	Inv. Gamma	0.033	2.0	5.1	0.25

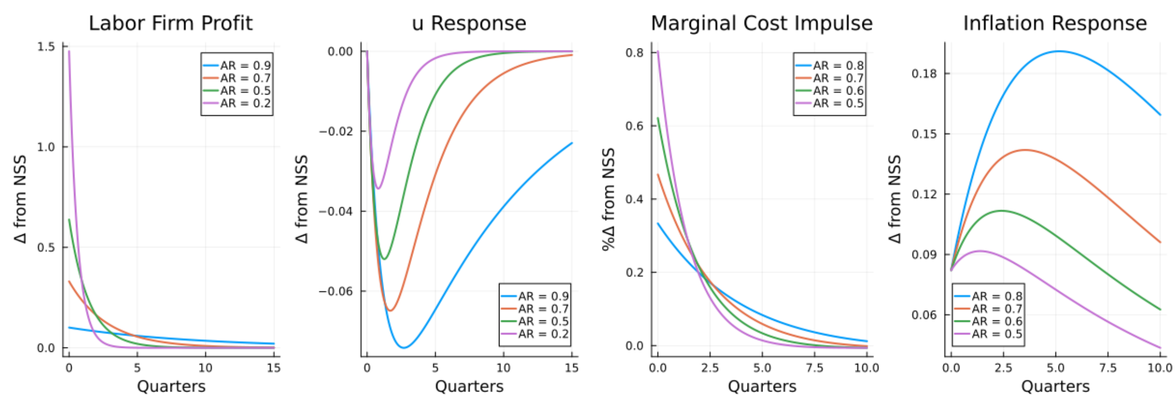

Figure 4: Impulse responses of unemployment and inflation in response to exogenous changes in worker profitability and total marginal cost.

into the labor agency's problem. The net present value of the increases are adjusted to be equal across the experiments, but the auto regressive (AR) persistences of the impulses are varied and range from AR(1) coefficients of 0.9 (highly persistent) to 0.2 (highly transitory). In the second plot, I display the resulting changes in the unemployment rate. Smaller but more persistent increases in per worker profit drive substantially more hiring and larger reductions in unemployment compared to large-but-fleeting profitability jumps.

I conduct a similar analysis of the equations that describe inflation. I generate jumps in marginal costs of varying persistence and normalize their net present value; I display these impulses in the third panel of Figure (4). The final plot in the panel depicts the effect of these marginal cost spikes on inflation.[4] Just as in the unemployment case, the response of inflation to increases in marginal cost is many times stronger when the persistence of the cost increase is larger, all else equal.

Both vacancy creation and inflation change according to strong forward-looking components in the model, even if the levels of unemployment and inflation have backward-looking components that induce a hump shape. Hiring and sticky price changes are an investment for firms that both take time and have persistent consequences. As such, even if stimulus checks drive large positive output gaps in the short term, forward-looking firms largely treat past output gaps as sunk. If output gaps fall rapidly or even become negative in the future after an initial peak, then this strongly attenuates the ability of stimulus checks to spur large movements in inflation and output compared to other macroeconomic events (like changes to monetary policy) that take years to fully play out.

4.2 Fiscal transfers with inattention

As documented in Auclert, Rognlie, and Straub (2020), greater degrees of inattention dampen the forward-looking general equilibrium effects that drive monetary policy, making it less stimulative. I show that the exact *opposite* is the case for deficit-financed fiscal transfers to households under an active monetary, passive fiscal policy mix. In Figure 5, I show the impulse responses in the model economy to a one-time increase in transfers equal to 1% of annual GDP on impact, which mean reverts with an AR coefficient of 0.16 to roughly simulate the speed at which transfers were disbursed during the COVID-19 pandemic. I plot a range of λ values; the extreme case of $\lambda = 10$ essentially corresponds to the full information model, while $\lambda = 0.01$ gives a half-life for the non-updating share of the population of over 17 years. The estimates calculated using the learning parameter in 3 is plotted in a solid line.

When stimulus checks go out, hand-to-mouth households spend them immediately. Nominal rigidities are high in the estimated version of the model, so inflation is muted and the Federal Reserve makes only miniscule changes to the nominal interest rate. However, for forward-looking households on the margin of spending or saving, the departure from rational expectations dampens two important channels. First, households could internalize how future aggregate demand may be high from the surge in spending today, which further increases current aggregate demand as households attempt to dis-save or borrow against their future high income. Conversely, households may anticipate that future tax revenue may have to be raised in order to pay off the future debt incurred from the government deficit, the force behind Ricardian equivalence. As illustrated using a tractable RANK model in Gabaix (2020), stronger information frictions should therefore further break Ricardian equivalence, amplifying fiscal policy. As such, dampening the first channel weakens the stimulus effect of stimulus checks, while dampening the second strengthens it.

As shown in Figure 5, dampening the second effect is clearly far more important quantitatively – even despite the fact that debt stabilization taxes in my model are carried out in a progressive

[4] For this exercise, I increase the slope of the Phillips Curve to a standard 0.10, although the IRF matching exercise demands a baseline calibration with a far flatter curve. In the next sections, I explore robustness of the results to various NKPC slopes.

manner. After stimulus checks are sent out, aggregate spending and income rise at virtually every horizon. Households see the income in their bank accounts, but when they are less attentive to the implications of fiscal stabilization, the more powerful fiscal transfer policy becomes. Quantitatively, moving from the full information to the estimated sticky expectation model, the peak of the response of real GDP rises by a factor of one third. After the first year, the total cumulative sum of output gaps is three times higher with the estimated information frictions than without, for a fiscal transfer multiplier of 0.35 compared to 0.11.

Despite the stimulative potential of fiscal policy, however, it tends to have a highly muted effect on unemployment in the model (although the effect is still greater as information frictions increase). Hiring in the search-and-matching block of the model involves a costly investment that takes time, and the average employer-employee relationship lasts several years. The stimulative effect of stimulus checks, in contrast, largely play out over just a few quarters. As a result, while stimulus checks are effective at increasing consumption and output (and therefore welfare) in the model economy, most of the increase in labor demand under the estimated parameters is on the intensive margin, with firms increasing hours per worker (or alternatively, worker effort intensity). Compared to monetary policy, transitory fiscal transfers boost the extensive margin of the labor market, employment, only relative modestly. Indeed, because the government's debt is repaid gradually over a long period of time, the persistent fall in aggregate demand after the initial boost in the stimulus checks actually drives unemployment slightly higher – although the effects are quantitatively very small.

Similarly, fiscal transfers also contribute only slightly to inflation in the first few quarters, and actually become very weakly deflationary as time goes on and debt repayment begins. Part of the reason the magnitude is small is due to the very low Phillips Curve slope implied by the model estimates. However, even reducing the nominal rigidities 50-fold leads to only a transitory inflationary peak of roughly 0.05% when all of the other parameters are kept at their estimated values. Rather, much like the unemployment response, the lack of inflationary pressure from the transfers is mostly due to the transitory nature of their timing – a theme explored in Kwicklis (2025a). Even with a backward-looking term, the Phillips Curve is forward-looking, and by the time firms are able to significantly adjust their prices, the output gaps of the fiscal stimulus are past and sunk, while the gradual (but prolonged) debt repayment lies in the future. Firms with nominal rigidities thus react less to the sharp stimulus than they do to the austerity period that follows.

