DISCUSSION PAPER SERIES

DP14364
(v. 2)

SHOCKS, FRICTIONS, AND INEQUALITY IN US BUSINESS CYCLES

Christian Bayer, Benjamin Born and Ralph Luetticke

MONETARY ECONOMICS AND FLUCTUATIONS

ISSN 0265-8003

SHOCKS, FRICTIONS, AND INEQUALITY IN US BUSINESS CYCLES

Christian Bayer, Benjamin Born and Ralph Luetticke

Discussion Paper DP14364
First Published 29 January 2020
This Revision 26 November 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

SHOCKS, FRICTIONS, AND INEQUALITY IN US BUSINESS CYCLES

Abstract

How much does inequality matter for the business cycle and vice versa? Using a Bayesian likelihood approach, we estimate a heterogeneous-agent New-Keynesian (HANK) model with incomplete markets and portfolio choice between liquid and illiquid assets. Even when the model is estimated on aggregate data only and with only a set of shocks and frictions along the lines of Smets and Wouters (2007), it reproduces observed US inequality dynamics. In other words, the model is successful in simultaneously accounting for cycle and distribution. Output stabilization via monetary or fiscal policy also stabilizes inequality.

JEL Classification: E32, E63, C11

Keywords: incomplete markets, business cycles, Monetary and Fiscal Policy, Bayesian estimation, Wealth Inequality, Income inequality

Christian Bayer - christian.bayer@uni-bonn.de
Universität Bonn and CEPR

Benjamin Born - b.born@fs.de
Frankfurt School of Finance & Management and CEPR

Ralph Luetticke - r.luetticke@ucl.ac.uk
UCL and CEPR

Acknowledgements
We would like to thank Luigi Bocola, Gregor Boehl, Edouard Challe, Marco Del Negro, Miguel Leon-Ledesma, Makoto Nakajima, Johannes Pfeifer, Mu-Chun Wang, Maximilian Weiß, and participants at various conferences and seminars for helpful comments and suggestions. Christian Bayer gratefully acknowledges support through the ERC-CoG project Liquid-House-Cycle funded by the European Union's Horizon 2020 Program under grant agreement No. 724204. We would like to thank Sharanya Pillai and Luis Calderon for excellent research assistance. Codes are available at https://github.com/BenjaminBorn/HANK_BusinessCycleAndInequality.

Shocks, Frictions, and Inequality in US Business Cycles*

Christian Bayer, Benjamin Born, and Ralph Luetticke

November 26, 2020

Abstract

How much does inequality matter for the business cycle and vice versa? Using a Bayesian likelihood approach, we estimate a heterogeneous-agent New-Keynesian (HANK) model with incomplete markets and portfolio choice between liquid and illiquid assets. Even when the model is estimated on aggregate data only and with only a set of shocks and frictions along the lines of Smets and Wouters (2007), it reproduces observed US inequality dynamics. In other words, the model is successful in simultaneously accounting for cycle and distribution. Output stabilization via monetary or fiscal policy also stabilizes inequality.

JEL codes: C11, D31, E32, E63

Keywords: Bayesian Estimation, Business Cycles, Income and Wealth Inequality, Incomplete Markets, Monetary and Fiscal Policy

---

*Bayer: University of Bonn, CEPR, CESifo, and IZA, christian.bayer@uni-bonn.de, Born: Frankfurt School of Finance & Management, CEPR, and CESifo, b.born@fs.de, Luetticke: University College London, CEPR, and CFM, r.luetticke@ucl.ac.uk. We would like to thank Luigi Bocola, Gregor Boehl, Edouard Challe, Marco Del Negro, Miguel Leon-Ledesma, Makoto Nakajima, Johannes Pfeifer, Mu-Chun Wang, Maximilian Weiß, and participants at various conferences and seminars for helpful comments and suggestions. Christian Bayer gratefully acknowledges support through the ERC-CoG project Liquid-House-Cycle funded by the European Union's Horizon 2020 Program under grant agreement No. 724204. We would like to thank Sharanya Pillai and Luis Calderon for excellent research assistance. Codes are available at `https://github.com/BenjaminBorn/HANK_BusinessCycleAndInequality`.

1 Introduction

A new generation of monetary business cycle models has become popular, featuring heterogeneous agents and incomplete markets (known as HANK models). This new class of models implies new transmission channels of monetary[1] and fiscal[2] policy, as well as new sources of business cycle fluctuations working through household portfolio decisions.[3] Much of this literature so far has focused on specific channels of transmission, shocks, or puzzles. In contrast, the present paper takes a more encompassing approach. Concretely, we answer three questions: First, to what extent does the inclusion of incomplete markets change our view of US business cycles? Second, can the model even account for US business cycle and inequality dynamics simultaneously? Third, which business cycle shocks and policies are important drivers of US inequality dynamics?

For this purpose, we study the business cycle using a technique that has become standard at least since Smets and Wouters' (2007) seminal paper, extending this technique to the analysis of HANK models: We estimate an incomplete markets model by a full information Bayesian likelihood approach using the state-space representation of the model. Specifically, we estimate an extension of the New-Keynesian incomplete markets model of Bayer et al. (2019). We add features such as capacity utilization, a frictional labor market with sticky wages, and progressive taxation, as well as the battery of shocks that drive business cycle fluctuations in estimated New-Keynesian models: aggregate and investment-specific productivity shocks, wage- and price-markup shocks, monetary- and fiscal-policy shocks, risk-premium shocks, and, as two additional incomplete-market-specific ones, shocks to the progressivity of taxes and shocks to idiosyncratic productivity risk.

In this model, precautionary motives play an important role for consumption-savings decisions. Since individual income is subject to idiosyncratic risk that cannot be directly insured and borrowing is constrained, households structure their savings decisions and portfolio allocations to optimally self-insure and achieve consumption smoothing. In particular, we assume that households can either hold liquid nominal bonds or invest in illiquid physical capital. Capital is illiquid because its market is segmented and households participate only from time to time. This portfolio-choice component, which gives rise to an endogenous liquidity premium, and the presence of occasional hand-to-mouth consumers lead the HANK model to have rich distributional dynamics in response to aggregate shocks.

[1]Auclert (2019) analyzes the redistributive effects of monetary policy, Kaplan et al. (2018) show the importance of indirect income effects, and Luetticke (2020) analyzes the portfolio rebalancing channel of monetary policy. McKay et al. (2016) study the effectiveness of forward guidance.

[2]Auclert et al. (2018), Bayer et al. (2020), and Hagedorn et al. (2019) discuss fiscal multipliers, McKay and Reis (2016, 2020) discuss the role of automatic stabilizers.

[3]Bayer et al. (2019) quantify the importance of shocks to idiosyncratic income risk, and Guerrieri and Lorenzoni (2017) look at the effects of shocks to the borrowing limit.

To answer the first question, whether the inclusion of incomplete markets changes our view of US business cycles, we estimate the HANK model using the same set of aggregate shocks and observables as in Smets and Wouters (2007), covering the time period of 1954 to 2019, and compare it to the representative household analogue (RANK). We find that both models tell a similar story about the US business cycle, but there are some differences. Shocks to the risk premium and monetary policy become more important for output and consumption growth at the expense of shocks to investment-specific technology. This holds true in particular for US recessions, for which the difference between the HANK and RANK model becomes larger. Changes in investment technology induce strong wealth effects via their effect on asset prices that lead to negative co-movement of investment with consumption in the short run. Key for this mechanism is household portfolio heterogeneity. However, when HANK and RANK are identically parameterized, the differences between the two models are more substantial. The estimation makes the dynamics of both models more similar. For example, the HANK model estimates real and nominal frictions that are smaller than those in the RANK model.

Turning to the second question, we find that the HANK model can simultaneously account for US business cycle and inequality dynamics from 1954 to 2019. Our model translates the business cycle shocks estimated from aggregate data into persistent movements in wealth and income inequality that are consistent with the U-shaped evolution of the shares of wealth and income held by the top 10% of US households.[4] Since our estimation so far does not use any cross-sectional information, the good fit of the cross-sectional data can be informally thought of as passing an over-identification test. In line with this, we find that what we infer about the aggregate shocks and frictions driving the US business cycle does not change when we include these distributional data in the set of observables.

Regarding the third question, what drives US inequality, we then re-estimate the model with two additional shocks that directly affect the distribution of income: the progressivity of taxes and idiosyncratic productivity risk, which are observable in the data. Including these shocks and the observables in the estimation provides a further improved fit of the shares of wealth and income held by the top 10% of households. Income risk, in particular, is important for income inequality and partly replaces risk-premium shocks in explaining aggregate consumption growth. Together, income risk and tax progressivity shocks explain one-fifth of fluctuations in aggregate consumption growth at business cycle frequencies.

[4] We focus on the top 10% shares taken from the World Inequality Database because these measures are available from the 1950s onward and are most consistent across alternative data sources such as the Survey of Consumer Finances. See Kopczuk (2015) and Bricker et al. (2016) for detailed comparisons of all available data sources.

We use this extended model for the historical decomposition of US inequality. We find that different shocks drive income and wealth inequality—measured by their top 10% shares. Since the 1980s, income inequality increased mostly because of higher profit margins (price markups), lower wage markups, and higher income risk. Together, these three shocks explain 90% of the increase in the top 10% income share from 1985 to 2019. By contrast, the two markup shocks and the income risk shock only explain one-third of the increase in wealth inequality. Shocks that directly affect asset returns, such as technology shocks, are central for understanding the evolution of wealth inequality. Key for this is portfolio heterogeneity between wealthy and poor households. Poor households predominantly hold liquid assets, while wealthy households hold their wealth in illiquid form. Take, for example, a negative shock to investment technology; this drives up the price of illiquid capital, but at the same time, reduces the future return to capital. Hence, the current owners of illiquid assets enjoy capital gains, while saving by the poor is discouraged because of the lower returns.

We also assess the importance of policy shocks and rules in shaping inequality over the business cycle. Monetary and fiscal policy shocks have contributed to the increase in inequality since the 1980s, in particular wealth inequality. We then assess counterfactual monetary and fiscal policy rules. Changing policy rules has a sizable and persistent effect on inequality. For example, a more hawkish monetary policy, i.e., a stronger reaction to inflation, would have led to output losses and increased income inequality in the 1970s with a protracted increase in wealth inequality until the 2000s. A more active monetary or fiscal policy response to the Great Recession would have decreased inequality and would have increased output. More generally, output stabilization via monetary or fiscal policy also typically stabilizes inequality.

To our knowledge, our paper is the first to provide an encompassing estimation of shocks and frictions using a HANK model with portfolio choice. Most of the literature on monetary heterogeneous-agent models has used a calibration approach.[5] Auclert et al. (2020) and Hagedorn et al. (2018) go beyond calibration but use one-asset HANK models. The latter provide parameter estimates based on impulse response function matching, while the former estimate the model using the MA-∞ representation in the sequence space. Using a state-space approach is key for us, as we need to deal with mixed-frequency data.

Our paper is also related to Chang et al. (2018) in the sense that it estimates a state-space model of both distributional (cross-sectional) data and aggregates. Chang et al. (2018) find that, in an SVAR sense, shocks to the cross-sectional distribution of income have only

[5]See, for example, Auclert et al. (2018); Ahn et al. (2018); Bayer et al. (2019); Broer et al. (2019); Challe and Ragot (2015); Den Haan et al. (2017); Ferriere and Navarro (2018); Gornemann et al. (2012); Guerrieri and Lorenzoni (2017); McKay et al. (2016); McKay and Reis (2016); Ravn and Sterk (2017); Sterk and Tenreyro (2018); Wong (2019).

a mild impact on aggregate time series. Our finding of structural estimates being relatively robust to the inclusion or exclusion of cross-sectional information resembles their results.[6] The estimated importance of cross-sectional shocks is also in line with the finding of Berger et al. (2020) who use a business-cycle accounting approach.