4.3 Fiscal transfers and the time series

In Figure 6, I detail historical decompositions of the federal funds rate, real GDP growth, and inflation inflation for the United States using the nine exogenous shock processes detailed in the model. The HANK model is evaluated at the minimum-distance coefficients obtained from the monetary policy IRF matching and, for the parameters related to shock persistences and variances, at the posterior modes obtained from the SMC estimation. As in Smets and Wouters (2007), and to emphasize the policy-related shocks, I bin the bond premium, investment, and government expenditure shocks as "demand" shocks, and similarly bin the wage and price markup shocks together. For comparison, I also solve and integrate a continuous time RANK model similar to the model solved in Smets and

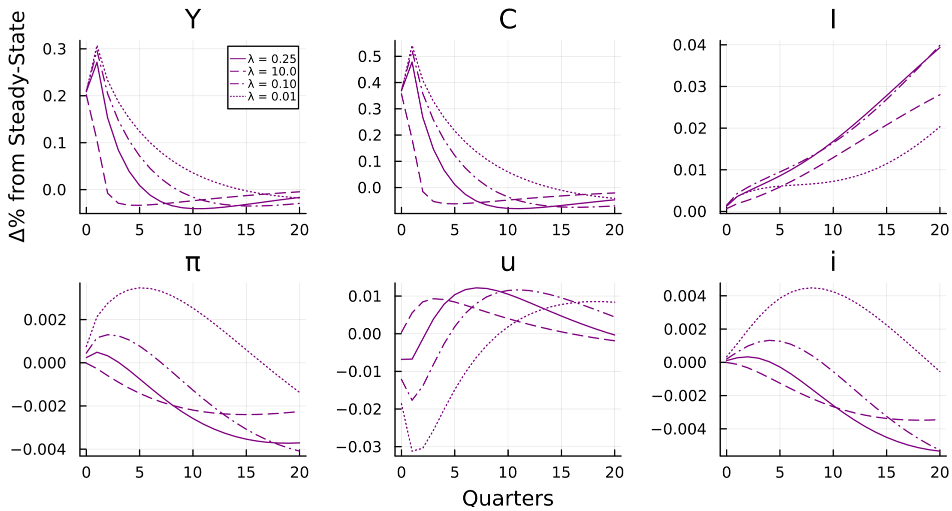

Figure 5: Impulse response functions to fiscal transfers to households with varying degrees of inattention. The solid line depicts the impulse response of the estimated model, while dashed and dotted lines represent alternative calibrations. Lower values of λ correspond to slower rates of household updating.

Wouters (2007). The decompositions through the lens of the RANK model are also displayed in Figure 6, on the right.

Despite the differences in the structure of the models, both the HANK and RANK frameworks tell a similar qualitative story regarding the U.S. macroeconomic policy following 1965. Both diagnose monetary policy as too "loose" in the 1970s and contributing to elevated levels of inflation, both identify contractionary interest rate hikes and subsequent disinflation following Paul Volker's appointment as Federal Reserve Chair in 1980, and both describe the 2010s as experiencing a large demand-side recession. For reasons previously discussed, the HANK model loads very little inflation and unemployment variation (the latter is not shown) onto stimulus transfers. Indeed, the HANK model does not attribute much GDP growth to fiscal transfers until the massive pandemic-era interventions.

Focusing on the period after 2020, the model suggests that stimulus checks significantly boosted the recovery. Figure 7 depicts the trajectory of real GDP, inflation, and unemployment following the 2020 recession, along with the relative contributions of fiscal and monetary policy to each line, according to the shocks after 2019 estimated via a Durbin and Koopman (2002) smoother. In keeping with the previous impulse response functions, the model suggests that the stimulus checks significantly boosted the recovery from the recession. In particular, the model predicts that by 2020q2 quarterly real GDP would have declined by 0.84 percentage points relative to 2019 had fiscal transfers checks not increased following the CARES act, roughly 9% of the realized decline. After the passage of the American Rescue Plan in the first quarter of 2021, real GDP per capita essentially returned to 2019 levels – but would have been roughly 2.5 percentage points lower in the model, had fiscal stimulus not taken place. Had this been this case, the economy would have taken an additional three months to return to its pre-2020 real GDP per capita levels.

To broadly assess how much worse the recession would have been without the stimulus checks, I integrate the difference between the realized GDP per capita line and then trend and compare that value to the blue area, the accumulated additional losses in the absence of the stimulus checks. Between

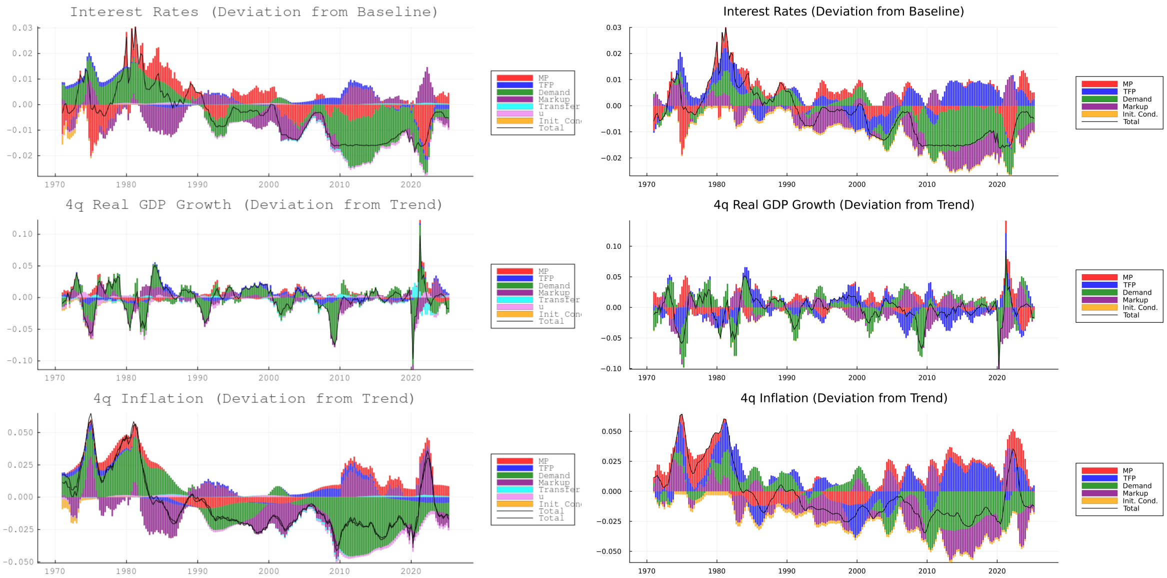

Figure 6: HANK (left) and RANK (right) historical shock decompositions of interest rates, output, and inflation.

the start of the recession and the start of 2022, the total sum of output losses is 25% smaller in the actual time series than in the model's simulation of a no-transfer scenario.

The second and third panels of Figure 7 indicate that the model ascribes very little movement in the 4-quarter inflation rate and the unemployment rate to fiscal transfers. Instead, the first panel of Figure 6 reveals that the model prefers to load much of the variation during that period in prices onto markup shocks – i.e., shocks that are largely unrelated to aggregate demand and output gaps. This is unsurprising, as the local projections estimates force the model to adopt a very flat Phillips Curve. Even so, if one dramatically increases the slope of the Phillips Curve to 0.10, then fiscal transfers still only account for less than 15% of inflation's 6% above-trend peak in the second quarter of 2022, as shown in the Appendix.

In contrast to the blue shaded area of the effect of fiscal transfers, the red shaded areas in Figure 7 denote the contribution of monetary policy to the macroeconomic time series. The monetary policy shocks are identified as deviations from the active slow-moving Taylor rule of equation (12); according to the estimated model, the Federal Reserve kept interest rates below the Taylor rule's proscription in the face of rising inflation and only reversed course in 2022, as shown in the graph on the left in Figure 6. During this time, and taking it as given that the agents in the model reacted to each instance that rates were below the Taylor target with surprise, the monetary policy innovations drove an increasing share of the post-COVID expansion. The effects were too lagged to significantly change the dynamics of the macroeconomy until the United States was well into its recovery, however, and only contribute to a modest increase in inflation and decrease in the unemployment rate long after both had respectively peaked. Following the decline in inflation after 2024, however, the after-effects of loose monetary policy explains most of the remaining persistence of inflation above trend.