Our findings provide new insights into the literature on the drivers of inequality.[7] Kaymak and Poschke (2016) and Hubmer et al. (2020) use quantitative models to study permanent changes in the US tax and transfer system and the variance of income. In terms of methods, these papers solve for steady-state transitions of calibrated models, while we estimate our model on US macro and micro time-series data. They find that tax and transfer changes can explain a significant part of the recent increase in wealth inequality. Ours is the first paper to quantify the distributional consequences of all standard business cycle shocks and estimate their importance in explaining US inequality. Having both changes in the US tax system and changes in income risk alongside standard business cycle shocks in our model, allows us to compare their relative importance for the evolution of income and wealth inequality. We find business cycle shocks, through their effect on asset prices, to be more important for wealth inequality than changes in taxes and income risk, but the latter two play a significant role for income inequality.

Focusing on the methodological contribution, Auclert et al. (2019) provide a fast estimation method for heterogeneous-agent models that requires a sequence space representation of the model and thus does not allow us to deal with missing or mixed-frequency data as we need to do here, when combining cross-sectional and aggregate data. Since this is the setup we are facing, we build on the solution method of Reiter (2009) using the dimensionality reduction approach of Bayer and Luetticke (2020) to make this feasible for estimation. We further exploit the fact that only a small fraction of the Jacobian of the non-linear difference equation that represents the model needs to be re-calculated during the estimation.

The remainder of this paper is organized as follows: Section 2 describes our model economy, its sources of fluctuations and frictions. It also provides details on the numerical solution method and estimation technique. Section 3 presents our model variants, the parameters that we calibrate to match steady-state targets, and prior distributions for the remaining parameters that we estimate. It also gives an overview of the data we employ in our estimation. Section 4 discusses the estimated shocks and frictions driving the US business cycle. Section 5 does so for US inequality. Section 6 concludes. An Appendix follows.

[6]Our approach is different and simpler than the method suggested by Liu and Plagborg-Møller (2019), which includes full cross-sectional information in the estimation of a heterogeneous-agent DSGE model. We, in contrast, only use the model to fit certain generalized cross-sectional moments.

[7]There is a growing literature on inequality dynamics. On the theory side, see, e.g., Gabaix et al. (2016). On the empirical side, see, e.g., Heathcote et al. (2010), Piketty and Saez (2003) or Saez and Zucman (2016).

2 Model

We model an economy composed of a firm sector, a household sector, and a government sector. The firm sector comprises (a) perfectly competitive intermediate goods producers who rent out labor services and capital; (b) final goods producers that face monopolistic competition, producing differentiated final goods out of homogeneous intermediate inputs; (c) producers of capital goods that turn consumption goods into capital subject to adjustment costs; (d) labor packers that produce labor services combining differentiated labor from (e) unions that differentiate raw labor rented out from households. Price setting for the final goods as well as wage setting by unions is subject to a pricing friction à la Calvo (1983).

Households earn income from supplying (raw) labor and capital and from owning the firm sector, absorbing all its rents that stem from the market power of unions and final goods producers, and decreasing returns to scale in capital goods production.

The government sector runs both a fiscal authority and a monetary authority. The fiscal authority levies taxes on labor income and distributed profits, issues government bonds, and adjusts expenditures to stabilize debt in the long run and aggregate demand in the short run. The monetary authority sets the nominal interest rate on government bonds according to a Taylor rule.

2.1 Households

The household sector is subdivided into two types of agents: workers and entrepreneurs. The transition between both types is stochastic. Both rent out physical capital, but only workers supply labor. The efficiency of a worker's labor evolves randomly exposing worker-households to labor-income risk. Entrepreneurs do not work, but earn all pure rents in our economy except for the rents of unions which are equally distributed across workers. All households self-insure against the income risks they face by saving in a liquid nominal asset (bonds) and a less liquid asset (capital). Trading illiquid assets is subject to random participation in the capital market.

To be specific, there is a continuum of ex-ante identical households of measure one, indexed by i. Households are infinitely lived, have time-separable preferences with time-discount factor β, and derive felicity from consumption c_{it} and leisure. They obtain income from supplying labor, n_{it}, from renting out capital, k_{it}, and from earning interest on bonds, b_{it}, and potentially from profits or union transfers. Households pay taxes on labor and profit income.

2.1.1 Productivity, labor supply, and labor income

A household's gross labor income $w_t n_{it} h_{it}$ is composed of the aggregate wage rate on raw labor, w_t, the household's hours worked, n_{it}, and its idiosyncratic labor productivity, h_{it}. We assume that productivity evolves according to a log-AR(1) process with time-varying volatility and a fixed probability of transition between the worker and the entrepreneur state:

$$\tilde{h}_{it} = \begin{cases} \exp\left(\rho_h \log \tilde{h}_{it-1} + \epsilon^h_{it}\right) & \text{with probability } 1-\zeta \text{ if } h_{it-1} \neq 0, \\ 1 & \text{with probability } \iota \text{ if } h_{it-1} = 0, \\ 0 & \text{else,} \end{cases} \tag{1}$$

with individual productivity $h_{it} = \frac{\tilde{h}_{it}}{\int \tilde{h}_{it} di}$ such that $\tilde{h}_{it}$ is scaled by its cross-sectional average, $\int \tilde{h}_{it} di$, to make sure that average worker productivity is constant. The shocks ϵ^h_{it} to productivity are normally distributed with variance $\sigma^2_{h,t}$.[8] With probability ζ households become entrepreneurs ($h = 0$). With probability ι an entrepreneur returns to the labor force with median productivity. In our baseline specification, an entrepreneur obtains a share of the pure rents (aside from union rents), Π^F_t, in the economy (from monopolistic competition in the goods sector and the creation of capital). We assume that the claim to the pure rent cannot be traded as an asset. Union rents, Π^U_t are distributed lump sum across workers, leading to labor-income compression. For tractability, we assume union profits to be taxed at a fixed rate independent of the recipient's labor income.

This modeling strategy serves two purposes. First and foremost, it generally solves the problem of the allocation of pure rents without distorting factor returns and without introducing another tradable asset.[9] Second, we use the entrepreneur state in particular – a

[8] In our baseline, we treat the variance as time fixed. We consider an extension where income risk follows a log-AR(1) process with endogenous feedback to aggregate output Y_{t+1} (hats denote log-deviations from the steady state):

$$\begin{aligned} \sigma^2_{h,t} &= \bar{\sigma}^2_h \exp \widehat{s}_t, \\ \widehat{s}_{t+1} &= \rho_s \widehat{s}_t + \Sigma_Y \widehat{Y}_{t+1} + \epsilon^\sigma_t \ , \end{aligned} \tag{1a}$$

i.e., at time t households observe a change in the variance of shocks that drive the next period's productivity.

[9] There are basically three possibilities for dealing with the pure rents: 1) Attribute them to capital and labor, but this affects their factor prices; 2) introduce a third asset that pays out rents as dividends and is priced competitively; or 3) distribute the rents in the economy to an exogenously determined group of households. The latter has the advantage that factor supply decisions remain the same as in any standard New-Keynesian framework and still avoids the numerical complexity of dealing with three assets. If one is willing to assume that pure and capital rents come in illiquid form, as in Kaplan et al. (2018), pure rents can be priced using the rate of return on capital. However, this approach requires not only assets (claims on physical capital and on future pure rents) to be illiquid, but also their corresponding asset income.

transitory state in which incomes are very high – to match the income and wealth distribution following the idea by Castaneda et al. (1998). The entrepreneur state does not change the asset returns or investment opportunities available to households. Our assumption on the distribution of pure rents implies that higher market power in the goods market leads to a higher concentration of income and higher mark-ups in the labor market leads to wage compression.[10]

With respect to leisure and consumption, households have Greenwood et al. (1988) (GHH) preferences and maximize the discounted sum of felicity:[11]

$$\mathbb{E}_0 \max_{\{c_{it}, n_{it}\}} \sum_{t=0}^{\infty} \beta^t u \left[c_{it} - G(h_{it}, n_{it}) \right]. \tag{2}$$

The maximization is subject to the budget constraints described further below. The felicity function u exhibits a constant relative risk aversion (CRRA) with risk aversion parameter $\xi > 0$,

$$u(x_{it}) = \frac{x_{it}^{1-\xi} - 1}{1 - \xi},$$

where $x_{it} = c_{it} - G(h_{it}, n_{it})$ is household i's composite demand for goods consumption c_{it} and leisure and G measures the disutility from work. Goods consumption bundles varieties j of differentiated goods according to a Dixit-Stiglitz aggregator:

$$c_{it} = \left(\int c_{ijt}^{\frac{\eta_t - 1}{\eta_t}} dj \right)^{\frac{\eta_t}{\eta_t - 1}}.$$

Each of these differentiated goods is offered at price p_{jt}, so that for the aggregate price level, $P_t = \left(\int p_{jt}^{1-\eta_t} dj \right)^{\frac{1}{1-\eta_t}}$, the demand for each of the varieties is given by

$$c_{ijt} = \left(\frac{p_{jt}}{P_t} \right)^{-\eta_t} c_{it}.$$

Assuming a (progressive) income-tax schedule (which we borrow from Benabou, 2002;

[10] Boar and Midrigan (2019) use a similar structure, where entrepreneurs retain a fraction of firm profits and thus the size of markups has an impact on inequality in the economy.

[11] The assumption of GHH preferences is mainly motivated by the fact that many estimated DSGE models of business cycles find small aggregate wealth effects in the labor supply; see, e.g., Schmitt-Grohé and Uribe (2012); Born and Pfeifer (2014). Unfortunately, it is not feasible to estimate the flexible form of preference of Jaimovich and Rebelo (2009), which also encompasses King et al. (1988) (KPR) preferences. This would require solving the stationary equilibrium in every likelihood evaluation, which is substantially more time consuming than solving for the dynamics around this equilibrium. However, we estimate a version with KPR preferences; see Section 4.5 and also Appendix F.2. The data clearly prefer the GHH specification over the KPR specification. What is more, the KPR version of the HANK model has more difficulty matching business cycle and inequality dynamics simultaneously.

Heathcote et al., 2017), a household's net labor income, y_{it}, is given by

$$y_{it} = (1-\tau_t^L)(w_t h_{it} n_{it})^{1-\tau_t^P} \ , \tag{3}$$

where w_t is the aggregate wage rate and τ_t^L and τ_t^P determine the level and the progressivity of the tax code. Given net labor income, the first-order condition for labor supply is

$$\frac{\partial G(h_{it}, n_{it})}{\partial n_{it}} = (1-\tau_t^P)(1-\tau_t^L)(w_t h_{it})^{1-\tau_t^P} n_{it}^{-\tau_t^P} = (1-\tau_t^P)\frac{y_{it}}{n_{it}} \ . \tag{4}$$

Assuming that G has a constant elasticity w.r.t. n, $\frac{\partial G(h_{it}, n_{it})}{\partial n_{it}} = (1+\gamma)\frac{G(h_{it}, n_{it})}{n_{it}}$ with $\gamma > 0$, we can simplify the expression for the composite consumption good, x_{it}, making use of this first-order condition (4), and substitute $G(h_{it}, n_{it})$ out of the individual planning problem:

$$x_{it} = c_{it} - G(h_{it}, n_{it}) = c_{it} - \frac{1-\tau_t^P}{1+\gamma} y_{it}. \tag{5}$$

When the Frisch elasticity of labor supply is constant and the tax schedule has the form (3), the disutility of labor is always a fraction of labor income and constant across households. Therefore, in both the household's budget constraint and felicity function, only after-tax income enters and neither hours worked nor productivity appears separately.

What remains to be determined is individual and aggregate effective labor supply. Without further loss of generality, we assume $G(h_{it}, n_{it}) = h_{it}^{1-\bar{\tau}^P} \frac{n_{it}^{1+\gamma}}{1+\gamma}$, where $\bar{\tau}^P$ is the stationary equilibrium level of progressivity of the tax code. This functional form simplifies the household problem in the stationary equilibrium as h_{it} drops out from the first-order condition and all households supply the same number of hours $n_{it} = N(w_t)$. Total effective labor input, $\int n_{it} h_{it} di$, is hence also equal to $N(w_t)$ because we normalized $\int h_{it} di = 1$.

Importantly, this means that we can read off average productivity risk from the estimated income risk series in the literature. Without scaling the labor disutility by productivity, we would need to translate productivity risk to income risk through the endogenous hour response. In our baseline, we keep tax progressivity constant, $\tau_t^P = \bar{\tau}^P$.[12]

[12] In an extension, we allow it to vary over time. There, when tax progressivity does not coincide with its stationary equilibrium value, individual hours worked differ across agents and are given by

$$n_{it} = \left[(1-\tau_t^P)(1-\tau_t^L)\right]^{\frac{1}{\gamma+\tau_t^P}} h_{it}^{\frac{\bar{\tau}^P-\tau_t^P}{\gamma+\tau_t^P}} w_t^{\frac{1-\tau_t^P}{\gamma+\tau_t^P}} \ , \tag{4a}$$

such that aggregate effective hours are given by

$$N_t = \int n_{it} h_{it} = \left[(1-\tau_t^P)(1-\tau_t^L)\right]^{\frac{1}{\gamma+\tau_t^P}} w_t^{\frac{1-\tau_t^P}{\gamma+\tau_t^P}} \underbrace{\int h_{it}^{\frac{\gamma+\bar{\tau}^P}{\gamma+\tau_t^P}}}_{:=H_t} \ . \tag{4b}$$

Household after-tax labor income, plugging in the optimal supply of hours, is then:

$$y_{it} = (1-\tau_t^L)(w_t h_{it} n_{it})^{1-\tau_t^P} = (1-\tau_t^L)^{\frac{1+\gamma}{\gamma+\tau_t^P}} (1-\tau_t^P)^{\frac{1-\tau_t^P}{\gamma+\tau_t^P}} w_t^{\frac{1+\gamma}{\gamma+\tau_t^P}(1-\tau_t^P)} h_{it}^{\frac{\gamma+\bar{\tau}^P}{\gamma+\tau_t^P}(1-\tau_t^P)} . \tag{6}$$

2.1.2 Consumption, savings, and portfolio choice

Given this labor income, households optimize intertemporally subject to their budget constraint:

$$\begin{aligned} c_{it} + b_{it+1} + q_t k_{it+1} = b_{it} \tfrac{R(b_{it}, R_t^b, A_t)}{\pi_t} + (q_t + r_t) k_{it} + y_{it} \\ + \mathbb{I}_{h_{it} \neq 0}(1-\tau_t)\Pi_t^U + \mathbb{I}_{h_{it}=0}(1-\tau_t^L)(\Pi_t^F)^{1-\tau_t^P}, \qquad k_{it+1} \geq 0, \quad b_{it+1} \geq \underline{B} \, , \end{aligned}$$

where Π_t^U is union profits taxed at the average tax rate τ_t (see Equation 23), Π_t^F is firm profits, b_{it} is real bond holdings, k_{it} is the amount of illiquid assets, q_t is the price of these assets, r_t is their dividend, $\pi_t = \frac{P_t}{P_{t-1}}$ is realized inflation, and R is the nominal interest rate on bonds, which depends on the portfolio position of the household and the central bank's interest rate R_t^b, which is set one period before. All households that do not participate in the capital market ($k_{it+1} = k_{it}$) still obtain dividends and can adjust their bond holdings. Depreciated capital has to be replaced for maintenance, such that the dividend, r_t, is the net return on capital. Holdings of bonds have to be above an exogenous debt limit $\underline{B}$, and holdings of capital have to be non-negative.

Substituting the expression $c_{it} = x_{it} + \frac{1-\tau_t^P}{1+\gamma} y_{it}$ for consumption, we obtain the budget constraint for the composite leisure-consumption good:

$$\begin{aligned} x_{it} + b_{it+1} + q_t k_{it+1} = b_{it} \tfrac{R(b_{it}, R_t^b, A_t)}{\pi_t} + (q_t + r_t) k_{it} + \tfrac{\tau_t^P + \gamma}{1+\gamma} y_{it} \\ + \mathbb{I}_{h_{it} \neq 0}(1-\tau_t)\Pi_t^U + \mathbb{I}_{h_{it}=0}(1-\tau_t^L)(\Pi_t^F)^{1-\tau_t^P}, \qquad k_{it+1} \geq 0, \quad b_{it+1} \geq \underline{B} \, . \end{aligned} \tag{7}$$

Households make their savings choices and their portfolio choice between liquid bonds and illiquid capital in light of a capital market friction that renders capital illiquid because participation in the capital market is random and i.i.d. in the sense that only a fraction, λ, of households are selected to be able to adjust their capital holdings in a given period.

What is more, we assume that there is a wasted intermediation cost that drives a wedge between the government bond yield R_t^b and the interest paid by/to households R_t. This wedge, A_t, is given by a time-varying term plus a constant, $\overline{R}$, when households resort to

Here H_t measures how the tax progressivity influences the (hours-weighted) average labor productivity. Scaling of the disutility of labor by $h_{it}^{1-\bar{\tau}^P}$ is thus a normalization of H_t to one in the stationary equilibrium.

unsecured borrowing. This means that we specify:

$$R(b_{it}, R_t^b, A_t) = \begin{cases} R_t^b A_t & \text{if } b_{it} \geq 0 \\ R_t^b A_t + \overline{R} & \text{if } b_{it} < 0. \end{cases}$$

The extra wedge for unsecured borrowing creates a mass of households with zero unsecured credit but with the possibility to borrow, though at a penalty rate. The "risk-premium shock" A_t (as in, for example, Smets and Wouters, 2007) is technically modeled here as an intermediation efficiency wedge, the cost of a banking sector turning government bonds into deposits.[13] This cost follows an AR(1) process in logs and fluctuates in response to shocks, ϵ_t^A. If A_t goes down, households will implicitly demand fewer government bonds and find it more attractive to save in (illiquid) real capital.

Since a household's saving decision—(b'_a, k') for the case of adjustment and (b'_n, k) for non-adjustment—will be some non-linear function of that household's wealth and productivity, inflation and all other prices will be functions of the joint distribution, Θ_t, of (b, k, h) in t. This makes Θ a state variable of the household's planning problem and this distribution evolves as a result of the economy's reaction to aggregate shocks. For simplicity, we summarize all effects of aggregate state variables, including the distribution of wealth and income, by writing the dynamic planning problem with time-dependent continuation values.

This leaves us with three functions that characterize the household's problem: value function V^a for the case where the household adjusts its capital holdings, the function V^n for the case in which it does not adjust, and the expected continuation value, $\mathbb{W}$, over both:

$$\begin{aligned} V_t^a(b, k, h) =& \max_{b'_a, k'} u[x(b, b'_a, k, k', h)] + \beta \mathbb{E}_t \mathbb{W}_{t+1}(b'_a, k', h') \\ V_t^n(b, k, h) =& \max_{b'_n} u[x(b, b'_n, k, k, h)] + \beta \mathbb{E}_t \mathbb{W}_{t+1}(b'_n, k, h') \\ \mathbb{W}_{t+1}(b', k', h') =& \lambda V_{t+1}^a(b', k', h') + (1-\lambda) V_{t+1}^n(b', k', h') \end{aligned} \tag{8}$$

Expectations about the continuation value are taken with respect to all stochastic processes conditional on the current states. Maximization is subject to the corresponding budget constraint.

2.2 Firm sector

The firm sector consists of four sub-sectors: (a) a labor sector composed of "unions" that differentiate raw labor and labor packers who buy differentiated labor and then sell labor

[13] In a representative agent framework, this strategy and modeling A_t as a preference shock are equivalent.

services to intermediate goods producers, (b) intermediate goods producers who hire labor services and rent out capital to produce goods, (c) final goods producers who differentiate intermediate goods and then sell them to goods bundlers, who finally sell them as consumption goods to households, and to (d) capital goods producers, who turn bundled final goods into capital goods.

When profit maximization decisions in the firm sector require intertemporal decisions (i.e. in price and wage setting and in producing capital goods), we assume for tractability that they are delegated to a mass-zero group of households (managers) that are risk neutral and compensated by a share in profits.[14] They do not participate in any asset market and have the same discount factor as all other households. Since managers are a mass-zero group in the economy, their consumption does not show up in any resource constraint and all but the unions' profits go to the entrepreneur households (whose $h = 0$). Union profits go lump sum to worker households. In our baseline all profit incomes are paid out in the very period they are earned. In an extension, we introduce retained earnings such that the profit payments of unions and firms to households are smoothed over time.

2.2.1 Labor packers and unions

Worker households sell their labor services to a mass-one continuum of unions indexed by j, each of whom offers a different variety of labor to labor packers who then provide labor services to intermediate goods producers. Labor packers produce final labor services according to the production function

$$N_t = \left(\int \hat{n}_{jt}^{\frac{\zeta_t - 1}{\zeta_t}} \, dj \right)^{\frac{\zeta_t}{\zeta_t - 1}}, \tag{9}$$

out of labor varieties $\hat{n}_{jt}$. Cost minimization by labor packers implies that each variety of labor, each union j, faces a downward-sloping demand curve

$$\hat{n}_{jt} = \left(\frac{W_{jt}}{W_t^F} \right)^{-\zeta_t} N_t,$$

where W_{jt} is the *nominal* wage set by union j and W_t^F is the nominal wage at which labor packers sell labor services to final goods producers.

Since unions have market power, they pay the households a wage lower than the price

[14] Since we solve the model by a first-order perturbation in aggregate shocks, the assumption of risk-neutrality only serves as a simplification in terms of writing down the model. With a first-order perturbation we have certainty equivalence and time fluctuations in stochastic discount factors become irrelevant. Of course in the cross-section, households have different stochastic discount factors depending on whether they are borrowing-constrained or not. Yet, this does not change the dynamics of the Phillips curves up to a first-order approximation beyond the steady-state non-stochastic discount factor.

at which they sell labor to labor packers. Given the nominal wage W_t at which they buy labor from households and given the *nominal* wage index W_t^F, unions seek to maximize their discounted stream of profits. However, they face a Calvo type (1983) of adjustment friction with indexation with the probability λ_w to keep wages constant. They therefore maximize

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \lambda_w^t \frac{W_t^F}{P_t} N_t \left\{ \left(\frac{W_{jt}\bar{\pi}_W^t}{W_t^F} - \frac{W_t}{W_t^F} \right) \left(\frac{W_{jt}\bar{\pi}_W^t}{W_t^F} \right)^{-\zeta_t} \right\}, \tag{10}$$

by setting W_{jt} in period t and keeping it constant except for indexation to $\bar{\pi}_W$, the steady-state wage inflation rate.