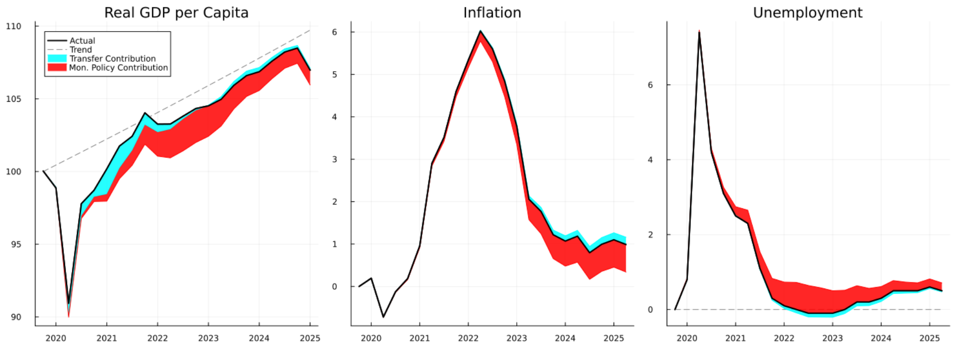

Figure 7: Decomposition of real GDP per capital, 4-quarter inflation, and the unemployment rate following 2020.

5 Conclusion

Using the methodology of Kwicklis (2025b), I build a medium-scale HANK model with search-and-matching frictions and sticky expectations that matches household MPC evidence and the slow-moving responses to identified monetary shocks of output, consumption, wages, hours, unemployment, and inflation. The model implies that sharply transitory transfers such as stimulus checks are noticeably more powerful under sticky expectations than under full information because households internalize the debt-stabilization implications only gradually. However, their inflationary and unemployment effects remain limited: the consumption-driven output gap is short-lived, while price adjustment and vacancy creation are slow, so firms do not strongly reprice and primarily adjust on intensive margins in the factor market (capital utilization and hours) rather than through new hiring.

Filtering postwar time series through the model, I find that routine movements in transfers account for little of the variance in real aggregates. In contrast, the pandemic-era transfers likely mitigated the depth of the recession but contributed only modestly to the subsequent inflation surge and the decline in unemployment.

How should the result that stimulus checks do little to drive inflation or reduce unemployment be interpreted? The slowness with which household expectations change following macroeconomic news and the heterogeneity of marginal propensities to consume are both features of the sticky expectations HANK model that appear to be well-grounded by the data. They are also features that increase the relative strength of stimulus checks in the model, bringing its predictions closer to conventional wisdom. Rather, other components of the model that are more standard in the DSGE literature – namely, the search-and-matching block and the partially indexed New Keynesian Phillips Curve – have dynamics that limit the ability of short-term jumps in aggregate demand to propagate into other parts of the model. This warrants further investigation into the frictions in the firms' problems, and may call for new modeling assumptions on the firms' side as well. Namely, the process by which hiring managers and price setters form their own expectations may also warrant information frictions – perhaps even instead of adjustment costs. Data from the Survey of Business Uncertainty (SBU) and other sources could potentially better calibrate such frameworks.

There are many more technical directions for further investigation as well. My model is completely linear with respect to macroeconomic aggregates, and while it matches interest rates in the zero lower bound (ZLB) eras of the 21st century, it does so only by engineering unexpected interest rate movements as fresh shocks. It may be possible, however, to adapt my state space framework to study the ZLB in otherwise active monetary/passive fiscal policy regimes explicitly, as in Alves and Violante (2025), or to study policy regime switching models like Bianchi and Ilut (2017). These questions of state dependence and nonlinearity are likely particularly important for examining major macroeconomic events.

Overall, this paper clarifies a quantitative view: an estimated HANK model with large MPCs but otherwise standard elements fitted to the data would *not* have predicted substantial inflation in response to the disbursal of stimulus in the COVID era, even with sticky expectations amplifying the effect. Whether this reflects a model misspecification or an accurate implication depends on how standard firm-side elements can be verified empirically.

References

Acharya, Sushant, William Chen, Marco Del Negro, Keshav Dogra, Aidan Gleich, Shlok Goyal, Ethan Matlin, Donggyu Lee, Reca Sarfati, and Sikata Sengupta (Aug. 2023). *Estimating HANK for Central Banks*. Tech. rep. 1071. Federal Reserve Bank of New York.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2021). "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach". In: *The Review of Economic Studies* 89.1, pp. 45–86. ISSN: 0034-6527.

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf (2018). "When Inequality Matters for Macro and Macro Matters for Inequality". In: *NBER Macroeconomics Annual* 32, pp. 1–75.

Alves, Felipe and Giovanni L. Violante (Feb. 2025). *Monetary Policy Under Okun's Hypothesis*. Working Paper 33488. National Bureau of Economic Research.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub (2020). *Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model*. Working Paper 26647. National Bureau of Economic Research.

— (2024). "The Intertemporal Keynesian Cross". In: *Journal of Political Economy* 132.12, pp. 4068–4121.

Baker, Scott R., Robert A. Farrokhnia, Steffen Meyer, Michaela Pagel, and Constantine Yannelis (Nov. 2023). "Income, Liquidity, and the Consumption Response to the 2020 Economic Stimulus Payments". In: *Review of Finance* 27.6, pp. 2271–2304.

Bardóczy, Bence and João Guerreiro (2024). *Unemployment Insurance in Macroeconomic Stabilization with Imperfect Expectations*. Working Paper. Federal Reserve Board and University of California, Los Angeles.

Bayer, Christian, Benjamin Born, and Ralph Luetticke (May 2024). "Shocks, Frictions, and Inequality in US Business Cycles". In: *American Economic Review* 114.5, pp. 1211–47.

Bayer, Christian and Ralph Luetticke (2020). "Solving discrete time heterogeneous agent models with aggregate risk and many idiosyncratic states by perturbation". In: *Quantitative Economics* 11.4, pp. 1253–1288.

Bianchi, Francesco, Renato Faccini, and Leonardo Melosi (May 2023). "A Fiscal Theory of Persistent Inflation*". In: *The Quarterly Journal of Economics* 138.4, pp. 2127–2179. ISSN: 0033-5533.

Bianchi, Francesco and Cosmin L Ilut (2017). "Monetary/Fiscal policy mix and agents' beliefs". In: *Review of Economic Dynamics* 26, pp. 113–139. ISSN: 1094-2025.

Cai, Michael, Marco Del Negro, Edward Herbst, Ethan Matlin, Reca Sarfati, and Frank Schorfheide (Sept. 2020). "Online estimation of DSGE models". In: *The Econometrics Journal* 24.1, pp. C33–C58. ISSN: 1368-4221.

Campbell, John and Angus Deaton (July 1989). "Why is Consumption So Smooth?" In: *The Review of Economic Studies* 56.3, pp. 357–373. ISSN: 0034-6527.

Campbell, John and N. Gregory Mankiw (1989). "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence". In: *NBER Macroeconomics Annual* 4, pp. 185–216.

Carroll, Christopher D. (Feb. 2003). “Macroeconomic Expectations of Households and Professional Forecasters*”. In: *The Quarterly Journal of Economics* 118.1, pp. 269–298. ISSN: 0033-5533.

Carroll, Christopher D., Edmund Crawley, Jiri Slacalek, Kiichi Tokuoka, and Matthew N. White (2020). “Sticky Expectations and Consumption Dynamics”. In: *American Economic Journal: Macroeconomics* 12.3, pp. 40–76.

Christensen, Bent Jesper, Luca Neri, and Juan Carlos Parra-Alvarez (2024). “Estimation of continuous-time linear DSGE models from discrete-time measurements”. In: *Journal of Econometrics* 244.2, p. 105871. ISSN: 0304-4076.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (2005). “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy”. In: *Journal of Political Economy* 113.1, pp. 1–45. ISSN: 00223808, 1537534X.

Christiano, Lawrence J., Martin Eichenbaum, and Mathias Trabandt (2016). “Unemployment and Business Cycles”. In: *Econometrica* 84.4, pp. 1523–1569.

Diamond, Peter A. (1982). “Aggregate Demand Management in Search Equilibrium”. In: *Journal of Political Economy* 90.5, pp. 881–894. ISSN: 00223808, 1537534X.

Durbin, J. and S. J. Koopman (2002). “A Simple and Efficient Simulation Smoother for State Space Time Series Analysis”. In: *Biometrika* 89.3, pp. 603–615. ISSN: 00063444.

Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik (Oct. 2021). “MPC Heterogeneity and Household Balance Sheets”. In: *American Economic Journal: Macroeconomics* 13.4, pp. 1–54.

Floden, Martin and Jesper Lindé (2001). “Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance?” In: *Review of Economic Dynamics* 4.2, pp. 406–437. ISSN: 1094-2025.

Friedman, Milton (May 1957). *The Permanent Income Hypothesis.*

Gabaix, Xavier (Aug. 2020). “A Behavioral New Keynesian Model”. In: *American Economic Review* 110.8, pp. 2271–2327.

Hagedorn, Marcus, Iourii Manovskii, and Kurt Mitman (Feb. 2019). *The Fiscal Multiplier.* Working Paper 25571. National Bureau of Economic Research.

Herbst, Edward and Frank Schorfheide (2014). “Sequential Monte Carlo Sampling for DSGE Models”. In: *Journal of Applied Econometrics* 29.7, pp. 1073–1098.

Jordá, Óscar (Mar. 2005). “Estimation and Inference of Impulse Responses by Local Projections”. In: *American Economic Review* 95.1, pp. 161–182.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante (2018). “Monetary Policy According to HANK”. In: *American Economic Review* 108.3, pp. 697–743.

Krusell, Per and Anthony A. Smith (1998). “Income and Wealth Heterogeneity in the Macroeconomy”. In: *Journal of Political Economy* 106.5, pp. 867–896.

Kwicklis, Noah (2025a). “Active vs. passive policy and the trade-off between output and inflation in HANK”. In: *Journal of Monetary Economics* 151, p. 103732. ISSN: 0304-3932.

— (2025b). *State Space HANK and Sticky Expectations.* Working Paper. University of California, Los Angeles.

Leeper, Eric M. (1991). "Equilibria under 'active' and 'passive' monetary and fiscal policies". In: *Journal of Monetary Economics* 27.1, pp. 129–147. ISSN: 0304-3932.

Mankiw, N. Gregory and Ricardo Reis (2002). "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve*". In: *The Quarterly Journal of Economics* 117.4, pp. 1295–1328. ISSN: 0033-5533.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson (2016). "The Power of Forward Guidance Revisited". In: *American Economic Review* 106.10, pp. 3133–58.

Mortensen, Dale T. and Christopher A. Pissarides (1994). "Job Creation and Job Destruction in the Theory of Unemployment". In: *The Review of Economic Studies* 61.3, pp. 397–415. ISSN: 00346527, 1467937X.

Newey, Whitney K. and Kenneth D. West (1987). "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". In: *Econometrica* 55.3, pp. 703–708. ISSN: 00129682, 14680262.

Orchard, Jacob D, Valerie Ramey, and Johannes F Wieland (Feb. 2025). "Micro MPCs and Macro Counterfactuals: The Case of the 2008 Rebates*". In: *The Quarterly Journal of Economics* 140.3, pp. 2001–2052. ISSN: 0033-5533.

Pissarides, Christopher A. (1986). "Unemployment and Vacancies in Britain". In: *Economic Policy* 1.3, pp. 499–541. ISSN: 0266-4658.

Ramey, Valerie (2016). "Macroeconomic Shocks and Their Propagation". In: *Handbook of Macroeconomics*. Vol. 2. Elsevier. Chap. Chapter 2, pp. 71–162.

Reis, Ricardo (July 2006). "Inattentive Producers". In: *The Review of Economic Studies* 73.3, pp. 793–821. ISSN: 0034-6527.

Reiter, Michael (2009). "Solving heterogeneous-agent models by projection and perturbation". In: *Journal of Economic Dynamics and Control* 33.3, pp. 649–665. ISSN: 0165-1889.

Romer, Christina D. and David H. Romer (Sept. 2004). "A New Measure of Monetary Shocks: Derivation and Implications". In: *American Economic Review* 94.4, pp. 1055–1084.

Rotemberg, Julio J. (1982). "Sticky Prices in the United States". In: *Journal of Political Economy* 90.6, pp. 1187–1211. ISSN: 00223808, 1537534X.

Smets, Frank and Raf Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach". In: *American Economic Review* 97.3, pp. 586–606.

— (2024). *Fiscal Backing, Inflation and US Business Cycles*. Discussion Paper 19791. Paris & London: CEPR.

Thornton, Michael A. and Marcus J. Chambers (2011). *On the Equivalence of Discrete Time Representations of Continuous Time ARMA Processes*. Working Paper. University of Reading and University of Essex.

A Appendix to Section 2: Medium-Scale HANK Model Derivation

A.1 Labor Agency Sector

Hiring in the labor market is managed identical labor agencies indexed by j. Each one hires such that their number of employees $N_t(j)$ evolves according to

$$\dot{N}_t(j) = \lambda_{v,t} v_t(j) - \lambda_{EU} N_t(j)$$

If hiring firm j internalizes a downward-sloping demand curve for its laborers' hours $h_t(j)$, then a CES demand system implies

$$h_t(j) = \left(\frac{r_t^l(j)}{r_t^l}\right)^{-\varepsilon_\ell} h_t$$

such that

$$\begin{aligned} r_{a,t} J_t(N_t(j)) = \max_{v_t(j), r_t^l(j)} &\left\{ (r_t^l(j) - w_t) \left(\frac{r_t^l(j)}{r_t^l}\right)^{-\varepsilon_l} N_t(j) Z_t h_t - \frac{\phi_h}{2} (h_t(j) - h^*)^2 N_t(j) Z_t \right. \\ &\left. - \eta_v v_t(j) - \eta_l \lambda_{v,t} v_t(j) + J_t'(N_t(j))[\lambda_{v,t} v_t(j) - \lambda_{EU} N_t(j)] + \partial_t J_t(N_t(j)) \right\} \end{aligned}$$

$$\begin{aligned} r_{a,t} J_t(N_t(j)) = \max_{v_t(j), r_t^l(j)} &\left\{ (r_t^l(j) - w_t) \left(\frac{r_t^l(j)}{r_t^l}\right)^{-\varepsilon_l} N_t(j) Z_t h_t - \frac{\phi_h}{2} \left(\left(\frac{r_t^l(j)}{r_t^l}\right)^{-\varepsilon_\ell} h_t - h^* \right)^2 N_t(j) Z_t h_t \right. \\ &\left. - \eta_v v_t(j) - \eta_l \lambda_{v,t} v_t(j) + J_t'(N_t(j))[\lambda_{v,t} v_t(j) - \lambda_{EU} N_t(j)] + \partial_t J_t(N_t(j)) \right\} \end{aligned}$$