Since all unions are symmetric, we focus on a symmetric equilibrium and obtain the linearized wage Phillips curve from the corresponding first-order condition as follows, leaving out all terms irrelevant at a first-order approximation around the stationary equilibrium:

$$\log\left(\frac{\pi_t^W}{\bar{\pi}_W}\right) = \beta \mathbb{E}_t \log\left(\frac{\pi_{t+1}^W}{\bar{\pi}_W}\right) + \kappa_w \left(mc_t^w - \frac{1}{\mu_t^W} \right), \tag{11}$$

with $\pi_t^W := \frac{W_t^F}{W_{t-1}^F} = \frac{w_t^F}{w_{t-1}^F}\pi_t^Y$ being wage inflation, w_t and w_t^F being the respective *real* wages for households and firms, $mc_t^w = \frac{w_t}{w_t^F}$ is the actual and $\frac{1}{\mu_t^W} = \frac{\zeta_t - 1}{\zeta_t}$ being the target mark-down of wages the unions pay to households, W_t, relative to the wages charged to firms, W_t^F and $\kappa_w = \frac{(1-\lambda_w)(1-\lambda_w\beta)}{\lambda_w}$. This target fluctuates in response to markup shocks, $\epsilon_t^{\mu W}$, and follows a log AR(1) process.[15] In our baseline, profits paid to households therefore are $\Pi_t^U = (w_t^F - w_t)N_t$.

2.2.2 Final goods producers

Similar to unions, final goods producers differentiate a homogeneous intermediate good and set prices. They face a downward-sloping demand curve

$$y_{jt} = (p_{jt}/P_t)^{-\eta_t} Y_t$$

for each good j and buy the intermediate good at the nominal price MC_t. As we do for unions, we assume price adjustment frictions à la Calvo (1983) with indexation.

Under this assumption, the firms' managers maximize the present value of real profits

[15] Including the first-order irrelevant terms, the Phillips curve reads

$$\log\left(\frac{\pi_t^W}{\bar{\pi}^W}\right) = \beta \mathbb{E}_t \left[\log\left(\frac{\pi_{t+1}^W}{\bar{\pi}^W}\right) \frac{1-\tau_{t+1}}{1-\tau_t} \frac{\zeta_{t+1}}{\zeta_t} \frac{W_{t+1}^F P_t}{W_t^F P_{t+1}} \frac{N_{t+1}}{N_t} \right] + \kappa_w \left(\frac{w_t}{w_t^F} - \frac{1}{\mu_t^W} \right),$$

where τ_t is the average income tax.

given this price adjustment friction, i.e., they maximize:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \lambda_Y^t (1-\tau_t^L) Y_t^{1-\tau_t^P} \left\{ \left(\frac{p_{jt}\bar{\pi}_Y^t}{P_t} - \frac{MC_t}{P_t} \right) \left(\frac{p_{jt}\bar{\pi}^t}{P_t} \right)^{-\eta_t} \right\}^{1-\tau_t^P}, \tag{12}$$

with a time-constant discount factor.

The corresponding first-order condition for price setting implies a Phillips curve

$$\log\left(\frac{\pi_t}{\bar{\pi}}\right) = \beta \mathbb{E}_t \log\left(\frac{\pi_{t+1}}{\bar{\pi}}\right) + \kappa_Y \left(mc_t - \frac{1}{\mu_t^Y} \right), \tag{13}$$

where we again dropped all terms irrelevant for a first-order approximation and have $\kappa_Y = \frac{(1-\lambda_Y)(1-\lambda_Y \beta)}{\lambda_Y}$. Here, π_t is the gross inflation rate of final goods, $\pi_t := \frac{P_t}{P_{t-1}}$, $mc_t := \frac{MC_t}{P_t}$ is the real marginal costs, $\bar{\pi}$ is steady-state inflation, and $\mu_t^Y = \frac{\eta_t}{\eta_t - 1}$ is the target markup. As for the unions, this target fluctuates in response to markup shocks, $\epsilon^{\mu Y}$, and follows a log AR(1) process. In our baseline, profits paid to households therefore are $\Pi_t^F = (1 - mc_t)Y_t$.[16]

2.2.3 Intermediate goods producers

Intermediate goods are produced with a constant returns to scale production function:

$$Y_t = Z_t N_t^{\alpha} (u_t K_t)^{(1-\alpha)},$$

where Z_t is total factor productivity and follows an autoregressive process in logs, and $u_t K_t$ is the effective capital stock taking into account utilization u_t, i.e., the intensity with which the existing capital stock is used. Using capital with an intensity higher than normal results in increased depreciation of capital according to $\delta(u_t) = \delta_0 + \delta_1 (u_t - 1) + \delta_2/2 \, (u_t - 1)^2$, which, assuming $\delta_1, \delta_2 > 0$, is an increasing and convex function of utilization. Without loss of generality, capital utilization in the steady state is normalized to 1, so that δ_0 denotes the steady-state depreciation rate of capital goods.

Let mc_t be the relative price at which the intermediate good is sold to final goods producers. The intermediate goods producer maximizes profits,

$$mc_t Z_t Y_t - w_t^F N_t - [r_t + q_t \delta(u_t)] \, K_t,$$

where r_t^F and q_t are the rental rate of firms and the (producer) price of capital goods re-

[16] In an extension, we allow for smoothed payments of profits. Here, firms retain a fraction $1-\omega^F$ of their period profits to buy bonds and pay out to entrepreneurs a fraction ω^F of retained earnings. Also, unions then retain a fraction $1-\omega^U$ of their period profits to buy bonds and pay out to workers a fraction ω^U of retained earnings.

spectively. The intermediate goods producer operates in perfectly competitive markets, such that the real wage and the user costs of capital are given by the marginal products of labor and effective capital:

$$w_t^F = \alpha mc_t Z_t \left(\frac{u_t K_t}{N_t}\right)^{1-\alpha}, \tag{14}$$

$$r_t + q_t \delta(u_t) = u_t (1-\alpha) mc_t Z_t \left(\frac{N_t}{u_t K_t}\right)^{\alpha}. \tag{15}$$

We assume that utilization is decided by the owners of the capital goods, taking the aggregate supply of capital services as given. The optimality condition for utilization is given by

$$q_t \left[\delta_1 + \delta_2 (u_t - 1)\right] = (1-\alpha) mc_t Z_t \left(\frac{N_t}{u_t K_t}\right)^{\alpha}, \tag{16}$$

i.e., capital owners increase utilization until the marginal maintenance costs equal the marginal product of capital services.

2.2.4 Capital goods producers

Capital goods producers take the relative price of capital goods, q_t, as given in deciding about their output, i.e., they maximize

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t I_t \left\{ \Psi_t q_t \left[1 - \frac{\phi}{2}\left(\log \frac{I_t}{I_{t-1}}\right)^2\right] - 1 \right\}, \tag{17}$$

where Ψ_t governs the marginal efficiency of investment à la Justiniano et al. (2011), which follows an AR(1) process in logs and is subject to shocks ϵ_t^{Ψ}.[17]

Optimality of the capital goods production requires (again dropping all terms irrelevant up to first order)

$$\Psi_t q_t \left[1 - \phi \log \frac{I_t}{I_{t-1}}\right] = 1 - \beta \mathbb{E}_t \left[\Psi_{t+1} q_{t+1} \phi \log \left(\frac{I_{t+1}}{I_t}\right)\right], \tag{18}$$

and each capital goods producer will adjust its production until (18) is fulfilled.

Since all capital goods producers are symmetric, we obtain as the law of motion for

[17] This shock has to be distinguished from a shock to the relative price of investment, which has been shown in the literature (Justiniano et al., 2011; Schmitt-Grohé and Uribe, 2012) to not be an important driver of business cycles as soon as one includes the relative price of investment as an observable. We therefore focus on the MEI shock.

aggregate capital

$$K_t - (1 - \delta(u_t))\, K_{t-1} = \Psi_t \left[1 - \frac{\phi}{2} \left(\log \frac{I_t}{I_{t-1}} \right)^2 \right] I_t \ . \tag{19}$$

The functional form assumption implies that investment adjustment costs are minimized and equal to 0 in the steady state.

2.3 Government

The government operates a monetary and a fiscal authority. The monetary authority controls the nominal interest rate on liquid assets, while the fiscal authority issues government bonds to finance deficits, chooses both the average tax rate in the economy and the tax progressivity, and makes expenditures.

We assume that monetary policy sets the nominal interest rate following a Taylor-type (1993) rule with interest rate smoothing:

$$\frac{R^b_{t+1}}{\bar{R}^b} = \left(\frac{R^b_t}{\bar{R}^b} \right)^{\rho_R} \left(\frac{\pi_t}{\bar{\pi}} \right)^{(1-\rho_R)\theta_\pi} \left(\frac{Y_t}{Y^*_t} \right)^{(1-\rho_R)\theta_Y} \epsilon^R_t . \tag{20}$$

The coefficient $\bar{R}^b \geq 0$ determines the nominal interest rate in the steady state. The coefficients $\theta_\pi, \theta_Y \geq 0$ govern the extent to which the central bank attempts to stabilize inflation and the output gap, where the gap, $\frac{Y_t}{Y^*_t}$, is defined relative to what output would be at stationary equilibrium markups, Y^*_t. $\rho_R \geq 0$ captures interest rate smoothing.

We assume that government debt evolves according to the following rule (c.f. Woodford, 1995):

$$\frac{B_{t+1}}{B_t} = \left(\frac{B_t}{\bar{B}} \right)^{-\gamma_B} \left(\frac{\pi_t}{\bar{\pi}} \right)^{\gamma_\pi} \left(\frac{Y_t}{Y^*_t} \right)^{\gamma_Y} D_t \ , \qquad D_t = D^{\rho_D}_{t-1} \epsilon^D_t \, , \tag{21}$$

where D_t is a persistent shock to the government's structural deficit. Besides issuing bonds, the government uses tax revenues T_t, defined below, to finance government consumption, G_t, and interest on debt. The parameters γ_B, γ_π, and γ_Y measure, respectively, how the deficit reacts to outstanding debt, inflation, and the output gap.

The government sets the average tax rate in the economy according to a similar rule

$$\frac{\tau_t}{\bar{\tau}} = \left(\frac{\tau_{t-1}}{\bar{\tau}} \right)^{\rho_\tau} \left(\frac{B_t}{\bar{B}} \right)^{(1-\rho_\tau)\gamma^\tau_B} \left(\frac{Y_t}{Y^*_t} \right)^{(1-\rho_\tau)\gamma^\tau_Y} \ . \tag{22}$$

The level parameter of the tax code τ^L_t adjusts such that the average tax rate on income equals this target level:

$$\tau_t = \frac{\mathbb{E}_t\left(w_t n_{it} h_{it} + \mathbb{I}_{h_{it}=0}\Pi_t^F\right) - \tau_t^L \mathbb{E}_t\left(w_t n_{it} h_{it} + \mathbb{I}_{h_{it}=0}\Pi_t^F\right)^{\tau_t^P}}{\mathbb{E}_t\left(w_t n_{it} h_{it} + \mathbb{I}_{h_{it}=0}\Pi_t^F\right)}\ , \tag{23}$$

where $\mathbb{E}_t$ is the expectation operator, which here gives the cross-sectional average, and τ_t^P is constant in the baseline.[18] Total government tax revenues T_t are then

$$T_t = \tau_t\left(w_t n_{it} h_{it} + \mathbb{I}_{h_{it}\neq 0}\Pi_t^U + \mathbb{I}_{h_{it}=0}\Pi_t^F\right)$$

and the government budget constraint determines government spending residually: $G_t = B_{t+1} + T_t - R_t^b/\pi_t B_t$.