FOC:

$$\left(\frac{r_t^l(j)}{r_t^l}\right)^{-\varepsilon_l} + (r_t^l(j) - w_t) \left(\frac{r_t^l(j)}{r_t^l}\right)^{-\varepsilon_l} (-\varepsilon_l) \frac{1}{r_t^l} - \phi_h \left(\left(\frac{r_t^l(j)}{r_t^l}\right)^{-\varepsilon_\ell} h_t - h^* \right) \left(\frac{r_t^l(j)}{r_t^l}\right)^{-\varepsilon_\ell} (-\varepsilon_l) \frac{1}{r_t^l(j)} = 0$$

In a symmetric equilibrium,

$$\begin{aligned} 1 &= \varepsilon_l (r_t^l - w_t) \frac{1}{r_t^l} - \varepsilon_l \phi_h (h_t - h^*) \frac{1}{r_t^l} \\ r_t^l (\varepsilon_l - 1) &= \varepsilon_l w_t + \varepsilon_l \phi_h (h_t - h^*) \\ r_t^l &= \frac{\varepsilon_l}{\varepsilon_l - 1} [w_t + \phi_h (h_t - h^*)] \end{aligned}$$

The FOC with respect to $v_t(j)$ is (in a symmetric equilibrium, dropping the j subscripts)

$$-(\eta_v + \eta_l \lambda_{v,t}) + J_t'(N_t) \lambda_{v,t} = 0$$

such that the marginal benefit of posting a vacancy is equated to its marginal cost:

$$J_t'(N_t) \lambda_{v,t} = (\eta_v + \eta_l \lambda_{v,t})$$

$$\Rightarrow \quad J_t'(N_t) = \frac{\eta_v}{\lambda_{v,t}} + \eta_l$$

Totally differentiating:

$$d(J_t'(N_t)) = d\left(\frac{\eta_v}{\lambda_{v,t}}\right) = -\left(\frac{\eta_v}{\lambda_{v,t}^2}\right) d\lambda_{v,t}$$

The envelope condition indicates

$$r_{a,t}J_t'(N_t) = (r_t^l - w_t)Z_t h_t - \lambda_{EU}J_t'(N_t) + \underbrace{J_t''(N_t)[\lambda_{v,t}v_t - \lambda_{EU}N_t] + \partial_t J_t'(N_t)}_{\mathbb{E}[dJ_t'(N_t)]/dt}$$

Thus

$$(r_{a,t} + \lambda_{EU})\left[\frac{\eta_v}{\lambda_{v,t}} + \eta_l\right] = (r_t^l - w_t)Z_t h_t - \left(\frac{\eta_v}{\lambda_{v,t}^2}\right)\frac{\mathbb{E}_t[d\lambda_{v,t}]}{dt}$$

Rearranging:

$$(r_{a,t} + \lambda_{EU})\left[1 + \frac{\eta_l}{\eta_v}\lambda_{v,t}\right]\lambda_{v,t} = (r_t^l - w_t)Z_t h_t\left(\frac{\lambda_{v,t}}{\eta_v}\right)\lambda_{v,t} - \frac{\mathbb{E}_t[d\lambda_{v,t}]}{dt}$$

such that

$$\frac{\mathbb{E}_t[d\lambda_{v,t}]}{dt}\frac{1}{\lambda_{v,t}} = (r_t^l - w_t)Z_t h_t\left(\frac{\lambda_{v,t}}{\eta_v}\right) - (r_{a,t} + \lambda_{EU})\left[1 + \frac{\eta_l}{\eta_v}\lambda_{v,t}\right]$$

Note that this means that in the NSS,

$$\frac{1}{\varepsilon_l}wZh\left(\frac{\lambda_v}{\eta_v}\right) = (r_a + \lambda_{EU})\left[1 + \frac{\eta_l}{\eta_v}\lambda_v\right]$$

To simplify the calibration, I set ε_l such that steady-state hours worked are equal to 1.0:

$$\varepsilon_l = \frac{wZh\left(\frac{\lambda_v}{\eta_v}\right)}{(r_a + \lambda_{EU})\left[1 + \frac{\eta_l}{\eta_v}\lambda_v\right]}$$

Note that because r_a and λ_{EU} are relatively small, this implies ε_l is large – such that labor hours are very substitutable across firms.

A.2 Mutual Fund Problem

If there are adjustment costs to altering capital, each mutual fund solves

$$\begin{aligned}
\Pi_0(I_0, K_0) = \max_{(I_t,u_t)_{t\geq 0}} \ & \mathbb{E}_0 \int_0^\infty e^{-\int_0^t r_\tau d\tau}\Big[(r_t^k u_t K_t - I_t - \psi(u_t)K_t\Big]dt \\
\text{s.t. } & \frac{dK_t}{dt} = e^{\zeta_t^I}[1-\Phi(I_t, \bar{I}_t)]I_t - \delta K_t \\
& \frac{d\bar{I}_t}{dt} = \kappa_I(I_t - \overline{I}_t)
\end{aligned}$$

Let $k_t \equiv \frac{K_t}{A_t}$ and $x_t \equiv \frac{I_t}{A_t}$, where $A_t = A_0 e^{g_Y t}$. Note that this means $\dot{x}_t = \dot{I}_t/A_t - x_t g_Y$, such that $\dot{I}_t/I_t = \dot{x}_t/x_t + g_Y$. Similarly, $\dot{k}_t = x_t - (\delta + g_Y)k_t$. The stationary problem is

$$\begin{aligned}
\Pi_0(x_0, k_0) = \max_{(x_t,u_t)_{t\geq 0}} \ & \mathbb{E}_0 \int_0^\infty e^{-\int_0^t (r_\tau - g_Y) d\tau}\Big[(r_t^k u_t k_t - x_t - \psi(u_t)k_t\Big]dt \\
\text{s.t. } & \frac{dk_t}{dt} = e^{\zeta_t^I}[1-\Phi(x_t, \overline{x}_t)]x_t - (\delta + g_Y)k_t \\
& \frac{d\overline{x}_t}{dt} = \kappa_I(x_t - \overline{x}_t)
\end{aligned}$$

Writing the problem in recursive form,

$$(r_t - g_Y)\Pi_t = \max_{x_t,u_t} r_t^k u_t k_t - x_t - \psi(u_t)k_t + \overbrace{\frac{\partial \Pi_t}{\partial \overline{x}_t}\kappa_I(\tilde{x}_t - \overline{x}_t) + \frac{\partial \Pi_t}{\partial k_t}(e^{\zeta_t^I}[1-\Phi(x_t,\overline{x}_t)]x_t - (\delta+g_Y)k_t) + \frac{\partial \Pi_t}{\partial t}}^{\mathcal{D}\Pi_t}$$

Taking FOCs,

$$-1 + \frac{\partial \Pi_t}{\partial k_t} e^{\zeta_t^I}[1 - \Phi(x_t, \overline{x}_t) - \partial_x \Phi(x_t, \overline{x}_t)x_t] = 0$$

$$r_t^k k_t - \psi'(u_t)k_t = 0$$

The last FOC equates the marginal value of renting more capital on the intensive margin with the cost of renting more intensive capital:

$$\psi'(u_t) = r_t^k$$

Rearranging the first FOC, and writing $\frac{\partial \Pi}{\partial k} \equiv q_t^K$ and $\frac{\partial \Pi}{\partial \overline{x}} \equiv \overline{q}_t^I$:

$$q_t^k e^{\zeta_t^I}[1 - \Phi(x_t, \overline{x}_t) - \partial_x \Phi(x_t, \overline{x}_t)x_t] = 1$$

Using the envelope theorem,

$$(r_t - g_Y)\partial_k \Pi_t = r_t^k u_t - \psi(u_t) + \partial_k\left[\frac{\partial \Pi_t}{\partial \overline{x}_t}\kappa_I(\tilde{x}_t - \overline{x}_t) + \frac{\partial \Pi_t}{\partial k_t}\left(e^{\zeta_t^I}[1-\Phi(x_t,\overline{x}_t)]x_t - (\delta+g_Y)k_t\right) + \frac{\partial \Pi_t}{\partial t}\right]$$

$$= r_t^k u_t - \psi(u_t) + \underbrace{\left[\frac{\partial \partial_k \Pi_t}{\partial \overline{x}_t}\kappa_I(\tilde{x}_t - \overline{x}_t) + \frac{\partial \partial_k \Pi_t}{\partial k_t}(e^{\zeta_t^I}[1-\Phi(x_t,\dot{x}_t)]x_t - (\delta+g_Y)k_t) + \frac{\partial \partial_k \Pi_t}{\partial t}\right]}_{\frac{\mathbb{E}[d\partial_k \Pi_t]}{dt}} - (\delta+g_Y)\frac{\partial \Pi_t}{\partial k_t}$$

such that

$$(r_t - g_Y)q_t^k = r_t^k u_t - (\delta + g_Y)q_t^k - \psi(u_t) + \frac{\mathbb{E}[dq_t^k]}{dt}.$$

$$\Rightarrow \frac{\mathbb{E}[dq_t^k]}{dt} = (r_t + \delta)q_t^k - r_t^k u_t + \psi(u_t).$$

A.2.1 Quadratic Costs

Suppose the adjustment costs take the form

$$\Phi(I_t, \overline{I}_t) = \frac{\phi_I}{2}\left(\frac{I_t}{\overline{I}_t} - 1\right)^2 = \frac{\phi_I}{2}\left(\frac{x_t}{\overline{x}_t} - 1\right)^2$$

such that

$$\partial_x \Phi(x_t, \overline{x}_t) = \phi_I \left(\frac{x_t}{\overline{x}_t} - 1\right)\frac{1}{\overline{x}_t}$$

Plugging these functional forms in, the investment choice is

$$q_t^k e^{\zeta_t^I}\left[1 - \frac{\phi_I}{2}\left(\frac{x_t}{\overline{x}_t} - 1\right)^2 - \phi_I\left(\frac{x_t}{\overline{x}_t} - 1\right)\frac{x_t}{\overline{x}_t}\right] = 1$$

Solving this quadratic expression,

$$x_t = \left\{\frac{2}{3} + \frac{1}{2}\left(\frac{16}{9} + \frac{8}{3\phi_I}\left[1 - \frac{\phi_I}{2} - \frac{1}{q_t^k e^{\zeta_t^I}}\right]\right)^{1/2}\right\}\overline{x}_t$$

A.3 Financier Households (external habit formation)

"Financier" households recieve dividend and interest income and maximize utility:

$$\max_{(c_t)_{t\geq 0}} \mathbb{E}_0 \left[\int_0^\infty e^{-\rho^* t} \left[e^{\zeta_t^c} \frac{(c_t^{FIN} - \beta_c \bar{c}_t)^{1-\gamma}}{1-\gamma} \right] \right]$$

$$\text{s.t. } \frac{da_t}{dt} = (i_t - \pi_t - g_Y) a_t + s_t - c_t^{FIN}$$

$$\frac{d\bar{c}_t}{dt} = \theta_c C_t^{FIN} - \theta_c \bar{c}_t$$

Where i is the nominal interest rate, w is the real wage, p is the price of the consumption good, and s is the dividend paid out by firms to shareholders. In equilibrium, $c_t^{FIN} = C_t^{FIN}$ – but this is not internalized by the individual household.

The household HJB is then

$$\begin{aligned} \rho^* V(a, \bar{c}^{FIN}; \zeta) = \max_{c,h} \Bigg\{ & \left[e^{\zeta_t^c} \frac{(c_t^{FIN} - \beta_c \bar{c}_t^{FIN})^{1-\gamma}}{1-\gamma} \right] \\ & + \partial_a V(a, \bar{c}^{FIN}; \zeta)[(i - \pi - g_Y)a + s - c^{FIN}] \\ & + \partial_{\bar{c}^{FIN}} V(a, \bar{c}^{FIN}; \zeta)[\theta_c C^{FIN} - \theta_c \bar{c}^{FIN}] \\ & + \underbrace{\frac{\mathbb{E}_t^\zeta dV(a, \bar{C}; \zeta)}{dt}}_{\mathcal{D}_\zeta V(a, \bar{C}; \zeta)} \Bigg\} \end{aligned}$$

Taking FOCs,

$$e^{\zeta_c}(c^{FIN} - \beta_c \bar{c}^{FIN})^{-\gamma} - \partial_a V(a, \bar{c}^{FIN}; \zeta) = 0$$

Using the Envelope Theorem,

$$\rho^* \partial_a V = \partial_a V(i - \pi) + \partial_a^2 V([i - \pi]a + s - c) + \theta \partial_a \partial_{\bar{c}} V(\theta_c c - \theta_c \bar{c}) + \partial_a \partial_t V$$

Thus, using $\mathcal{D}$ as the macro-variable infinitessimal generator,

$$\partial_a V(\rho^* - [i - \pi]) = \mathcal{D} \partial_a V = \frac{\mathbb{E}_t[d\partial_a V]}{dt}$$

Thus the Euler equation becomes

$$\frac{\mathbb{E}_t[d(c_t^{FIN} - \beta_c \bar{c}_t^{FIN})^{-\gamma}]}{dt} = (\rho^* - [i - \pi])(c_t^{FIN} - \beta_c \bar{c}_t^{FIN})^{-\gamma}$$

With the chain rule,

$$-\gamma(c_t^{FIN} - \beta_c \bar{c}_t^{FIN})^{-\gamma-1} \frac{\mathbb{E}_t[d(c_t^{FIN} - \beta \bar{c}_t^{FIN})]}{dt} = (\rho^* - [i - \pi])(c_t^{FIN} - \beta_c \bar{c}_t)^{-\gamma}$$

$$\Rightarrow \frac{\mathbb{E}_t[dc_t^{FIN}]}{dt} - \beta_c \frac{d\bar{c}_t^{FIN}}{dt} = \frac{1}{\gamma}([i - \pi] - \rho)(c_t^{FIN} - \beta_c \bar{c}_t)$$

Since $\frac{d\bar{c}_t}{dt} = -\theta(\bar{c}_t - c_t^{FIN})$, the financier agents' consumption thus evolves according to

$$\frac{\mathbb{E}_t[dc_t^{FIN}]}{dt} = \gamma^{-1}([i-\pi]-\rho^*)(c_t^{FIN} - \beta_c\bar{c}_t) - \beta_c\theta_c(\bar{c}_t^{FIN} - c_t^{FIN}) \tag{21}$$

B Appendix to Section 6: Persistence Intuition

B.1 Persistence and Unemployment

Recall that from the hiring firm's problem, labor market tightness evolves according to equation (6):

$$\frac{\mathbb{E}_t[d\theta_t]}{dt}\frac{1}{\theta_t} = \frac{1}{\iota}(r_t + \lambda_{EU})\left[1 + \frac{\eta_l}{\eta_v}\lambda_{v,t}\right] - \frac{1}{\iota}(r_t^l - w_t)h_t\left(\frac{\lambda_{v,t}}{\eta_v}\right)$$

where in equilibrium $\lambda_{v,t} = \Psi\theta_t^{-\iota}$. Log linearizing this expression around the steady-state with $\theta_{NSS} = 1$, $\theta_t = e^{\widehat{\theta}_t}$, $(r_t^l - w_t) = (r_{NSS}^l - w_{NSS})\exp(\widehat{r_t^l - w_t})$, $r_t = \widehat{r}_t + r_{NSS}$, and $h_t = h_{NSS}e^{\widehat{h}_t}$ yields

$$\begin{aligned}\frac{\mathbb{E}_t[d\widehat{\theta}_t]}{dt} =& \frac{r_{NSS}}{\iota}(1 + \frac{\eta_l}{\eta_v}\Psi)\widehat{r}_t - \frac{\Psi}{\eta_v}\left[(r_{NSS} + \lambda_{EU})\eta_l - (r_{NSS}^l - w_{NSS})h_{NSS}\right]\widehat{\theta}_t \\ &- \frac{\Psi}{\iota\eta_v}(r_{NSS}^l - w_{NSS})h_{NSS}(\widehat{[r_t - w_t]} + \widehat{h}_t)\end{aligned}$$