There are thus two shocks to government rules: monetary policy shocks, ϵ_t^R, and structural deficit shocks, ϵ_t^D. We assume these shocks to be log normally distributed with mean zero.

2.4 Goods, bonds, capital, and labor market clearing

The labor market clears at the competitive wage given in (14). The bond market clears whenever the following equation holds:

$$B_{t+1} = B^d(R_t^b, A_t, r_t, q_t, \Pi_t^F, \Pi_t^U, w_t, \pi_t, \tau_t, \tau_t^P, \tau_t^L, \Theta_t, \mathbb{W}_{t+1}) := \mathbb{E}_t\left[\lambda b_{a,t}^* + (1-\lambda) b_{n,t}^*\right], \tag{24}$$

where $b_{a,t}^*, b_{n,t}^*$ are functions of the states (b, k, h), and depend on how households value asset holdings in the future, $\mathbb{W}_{t+1}(b, k, h)$, and the current set of prices (and tax rates) $(R_t^b, A_t, r_t, q_t, \Pi_t^F, \Pi_t^U, w_t, \pi_t, \tau_t, \tau_t^P, \tau_t^L)$. Future prices do not show up because we can express the value functions such that they summarize all relevant information on the expected future price paths. Expectations in the right-hand-side expression are taken w.r.t. the distribution $\Theta_t(b, k, h)$. Equilibrium requires the total *net* amount of bonds the household sector demands, B^d, to equal the supply of government bonds. In gross terms there are more liquid assets in

[18] In an extension, the parameter τ_t^P that governs the progressivity of the tax schedule evolves according to

$$\frac{\tau_t^P}{\bar{\tau}^P} = \left(\frac{\tau_{t-1}^P}{\bar{\tau}^P}\right)^{\rho_P} \epsilon_t^P\ , \tag{21a}$$

where ϵ_t^P are shocks to tax progressivity.

circulation as some households borrow up to $\underline{B}$.[19]

Last, the market for capital has to clear:

$$K_{t+1} = K^d(R_t^b, A_t, r_t, q_t, \Pi_t^F, \Pi_t^U, w_t, \pi_t, \tau_t, , \tau_t^P, \tau_t^L, \Theta_t, \mathbb{W}_{t+1}) := \mathbb{E}_t[\lambda k_t{}^* + (1-\lambda)k] \;, \quad (25)$$

where the first equation stems from competition in the production of capital goods, and the second equation defines the aggregate supply of funds from households—both those that trade capital, λk_t^*, and those that do not, $(1-\lambda)k$. Again k_t^* is a function of the current prices and continuation values. The goods market then clears due to Walras' law, whenever labor, bonds, and capital markets clear.

2.5 Equilibrium

A *sequential equilibrium with recursive planning* in our model is a sequence of policy functions $\{x_{a,t}^*, x_{n,t}^*, b_{a,t}^*, b_{n,t}^*, k_t^*\}$, a sequence of value functions $\{V_t^a, V_t^n\}$, a sequence of prices $\{w_t, w_t^F, \Pi_t^F, \Pi_t^U, q_t, r_t, R_t^b, \pi_t, \pi_t^W, \tau_t, \tau_t^P, \tau_t^L\}$, a sequence of stochastic states $A_t, Z_t, \Psi_t, \mu_t, \mu W_t, D_t$ and shocks $\epsilon_t^A, \epsilon_t^Z, \epsilon_t^\Psi, \epsilon_t^{\mu Y}, \epsilon_t^{\mu W}, \epsilon_t^D, \epsilon_t^R$, aggregate capital and labor supplies $\{K_t, N_t\}$, distributions Θ_t over individual asset holdings and productivity, and expectations Γ for the distribution of future prices, such that

1. Given the functionals $\mathbb{E}_t\mathbb{W}_{t+1}$ for the continuation value and period-t prices, policy functions $\{x_{a,t}^*, x_{n,t}^*, b_{a,t}^*, b_{n,t}^*, k_t^*\}$ solve the households' planning problem; and given the policy functions $\{x_{a,t}^*, x_{n,t}^*, b_{a,t}^*, b_{n,t}^*, k_t^*\}$ and prices, the value functions $\{V_t^a, V_t^n\}$ are a solution to the Bellman equation (8).

2. Distributions of wealth and income evolve according to households' policy functions.

3. The labor, the final goods, the bond, the capital, and the intermediate goods markets clear in every period, interest rates on bonds are set according to the central bank's

[19] In the extension with retained earnings, see Footnote 16, the bond market clearing condition reads:

$$\begin{aligned} B_{t+1} &= B^d + B_{t+1}^U + B_{t+1}^F \\ B_{t+1}^U &= (1-\omega^U)\left(B_t^U \frac{R_t^b}{\pi_t} + (\bar{mc}^w - mc_t^w) w_t N_t\right) \\ B_{t+1}^F &= (1-\omega^F)\left(B_t^F \frac{R_t^b}{\pi_t} + (\bar{mc} - mc_t) Y_t\right), \end{aligned} \quad (24a)$$

where $B^{U,F}$ is demand for bonds from unions and firms. This amounts to payments to workers of $\Pi_t^U = \omega^U\left(B_t^U \frac{R_t^b}{\pi_t} + (1-mc_t^w) w_t N_t\right) + (1-\omega^U)(1-\bar{mc}^w) w_t N_t$ and to entrepreneurs of $\Pi_t^F = \omega^F\left(B_t^F \frac{R_t^b}{\pi_t} + (1-mc_t) Y_t\right) + (1-\omega^F)(1-\bar{mc}) Y_t$.

Taylor rule, fiscal policies are set according to the fiscal rules, and stochastic processes evolve according to their law of motion.

4. Expectations are model consistent.

2.6 Numerical solution and estimation technique

We solve the model by perturbation methods. We choose a first-order Taylor expansion around the stationary equilibrium following the method of Bayer and Luetticke (2020). This method replaces the value functions with linear interpolants and the distribution function with histograms to calculate a stationary equilibrium. Then it performs dimensionality reduction before linearization but after calculation of the stationary equilibrium. The dimensionality reduction is achieved by first writing the distribution of assets and income in terms of its copula and marginals, then representing the value functions and copula by their discrete cosine transformations (DCT) and finally perturbing only the largest coefficients of these transformations together with the marginal distributions of asset holdings and income. We solve the model originally on a grid of 80x80x22 points for liquid assets, illiquid assets, and income, respectively. The dimensionality-reduced number of states and controls in our baseline system is 929.

Approximating the sequential equilibrium in a linear state-space representation then boils down to the linearized solution of a non-linear difference equation

$$\mathbb{E}_t F(x_t, X_t, x_{t+1}, X_{t+1}, \sigma\Sigma\epsilon_{t+1}), \tag{26}$$

where x_t is "idiosyncratic" states and controls: the value and distribution functions, and X_t is aggregate states and controls: prices, quantities, productivities, etc. The error term ϵ_t represents fundamental shocks. Importantly, we can also order the equations in a similar way. The law of motion for the distribution and the Bellman equations describe a non-linear difference equation for the idiosyncratic variables, and all other optimality and market clearing conditions describe a non-linear difference equation for the aggregate variables. By introducing auxiliary variables that capture the mean of b, k, and h, we make sure that the distribution itself does not directly show up in any aggregate equation other than in the one for the summary variables. Yet, these equations are free of all model parameters.

This helps substantially in estimating the model. For each parameter draw, we need to calculate the Jacobian of F and then use the Klein algorithm (2000) (see also Schmitt-Grohé and Uribe, 2004) to obtain a linear state-space representation, which we then feed into a Kalman filter to obtain the likelihood of the data given our model. However, most model parameters do not show up in the Bellman equation. Only $\rho_h, \bar{\sigma}_h, \lambda, \beta, \gamma$, and ξ do, but we do not estimate these parameters but calibrate them from the stationary equilibrium.[20] Therefore, the Jacobian of the "idiosyncratic equations" is unaltered by all parameters that we estimate and we only need to calculate it once. Similarly, "idiosyncratic variables" (i.e., the value functions and the histograms) only affect the aggregate equations through their parameter-free effect on summary variables, such that this part of the Jacobian also does not need to be updated during the estimation.[21] This leaves us with the same number of derivatives to be calculated for every parameter draw during the estimation as in a representative agent model.[22] Still, solving for the state-space representation and evaluating the likelihood are substantially more time consuming, and computing the likelihood of a given parameter draw takes roughly 5 seconds on a workstation computer; 90% of the computing time goes into the Schur decomposition, which is still much larger because of the many additional "idiosyncratic" states (histograms) and controls (marginal value functions) the system contains.

We use a Bayesian likelihood approach as described in An and Schorfheide (2007) and Fernández-Villaverde (2010) for parameter estimation. In particular, we use the Kalman filter to obtain the likelihood from the state-space representation of the model solution and employ a standard random walk Metropolis-Hastings (RWMH) algorithm to generate draws from the posterior likelihood. Smoothed estimates of the states at the posterior mean of the parameters are obtained via a Kalman smoother of the type described in Koopman and Durbin (2000) and Durbin and Koopman (2012).

3 Model variants and parameterization

We estimate a number of model variants using a two-step procedure. First, we calibrate or fix all parameters that affect the steady state of the model, which is the same for all model variants (see Table 1). Second, we estimate by full-information methods all parameters that only matter for the dynamics of the model, i.e., the aggregate shocks and frictions. Table 2 summarizes the calibrated and externally chosen parameters, Table 3 shows the calibration targets, and Table 4 lists the estimated parameters. One period in the model refers to a quarter of a year.

[20]Note that the scaling of idiosyncratic risk, s_t, shows up in the Bellman equation, but similar to a price and not as a parameter.

[21]Note that this does not stop the underlying heterogeneity from influencing the estimation. The counterparts in a representative agent framework are the consumption Euler and the capital accumulation equations, which do not change in the estimated parameters; but of course that does not imply that they are unimportant for the estimation.

[22]Auclert et al. (2019) exploit a similar fact for estimation based on the sequence-space representation of a HANK model.

3.1 Model variants

Table 1: Model variants

Shocks \ Data	Aggregate	+Cross-sectional
Aggregate	HANK RANK	HANKX
+Cross-sectional		HANKX+

Table 1 gives an overview of the model variants that we estimate. The most basic model variant with complete markets and representative household (RANK) naturally is estimated on aggregate data only. As in Smets and Wouters (2007), we allow for aggregate shocks to TFP and investment-specific technology, to price and wage markups, to the risk premium and monetary policy, and to deficits.

Next, we add incomplete markets and portfolio choices to the model (HANK) and estimate this variant on the same data and the same sources of fluctuations. Beyond understanding the impact of incomplete markets on the business cycle, we use this model to obtain out-of-sample predictions for inequality. The third model we estimate adds cross-sectional information to the estimation sample but keeps the same sources of fluctuations (HANKX). We use this variant as an over-identification test, asking whether there is a tension between aggregate and cross-sectional information given the model.