Additionally, note that in the non-stochastic steady-state,

$$(r_{NSS} + \lambda_{EU})\left[1 + \frac{\eta_l}{\eta_v}\Psi\right] = (r_{NSS}^l - w_{NSS})h_{NSS}\left(\frac{\Psi}{\eta_v}\right)$$

$$r_{NSS} + \lambda_{EU} = \frac{\Psi}{\eta_v}\left[(r_{NSS}^l - w_{NSS})h_{NSS} - (r_{NSS} + \lambda_{EU})\eta_l\right]$$

such that the labor market tightness dynamics simplify to

$$\frac{\mathbb{E}_t[d\widehat{\theta}_t]}{dt} = \frac{r_{NSS}}{\iota}(1 + \frac{\eta_l}{\eta_v}\Psi)\widehat{r}_t + (r_{NSS} + \lambda_{EU})\widehat{\theta}_t - \frac{\Psi}{\iota\eta_v}(r_{NSS}^l - w_{NSS})h_{NSS}(\widehat{[r_t - w_t]} + \widehat{h}_t)$$

Integrating this equation forward from t to $+\infty$ and applying a transversality condition, labor market tightness may be expressed (in log deviation terms) as

$$\widehat{\theta}_t = \frac{1}{\iota}\mathbb{E}_t\int_t^{\infty} e^{-(r+\lambda_{EU})(s-t)}\left[\frac{\Psi}{\eta_v}(r_{NSS}^l - w_{NSS})h_{NSS}(\widehat{[r_s - w_s]} + \widehat{h}_s) - \left(1 + \frac{\eta_l}{\eta_v}\right)r_{NSS}\widehat{r}_s\right]ds$$

Labor market tightness may be expressed as the future present value (additionally discounted by the job termination rate) of future profits per worker, minus interest costs, all times the relevant elasticities.

To consider the dynamics of unemployment, first note that in level terms, the unemployment rate follows

$$\dot{u}_t = \lambda_{EU}(1 - u_t) - \lambda_{UE,t}u_t$$

Linearizing this expression with $\widehat{u} = u_t - u_{NSS}$ and $\lambda_{UE,t} = \Psi\theta_t^{1-\iota}$,

$$\frac{d\widehat{u}_t}{dt} = -\lambda_{EU}\widehat{u}_t - (1-\iota)\Psi u_{NSS}\widehat{\theta}_t - \Psi\widehat{u}_t.$$

Integrating forward from $s = 0$ to t,

$$\widehat{u}_t = \widehat{u}_0 - \Psi(1-\iota)u_{NSS} \int_0^t e^{-(\lambda_{EU}+\Psi)(t-s)} \widehat{\theta}_s ds.$$

Suppose the economy starts in the non-stochastic steady-state, such that $\widehat{u}_0 = 0$. Then, plugging in the formula for $\widehat{\theta}_s$ (along the deterministic transition dynamic):

$$\begin{aligned}\widehat{u}_t = & - u_{NSS}\frac{\Psi}{\iota} \int_0^t e^{-(\lambda_{EU}+\Psi)(t-s)} \int_s^\infty e^{-(r+\lambda_{EU})\tau} \\ & \left[\frac{\Psi}{\eta_v}(r^l_{NSS} - w_{NSS})h_{NSS}(\widehat{[r_\tau - w_\tau]} + \widehat{h}_\tau) - \left(1 + \frac{\eta_l}{\eta_v}\right) r_{NSS}\widehat{r}_\tau\right] d\tau \ ds.\end{aligned}$$

B.1.1 An exponential impulse

Suppose $(\widehat{r^l_t - w_t}) + \widehat{h}_t = Ae^{-\rho_1 t}$, while $\widehat{r}_t = 0$. Then labor market tightness will be

$$\begin{aligned}\widehat{\theta}_t = & \frac{1}{\iota}\frac{\Psi}{\eta_v}(r^l_{NSS} - w_{NSS})h_{NSS}A\mathbb{E}_t \int_t^\infty e^{-(r+\lambda_{EU})(s-t)}e^{-\rho_1 s}ds \\ = & \frac{1}{\iota}\frac{\Psi}{\eta_v}(r^l_{NSS} - w_{NSS})h_{NSS}A\frac{e^{-\rho_1 t}}{r + \lambda_{EU} + \rho_1}\end{aligned}$$

Plugging this into the unemployment expression,

$$\begin{aligned}\widehat{u}_t &= -\Psi(1-\iota)u_{NSS}\frac{1}{\iota}\frac{\Psi}{\eta_v}(r^l_{NSS} - w_{NSS})h_{NSS}A \int_0^t e^{-(\lambda_{EU}+\Psi)(t-s)} \frac{e^{-\rho_1 s}}{r + \lambda_{EU} + \rho_1} ds \\ &= -\Psi(1-\iota)u_{NSS}\frac{1}{\iota}\frac{\Psi}{\eta_v}(r^l_{NSS} - w_{NSS})h_{NSS}A \frac{e^{-\rho_1 t} - e^{-(\Psi+\lambda_{EU})t}}{(r + \lambda_{EU} + \rho_1)(r + \Psi - \rho_1)}.\end{aligned}$$

B.2 Persistence and Inflation

Consider the NKPC

$$\frac{\mathbb{E}[d\pi_t]}{dt} = r_t(\pi_t - \overline{\pi}_t) + \frac{\varepsilon}{\theta_\pi}\left[\frac{\varepsilon - 1}{\varepsilon} - m_t\right] + \kappa(\pi_t - \overline{\pi}_t),$$

where

$$\frac{d\overline{\pi}_t}{dt} = \kappa(\pi_t - \overline{\pi}_t)$$

and $\overline{\pi}_0 = 0$, $\mathbb{E}_t[\lim_{T\to\infty} e^{-\int_t^T r_s ds}\pi_T] = 0$. Rewriting the expression using $x_t \equiv \pi_t - \overline{\pi}_t$, such that

$$\frac{\mathbb{E}[dx_t]}{dt} = r_t x + \frac{\varepsilon}{\theta_\pi}\left[\frac{\varepsilon - 1}{\varepsilon} - m_t\right],$$

such that the ODE can be integrated forward to write

$$x_t = \frac{\varepsilon}{\theta_\pi}\int_t^\infty e^{-\int_t^s r_\tau d\tau}\left[m_s - \frac{\varepsilon - 1}{\varepsilon}\right] ds$$

The increase of inflation from its lagged values is proportional to the present value of the future marginal cost gaps.

From there, note that

$$\dot{\pi}_t = \dot{x}_t + \kappa x_t$$

while log-linearizing (with $x_t = 0$ in the NSS),

$$\dot{x}_t = r x_t - \frac{\varepsilon - 1}{\theta_\pi} \widehat{m}_t$$

From the first equation, assuming $\lim_{T\to\infty} \pi_T = 0$,

$$x_t = \frac{\varepsilon - 1}{\theta_\pi} \int_t^\infty e^{-r(s-t)} \widehat{m}_s ds$$

The backward-looking inflation average term is then

$$\overline{\pi}_t = \int_0^t \dot{\overline{\pi}}_s ds = \int_0^t \kappa x_s ds = \frac{\varepsilon - 1}{\theta_\pi} \kappa \int_0^t \int_s^\infty e^{-r(\tau - s)} \widehat{m}_\tau d\tau \; ds$$

and changing the bounds of the integral,

$$= \frac{\varepsilon - 1}{\theta_\pi} \frac{\kappa}{r} \left[\int_0^t (1 - e^{-rs}) \widehat{m}_s ds + \int_t^\infty (e^{-r(s-t)} - e^{-rs}) \widehat{m}_s ds \right]$$

Using the fact that

$$\pi_t = x_t + \overline{\pi}_t,$$

it follows that inflation is

$$\pi_t = \frac{\varepsilon - 1}{\theta_\pi} \int_t^\infty e^{-r(s-t)} \widehat{m}_s ds + \frac{\varepsilon - 1}{\theta_\pi} \frac{\kappa}{r} \left[\int_0^t (1 - e^{-rs}) \widehat{m}_s ds + \int_t^\infty (e^{-r(s-t)} - e^{-rs}) \widehat{m}_s ds \right].$$

Simplifying the algebra, inflation is then the sum of both a backward-looking and a forward-looking component.

$$\pi_t = \frac{\varepsilon - 1}{\theta_\pi} \frac{\kappa}{r} \left[\int_0^t (1 - e^{-rs}) \widehat{m}_s ds + \int_t^\infty \left[\left(1 + \frac{r}{\kappa} \right) e^{-r(s-t)} - e^{-rs} \right] \widehat{m}_s ds \right].$$

If the marginal costs jump and follow an exponential decay according to $\widehat{m}_t = m_0 e^{-\rho_2 s}$, then inflation at each point in time will be

$$\pi_t = \frac{\varepsilon - 1}{\theta_\pi} \frac{\kappa}{r} \left[\int_0^t (1 - e^{-rs}) m_0 e^{-\rho_2 s} ds + \int_t^\infty \left[\left(1 + \frac{r}{\kappa} \right) e^{-r(s-t)} - e^{-rs} \right] m_0 e^{-\rho_2 s} ds \right]$$

B.3 Determinacy for an exogenous impulse to marginal cost

The price-setting system can be expressed with three equations:

$$\frac{\mathbb{E}[d\pi_t]}{dt} = (r+\kappa)(\pi_t - \overline{\pi}_t) + \frac{\varepsilon - 1}{\theta_\pi}\widehat{m}_t$$

$$\frac{d\overline{\pi}_t}{dt} = \kappa(\pi_t - \overline{\pi}_t)$$

$$\frac{d\widehat{m}_t}{dt} = -\rho_2 \widehat{m}_t$$

Write the system as a matrix:

$$\begin{bmatrix} \mathbb{E}[d\pi_t] \\ d\overline{\pi}_t \\ d\widehat{m}_t \end{bmatrix} = \begin{bmatrix} r+\kappa & -r-\kappa & \frac{\varepsilon-1}{\theta_\pi} \\ \kappa & -\kappa & 0 \\ 0 & 0 & -\rho_2 \end{bmatrix} \begin{bmatrix} \pi_t \\ \overline{\pi}_t \\ \widehat{m}_t \end{bmatrix} dt$$

The characteristic polynomial is

$$\det(A - \lambda I) = \det \begin{bmatrix} r+\kappa-\lambda & -r-\kappa & \frac{\varepsilon-1}{\theta_\pi} \\ \kappa & -\kappa-\lambda & 0 \\ 0 & 0 & -\rho_2-\lambda \end{bmatrix}$$

$$= [(r+\kappa-\lambda)(-\kappa-\lambda) + \kappa(r+\kappa)](-\rho_2 - \lambda)$$

$$= \lambda(\lambda - r)(-\lambda - \rho_2)$$

Clearly, the roots are $-\rho_2 < 0$ (stable), $r > 0$ (explosive) and 0 (borderline).

C Model Robustness

Impulse responses to transfer shock by strength of nominal rigidities:

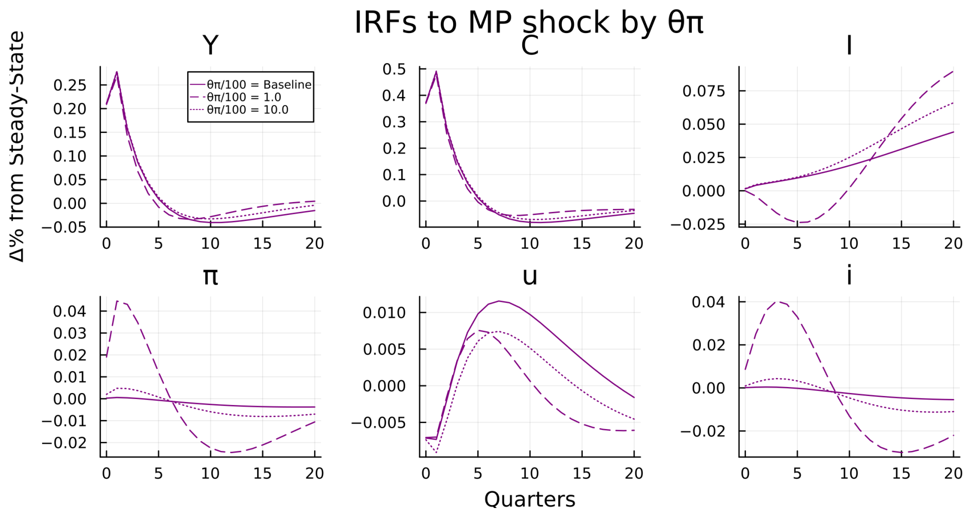

Figure 8: IRFs to a transfer shock, by degree of nominal rigidity θ_π.

Impulse responses to a transfer shock, by persistence of the transfer (rescaled to have the same net present value):

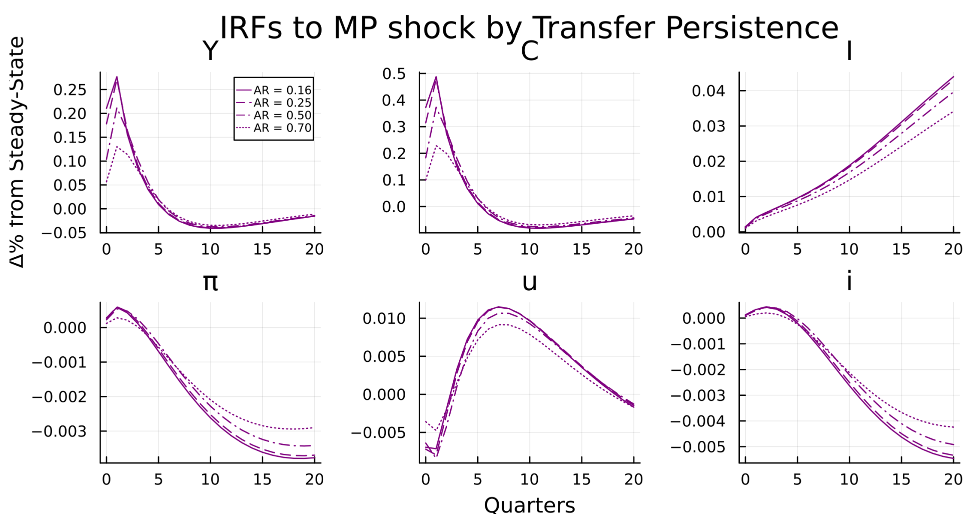

Figure 9: IRFs to a transfer shock, by the persistence of the shock (rescaled to have the same NPV).