Finally, we extend the shocks and allow for time variation in income risk and tax progressivity that directly affect the cross-section (HANKX+). To discipline the two processes we include additional data on income risk and tax progressivity. This variant also allows for retained earnings by unions and firms.[23]

3.2 Calibrated parameters

We fix a number of parameters either following the literature or targeting steady-state ratios; see Table 2 (all at quarterly frequency of the model).[24] For the household side, we set the relative risk aversion, ξ, to 4, which is common in the incomplete markets literature; see Kaplan and Violante (2014).[25] We set the Frisch elasticity to 0.5; see Chetty et al. (2011).

[23]See Footnotes 8, 12, 18, and 19 for the adjustments to the baseline model.

[24]Appendix B provides the table of steady-state parameters for the representative agent analogue of our model.

[25]We also estimate our model with a relative risk aversion of 2, see Appendix F.1.

Table 2: External/calibrated parameters (quarterly frequency)

Parameter	Value	Description	Target
Households			
β	0.984	Discount factor	see Table 3
ξ	4.00	Relative risk aversion	Kaplan and Violante (2014)
γ	2.00	Inverse of Frisch elasticity	Chetty et al. (2011)
λ	10.00%	Portfolio adj. prob.	see Table 3
ρ_h	0.98	Persistence labor income	Storesletten et al. (2004)
σ_h	0.12	STD labor income	Storesletten et al. (2004)
ζ	0.02%	Trans. prob. from W. to E.	see Table 3
ι	6.25%	Trans. prob. from E. to W.	Guvenen et al. (2014)
$\bar{R}$	1.71%	Borrowing penalty	see Table 3
Firms			
α	0.68	Share of labor	62% labor income
δ_0	1.75%	Depreciation rate	7.0% p.a.
$\bar{\eta}$	11	Elasticity of substitution	Price markup 10%
$\bar{\zeta}$	11	Elasticity of substitution	Wage markup 10%
Government			
$\bar{\tau}^L$	0.175	Tax rate level	$G/Y = 20\%$
$\bar{\tau}^P$	0.12	Tax progressivity	see text
$\bar{R}^b$	1.00	Nominal rate	see text
$\bar{\pi}$	1.00	Inflation	see text

We take estimates for idiosyncratic income risk from Storesletten et al. (2004), $\rho_h = 0.98$ and $\bar{\sigma}_h = 0.12$. Guvenen et al. (2014) provide the probability that a household will fall out of the top 1% of the income distribution in a given year, which we take as the transition probability from entrepreneur to worker, $\iota = 6.25\%$.

Table 3 summarizes the targeted moments. We match four targets: 1) average illiquid assets (K/Y=11.44), 2) average liquidity (B/Y=1.72), 3) the fraction of borrowers, 16%, and 4) the average top 10% share of wealth, which is 67%. This yields a discount factor of 0.984, a portfolio adjustment probability of 10%, a borrowing penalty of 1.71% quarterly (given a borrowing limit of two times average quarterly income), and a transition probability from worker to entrepreneur of 0.02%.[26]

For the firm side, we set the labor share in production, α, to 68% to match a labor income share of 62%, which corresponds to the average BLS labor share measure over 1954-2019. The depreciation rate is 1.75% per quarter. An elasticity of substitution between differentiated goods of 11 yields a markup of 10%. The elasticity of substitution between labor varieties is also set to 11, yielding a wage markup of 10%. All are standard values in the literature.

The government taxes labor and profit income using a non-linear tax schedule that approx-

[26]Detailed data sources can be found in Appendix A.

Table 3: Targeted moments

Targets	Model	Data	Source	Parameter
Mean illiquid assets (K/Y)	11.44	11.44	NIPA	Discount factor
Mean liquidity (B/Y)	1.72	1.72	FRED	Portfolio adj. probability
Top10 wealth share	0.67	0.67	WID	Fraction of entrepreneurs
Fraction borrowers	0.16	0.16	SCF	Borrowing penalty

imates the progressivity of the US tax system; see Heathcote et al. (2017). The progressivity parameter, $\tau^P = 0.12$, corresponds to the average value over 1954 – 2017. We follow Ferriere and Navarro (2018) in constructing a direct estimate for tax progressivity and extend their tax progressivity estimate until 2017. The level of taxes, τ^L, is set to clear the government budget constraint that corresponds to a government share of $G/Y = 20\%$. We set steady-state inflation to zero as we have assumed indexation to the steady-state inflation rate in the Phillips curves. We set the steady-state net interest rate on bonds to 0.0%, in order to capture the average federal funds rate in real terms minus output growth over 1954 – 2019.

3.3 Estimation data

We use quarterly US data from 1954Q3 to 2019Q4 and include the following eight observable time series in the baseline: the growth rates of per capita GDP, private consumption, investment, and wages, all in real terms; the logarithm of the level of per capita hours worked; the log difference of the GDP deflator; and the (shadow) federal funds rate. Our model is stationary so all growth rates are demeaned; see Appendix A.2 for a formal depiction of the vector of observables.

In the two extensions HANKX and HANKX+, we add more data with shorter and/or non-quarterly availability. First, we use cross-sectional information, wealth and income shares of the top 10% at an annual frequency, available from 1954 to 2019 from the World Inequality Database[27], but keep the set of parameters to be estimated unchanged (HANKX). The reason we focus on the top 10% wealth and income share is that this measure is most consistent across alternative, but less frequently available, data sources such as the Survey of Consumer Finances (SCF); see Kopczuk (2015).[28] Second, we allow for shocks to income risk and tax

[27]This database draws on work by Piketty, Saez, and Zucman; see, e.g., Piketty and Saez (2003) or Saez and Zucman (2016)

[28]We abstain from including other cross-sectional data in the estimation, such as the Panel Study of Income

progressivity (HANKX+). To identify these, we use a direct estimate for tax progressivity, extending the annual data series of Ferriere and Navarro (2018) based on Mertens and Montiel Olea (2018)'s measure of average marginal tax rates to 2017, see Appendix A.2. For income risk, we use estimates of the variance of idiosyncratic income, available at a quarterly frequency from 1983Q1 to 2013Q1, from Bayer et al. (2019) based on panel data in the Survey of Income and Program Participation (SIPP) .

3.4 Prior distributions

Columns 1-4 of Table 4 present the parameters we estimate and their assumed prior distributions. The posterior distribution is discussed in the next section. Where available, we use prior values that are standard in the literature and independent of the underlying data.

Following Justiniano et al. (2011), we impose a gamma distribution with prior mean of 5.0 and standard deviation of 2.0 for δ_2/δ_1, the elasticity of marginal depreciation with respect to capacity utilization, and a gamma prior with mean 4.0 and standard deviation of 2.0 for the parameter controlling investment adjustment costs, ϕ. For the slopes of price and wage Phillips curves, κ_Y and κ_w, we assume gamma priors with mean 0.1 and standard deviation 0.01, which corresponds to price and wage contracts having an average length of one year. Following Smets and Wouters (2007), the autoregressive parameters of the shock processes are assumed to follow a beta distribution with mean 0.5 and standard deviation 0.2. The standard deviations of the shocks follow inverse-gamma distributions with prior mean 0.1% and standard deviation 2%.[29]

Regarding policy, for the inflation and output feedback parameters in the Taylor rule, θ_π and θ_Y, we impose normal distributions with prior means of 1.7 and 0.13, respectively, while the interest rate smoothing parameter ρ_R has the same prior distribution as the persistence parameters of the shock processes. In the bond rule, the debt-feedback parameter γ_B is assumed to follow a gamma distribution with mean 0.10 such that the prior for the autocorrelation of debt is centered around 0.9 and standard deviation 0.08, implying a half-life of a deviation in debt of between one and eight years. The parameters governing feedback to inflation and output, γ_π and γ_Y, follow standard normal distributions. Similarly, the autoregressive parameters, in the tax rules, ρ_i where $i \in \{P, \tau\}$, are assumed to follow beta distributions (with mean 0.5 and standard deviation 0.2), while the feedback parameters, γ_Y^τ and γ_B^τ, follow standard normal distributions.

Dynamics (PSID) or the Survey of Consumer Finances (SCF) to avoid dealing with two measurements of the same model variable.

[29]In the HANKX+ extension, we use a higher prior mean for income risk shocks, s, given the evidence in Bayer et al. (2019).

In the HANKX+ extension, we allow for retained earnings and take as prior for the payment to workers or entrepreneurs, $\omega^{U,F}$, a uniform distribution between 0 and 1.

In our baseline we do not include measurement errors, but allow for these when including estimates of top income and wealth shares or tax progressivity and income risk as additional data.

Table 4: Prior and posterior distributions of estimated parameters

Parameter		Prior		Posterior			
	Distribution	Mean	Std. Dev.	RANK	HANK	HANKX	HANKX+
				Frictions			
δ_s	Gamma	5.00	2.00	1.929	1.226	1.274	1.289
				(1.450, 2.519)	(1.006, 1.468)	(1.061, 1.508)	(1.085, 1.511)
ϕ	Gamma	4.00	2.00	1.267	0.164	0.166	0.163
				(0.738, 2.160)	(0.119, 0.215)	(0.122, 0.218)	(0.120, 0.212)
κ	Gamma	0.10	0.01	0.083	0.097	0.097	0.097
				(0.069, 0.097)	(0.083, 0.112)	(0.083, 0.112)	(0.083, 0.112)
κ_w	Gamma	0.10	0.01	0.085	0.110	0.108	0.101
				(0.069, 0.103)	(0.092, 0.128)	(0.092, 0.125)	(0.085, 0.118)
				Debt and monetary policy rules			
ρ_R	Beta	0.50	0.20	0.818	0.788	0.791	0.799
				(0.792, 0.843)	(0.758, 0.816)	(0.761, 0.820)	(0.770, 0.827)
σ_R	Inv.-Gamma	0.10	2.00	0.226	0.257	0.258	0.256
				(0.208, 0.245)	(0.233, 0.283)	(0.234, 0.284)	(0.234, 0.282)
θ_π	Normal	1.70	0.30	2.106	2.457	2.493	2.525
				(1.906, 2.327)	(2.230, 2.708)	(2.259, 2.751)	(2.293, 2.779)
θ_Y	Normal	0.13	0.05	0.176	0.115	0.111	0.091
				(0.145, 0.209)	(0.087, 0.145)	(0.084, 0.139)	(0.063, 0.119)
γ_B	Gamma	0.10	0.08	0.442	0.294	0.297	0.148
				(0.332, 0.557)	(0.204, 0.399)	(0.223, 0.384)	(0.111, 0.190)
γ_π	Normal	0.00	1.00	-1.054	-1.078	-1.14	-1.147
				(-1.187, -0.922)	(-1.216, -0.945)	(-1.277, -1.005)	(-1.277, -1.021)
γ_Y	Normal	0.00	1.00	-0.852	-0.877	-0.834	-1.118
				(-0.955, -0.759)	(-0.992, -0.773)	(-0.942, -0.738)	(-1.262, -0.99)
ρ_D	Beta	0.50	0.20	0.988	0.994	0.992	0.989
				(0.978, 0.995)	(0.987, 0.998)	(0.983, 0.998)	(0.979, 0.996)
σ_D	Inv.-Gamma	0.10	2.00	0.242	0.291	0.308	0.287
				(0.203, 0.281)	(0.263, 0.323)	(0.278, 0.341)	(0.260, 0.317)
				Tax rules			
ρ_τ	Beta	0.50	0.20	0.294	0.401	0.405	0.481
				(0.123, 0.462)	(0.228, 0.551)	(0.246, 0.548)	(0.361, 0.598)
γ_B^τ	Normal	0.00	1.00	1.883	1.454	1.374	0.648
				(1.454, 2.365)	(1.085, 1.884)	(1.052, 1.750)	(0.498, 0.823)
γ_Y^τ	Normal	0.00	1.00	1.329	2.842	2.631	3.971
				(0.453, 2.259)	(1.915, 3.810)	(1.885, 3.443)	(3.237, 4.751)
ρ_P	Beta	0.50	0.20	–	–	–	0.971
							(0.950, 0.988)
σ_P	Inv.-Gamma	0.10	2.00	–	–	–	3.720
							(3.140, 4.407)

Table 4: Prior and posterior distributions of estimated parameters - continued

Parameter		Prior		Posterior			
	Distribution	Mean	Std. Dev.	RANK	HANK	HANKX	HANKX+
				Structural shocks			
ρ_A	Beta	0.50	0.20	0.903	0.943	0.967	0.981
				(0.879, 0.926)	(0.911, 0.973)	(0.935, 0.991)	(0.965, 0.994)
σ_A	Inv.-Gamma	0.10	2.00	0.245	0.199	0.180	0.143
				(0.211, 0.281)	(0.169, 0.230)	(0.152, 0.212)	(0.125, 0.164)
ρ_Z	Beta	0.50	0.20	0.996	0.997	0.995	0.995
				(0.994, 0.998)	(0.994, 0.999)	(0.991, 0.998)	(0.991, 0.998)
σ_Z	Inv.-Gamma	0.10	2.00	0.521	0.582	0.591	0.601
				(0.481, 0.563)	(0.540, 0.627)	(0.549, 0.636)	(0.559, 0.647)
ρ_Ψ	Beta	0.50	0.20	0.883	0.969	0.967	0.969
				(0.825, 0.928)	(0.953, 0.983)	(0.954, 0.978)	(0.956, 0.980)
σ_Ψ	Inv.-Gamma	0.10	2.00	5.246	2.158	2.171	2.004
				(4.338, 6.671)	(1.864, 2.484)	(1.895, 2.479)	(1.753, 2.278)
ρ_μ	Beta	0.50	0.20	0.915	0.899	0.902	0.898
				(0.876, 0.952)	(0.865, 0.929)	(0.867, 0.934)	(0.864, 0.929)
σ_μ	Inv.-Gamma	0.10	2.00	1.567	1.660	1.637	1.732
				(1.380, 1.785)	(1.487, 1.858)	(1.467, 1.832)	(1.546, 1.939)
$\rho_{\mu w}$	Beta	0.50	0.20	0.691	0.859	0.852	0.856
				(0.624, 0.752)	(0.814, 0.900)	(0.810, 0.890)	(0.814, 0.894)
$\sigma_{\mu w}$	Inv.-Gamma	0.10	2.00	8.640	5.930	6.085	6.968
				(7.010, 10.604)	(5.071, 6.939)	(5.256, 7.045)	(6.012, 8.063)
				Income risk process			
ρ_s	Beta	0.50	0.20	–	–	–	0.681
							(0.618, 0.736)
σ_s	Inv.-Gamma	1.00	2.00	–	–	–	68.826
							(61.173, 77.392)
Σ_Y	Normal	0.00	100.00	–	–	–	21.663
							(21.651, 21.674)
				Retained earnings processes			
ω^F	Uniform	0.50	0.29	–	–	–	0.115
							(0.074, 0.174)
ω^U	Uniform	0.50	0.29	–	–	–	0.153
							(0.126, 0.187)
				Measurement errors			
σ^{me}_{W10}	Inv.-Gamma	0.05	0.01	–	–	2.217	1.843
						(1.902, 2.577)	(1.552, 2.175)
σ^{me}_{I10}	Inv.-Gamma	0.05	0.01	–	–	6.243	3.647
						(5.405, 7.219)	(2.879, 4.580)
$\sigma^{me}_{\tau P}$	Inv.-Gamma	0.05	0.01	–	–	–	0.042
							(0.012, 0.110)
σ^{me}_{s}	Inv.-Gamma	0.05	0.01	–	–	–	5.045
							(3.604, 7.002)

Notes: The standard deviations of the shocks and measurement errors have been transformed into percentages by multiplying by 100. HANK and HANKX denote posterior estimates for the model without and with observable inequality series, respectively. HANKX+ additionally includes retained earnings, and shocks to income risk and tax progressivity and their observables.

4 US business cycles

One key advantage of HANK models is that we can use them to understand the distributional consequences of business cycle shocks and policies. This raises three questions. First, to what extent does the inclusion of incomplete markets change our view of the business cycle compared to a representative-agent model? Second, can the model simultaneously account for inequality and business cycle dynamics or does the model need to make compromises on the one to accommodate the other? Third, which business cycle shocks and policies are the most important drivers of inequality dynamics?

In this section, we answer the first two questions by comparing parameter estimates, variance decompositions, and historical decompositions of US business cycles for the estimated HANK model without (HANK) and with (HANKX) data on inequality and compare them to the otherwise equivalent representative-agent model (RANK) – with all three models being subject to aggregate shocks only. In Section 5, we dig deeper into the details of how the business cycle drives US inequality and add shocks to income risk and tax progressivity (using the HANKX+ model), giving an answer to the third question.

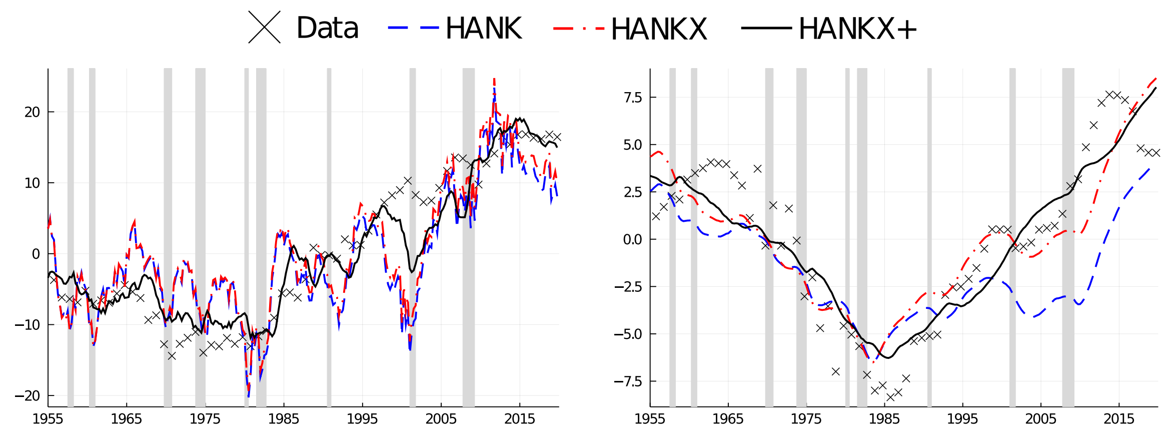
4.1 Estimation results

Table 4, columns 5-7, reports the posterior distributions across the three estimation variants: RANK, HANK, and HANKX. Checks on the convergence of the estimator are provided in Appendix C.[30] The parameter estimates are broadly in line with the RANK estimates (our priors are taken from this literature) and typically of the same order of magnitude, but there are notable differences. These differences reflect the fact that the HANK model under the RANK parameterization yields significantly different aggregate dynamics (see Appendix D.1), but the models need to fit the same data.

In particular, the estimated frictions in RANK are larger. The RANK model estimates larger investment adjustment costs and costs of capacity utilization. In part this reflects the portfolio adjustment costs at the household level that generate inertia in aggregate investment. Our estimates for nominal frictions for the HANK model imply price and wage adjustments roughly every four quarters on average. For the RANK model the implied average time between adjustments is slightly longer. In terms of shocks, the estimated persistence and variance for the seven "standard" shocks are comparable to the results of Smets

[30]The estimation is conducted with five parallel RWMH chains starting from an over-dispersed target distribution after an extensive mode search. After burn-in, 250,000 draws from the posterior are used to compute the posterior statistics. The acceptance rates across chains are between 20% and 30%. Appendix C provides Gelman and Rubin (1992) and Geweke (1992) convergence statistics as well as traceplots of individual parameters.

Figure 1: US inequality – data vs. model



(a) Top 10% income share

(b) Top 10% wealth share

Notes: Data (crosses) correspond to log-deviations of the annual observations of the share of pre-tax income and wealth held by the top 10% in each distribution in the US taken from the World Inequality Database. HANK (dashed) corresponds to the smoothed states of both implied by the estimated model w/o inequality data, HANKX (dash-dotted) with inequality data, and HANKX+ (solid) with inequality data, shocks to income risk and tax progressivity (plus their observables), as well as retained earnings. Shaded areas correspond to NBER-dated recessions.

and Wouters (2007). The persistence estimates under the HANK specification range from 0.997 for TFP to 0.859 for wage markups. The variance ranges from 0.2% for risk-premium shocks to 5.9% for wage-markup shocks. In terms of policy rules, we also obtain comparable results. The Taylor rule coefficients on inflation and output deviations are 2.5 and 0.1, and there is substantial inertia with a coefficient of 0.8 for interest smoothing. The fiscal rule that governs deficits and hence government spending exhibits a countercyclical response to inflation and output deviations, -1.1 and -0.9, and features persistence as well. The tax rule that governs average taxation has similar properties. Average tax rates rise when output or debt is high, but tax level changes do not have a strong persistence beyond the persistence of output and debt.

Figure 1 plots the top 10% income and wealth shares in the US data and for all HANK variants based on the Kalman smoother. Strikingly, even when estimated only on aggregate data, the business cycle shocks and frictions lead the HANK model to reproduce the observed U-shaped evolution of inequality from 1954 to 2019. Since our baseline estimation (HANK) does not use any cross-sectional information, the good fit of the cross-sectional data can be informally thought of as passing an over-identification test. In turn, adding data on inequality has little effect on the estimated parameters (HANKX). In other words, the estimated

Figure 2: Variance decompositions: Output and consumption growth

Notes: Conditional variance decompositions at a 4-quarter forecast horizon for the estimated 1) RANK model, 2) HANK model without inequality data, and 3) the HANK model with inequality data (HANKX).

business cycle shocks and frictions in combination with incomplete markets do a good job of matching the evolution of wealth and income inequality over the last 65 years. Hence, the HANK model is able to match both inequality and business cycle dynamics without needing to compromise on one or the other. Adding retained earnings and fluctuations in income risk and tax progressivity as well as the corresponding data for the latter two (HANKX+) further increases the fit with the inequality data. We will explore the inequality results in detail in Section 5.

4.2 Variance decompositions

Next, we show what the estimated parameters imply for our view of US business cycles by looking at variance decompositions at business cycle frequency.[31] Figure 2 shows the decompositions of the growth rates of output and consumption.

As in the representative agent literature (see, e.g., Smets and Wouters, 2007; Justiniano et al., 2011), TFP and investment-specific technology shocks are the most important drivers of output growth, each explaining roughly 25% of its conditional variance at business cycle frequency. Compared to the RANK model, the HANK model puts less emphasis on

[31] We report the conditional variances at four quarters.

investment-specific technology shocks and more emphasis on shocks to the risk-premium, monetary policy, and markups. The HANK model implies additional cross-equation restrictions that matter in particular for the investment-specific technology shock.

A negative investment technology shock makes the production of investment goods less efficient, it discourages investment and thereby leads to an output decline. This is true in both HANK and RANK models. As the price of investment goods goes up after a negative investment technology shock, there is an additional effect in our HANK model: the price increase of capital has a positive wealth effect, which stabilizes consumption and therefore investment technology shocks contribute positively to consumption through the lens of the HANK model. This feature reduces the co-movement of investment and consumption in the short run; see also the impulse response functions to investment technology shocks in Appendix D.3. In models in which households do not directly hold capital this mechanism is muted.[32]

Looking at the decomposition of consumption growth, the HANK model again implies a larger role of risk-premium and monetary shocks at the expense of the importance of TFP shocks. The smaller role of TFP shocks reflects the fact that households under incomplete markets do not initially respond in their consumption one-for-one to a permanent increase in income; see Blundell et al. (2008). Since we estimate TFP shocks to be fairly persistent, this effect is important. The fact that we find monetary shocks to be more important for consumption is in line with Kaplan et al. (2018) and Luetticke (2020). In the next section, we revisit this question with income risk shocks and find that they partly replace risk-premium shocks.

Consistent with estimating very similar shocks and frictions, as discussed before, we find close to identical variance decompositions for the estimations across both HANK variants with and without inequality data (HANK vs. HANKX).

4.3 Historical decompositions

While the variance decompositions help us understand the average business cycle implied by the model, a historical decomposition tells us how the model views the actual cycles that the US economy has gone through. We summarize the historical decomposition of NBER-dated recessions in Table 5. We report historical decompositions for all observables in Appendix E. Again, we find that all model variants paint overall similar pictures of US recessions. Having said this, the difference in variance decompositions that we documented before for

[32]This partly explains the difference with Auclert et al. (2020), who find that investment-specific technology shocks are important in their HANK model with a mutual fund that holds all assets, instead of our two-asset structure that increases the role of the wealth effect of higher capital prices.

Table 5: Contribution of shocks to US recessions

	Output growth			Consumption growth		
Shocks	RANK	HANK	HANKX	RANK	HANK	HANKX
TFP, ϵ^Z	-0.19	-0.23	-0.27	-0.24	-0.16	-0.19
Inv.-spec. tech., ϵ^Ψ	-0.51	-0.30	-0.27	-0.19	0.07	0.10
Price markup, $\epsilon^{\mu Y}$	0.00	0.01	-0.02	-0.03	-0.07	-0.09
Wage markup, $\epsilon^{\mu W}$	-0.16	-0.26	-0.25	-0.31	-0.47	-0.45
Risk premium, ϵ^A	-0.24	-0.38	-0.35	-0.39	-0.48	-0.47
Monetary policy, ϵ^R	0.07	0.13	0.13	0.18	0.25	0.25
Structural deficit, ϵ^G	0.02	0.03	0.03	-0.03	-0.15	-0.15

Notes: The table displays the average contribution of the various shocks during an NBER-dated recession that result from our historical shock decomposition. Values are calculated by averaging the value of each shock component over all NBER recession quarters. To improve readability, we normalized the size of the overall contraction to -1%. In the data, the average is -1.24% for output and -0.5% for consumption.

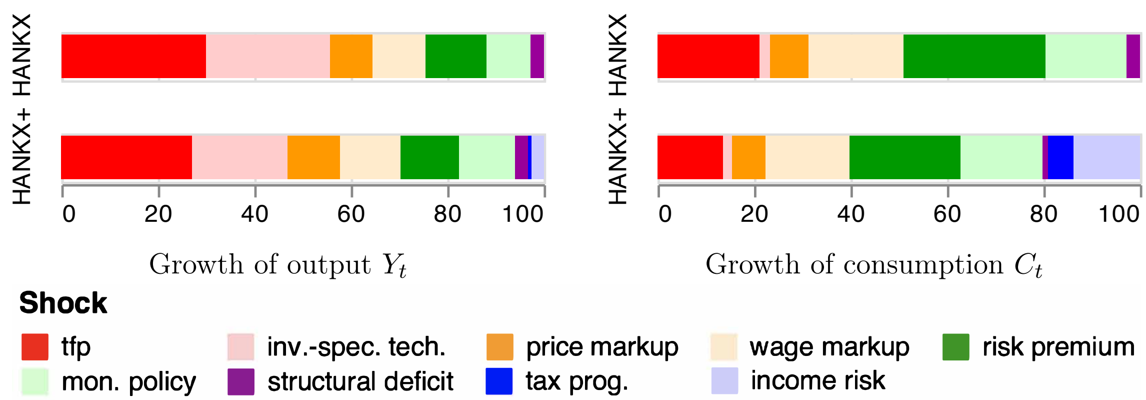
the average cycle are more pronounced in US recessions. The technology shocks are less important under the HANK model, in particular for consumption growth. By contrast, the risk-premium shock, as a prominent demand-side shock, but also wage markup shocks become more important through the lens of the HANK model for both consumption and output growth in US recessions.

4.4 Adding cross-sectional shocks

Importantly, the HANK model with two assets allows us to put more structure on the risk-premium shock. This shock captures any wedge between the return on liquid and illiquid assets that households require. In an incomplete markets model, this can be the result of income risk shocks (Bayer et al., 2019) or changes in tax progressivity. Both can be directly observed and, hence, we use them to estimate the HANKX+ variant of our model.

Figure 3 shows the variance decomposition for this estimation. The new shocks pick up a part of what is estimated as the more unspecific risk-premium shock in the baseline specification, so that its importance declines. Together, the shocks to income risk and tax progressivity explain one-fifth of consumption growth.

Figure 3: Variance decompositions: Aggregate vs. cross-sectional shocks



Notes: Conditional variance decompositions at a 4-quarter forecast horizon. HANKX corresponds to the estimated model with inequality data. HANKX+ corresponds to the estimated model with inequality data, shocks to income risk and tax progressivity (plus their observables), as well as retained earnings.

4.5 Robustness

Our estimation technique does not allow us to estimate parameters that change the stationary equilibrium (those in Table 2). Therefore, we consider two robustness checks for the HANK model. First, we ask whether our findings change when we assume a coefficient of relative risk aversion ξ, of 2 instead of 4. We find our results to be robust; in particular, the variance decompositions of output and consumption growth are almost unchanged: See Appendix F.1.

In a similar way, we also compare a variant of the model where we use King et al. (1988) (KPR) instead of Greenwood et al. (1988) (GHH) preferences. KPR preferences introduce a wealth effect in labor supply such that consumption-poor households supply more labor, and hence, aggregate labor supply increases when consumption falls. In line with the literature on estimated DSGE models that finds small aggregate wealth effects on the labor supply (see, e.g., Schmitt-Grohé and Uribe, 2012; Born and Pfeifer, 2014), we also find that the data clearly prefer the GHH specification over the KPR specification of the HANK model. The marginal data density for the model with KPR preferences is 6390 vs. 6467 for GHH preferences (based on the modified harmonic mean estimator of Geweke, 1999). Unsurprisingly, given the lower marginal data density, the drivers of US business cycles look somewhat different with KPR preferences. Compared to our baseline results, investment-specific technology shocks become more important for output growth and TFP shocks more important for consumption growth.

What is more, the KPR version of the HANK model produces counterfactual inequality dynamics and, hence, fails our informal "over-identification" test; see Figure F.27 in Appendix F.2. Vice versa, however, this shows that matching inequality dynamics alongside aggregate dynamics is not a simple implication of market incompleteness and the distribution of incomes that we assume.

5 US inequality

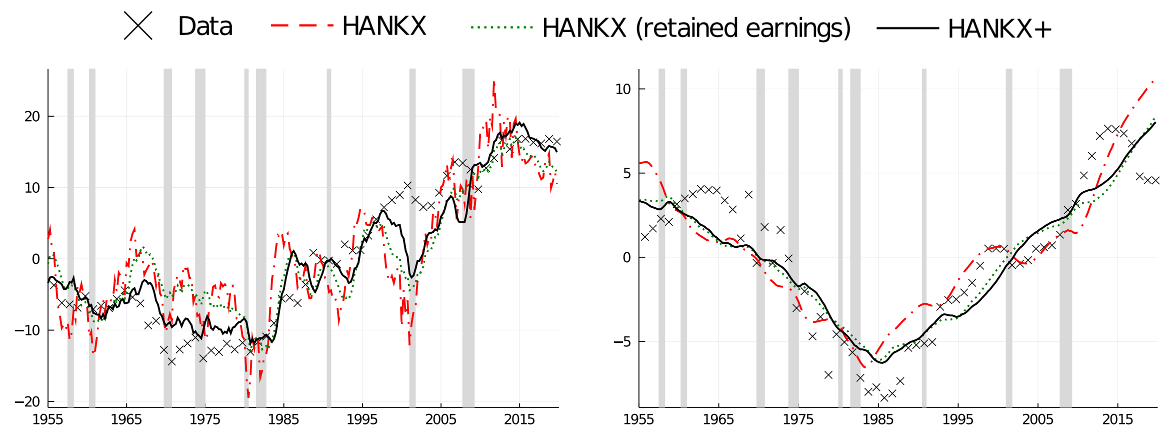
We have already seen in the previous section our HANK model can simultaneously match aggregate and inequality dynamics. In this section, we use the HANKX+ model to understand the propagation from aggregate and cross-sectional shocks to inequality, the drivers of US inequality through the lens of the model, and the importance of policy rules for the evolution of inequality.

The model closely replicates the increase in the top 10% share in income and wealth in the US data from 1985 to 2019, which increased by 22 and 12 percent, respectively. This holds true for all model variants; see Figure 4. However, the HANKX model with only aggregate shocks yields too much short-run movement in income inequality, which appears very smooth in the data. The HANKX+ model removes most of the short-run volatility through the introduction of retained earnings. To highlight this, we also estimate an intermediate variant where we only add retained earnings to the HANKX model (dotted line). With retained earnings, every quarter firms and unions pay out their profits in stationary equilibrium plus a fraction of the deviations thereof; see Footnote 19. The remainder is retained and paid out later. We estimate this pay-out fraction, $\omega^{U,F}$, to be around 13% quarterly; see Table 4 Column 8. This payout policy implies that 29% of the deviations in profits over the fiscal year are paid out to households over the course of a year.[33] Given that profits are not accounted for in the income-inequality data at the time they are accrued but at the time they are paid out to households, this smoothes the inequality response to markup fluctuations.

The HANKX+ model, in addition to retained earnings, features shocks to income risk and tax progressivity that directly affect the distribution of income. Both of these shocks are pinned down by two additional observables. We estimate the tax shocks to be very persistent, 0.971, and the income risk shocks to be less persistent, 0.681, but with a sizable variance of 69%. This is in line with the size of income risk variations reported in Storesletten et al. (2004) over the business cycle. We find that fluctuations in income risk are effectively

[33] This is broadly in line with estimates from US Compustat data. Kahle and Stulz (2020) find that payouts relative to operating income vary over time between 20% and 35%, see their Table 1. They estimate an elasticity of payouts to operating income of roughly 60%, see their Table 2, implying a payout of profit variations of 15 to 20 cents per dollar.

Figure 4: US inequality – data vs. extended model



(a) Top 10% income share

(b) Top 10% wealth share

Notes: Data (crosses) correspond to log-deviations of the annual observations of the share of pre-tax income and wealth held by the top 10% in each distribution in the US taken from the World Inequality Database. HANKX (dash-dotted) corresponds to the smoothed states of the estimated model with inequality data, HANKX (retained earnings) also has retained earnings (dotted), and HANKX+ (solid) with inequality data, shocks to income risk and tax progressivity (plus their observables), as well as retained earnings. Shaded areas correspond to NBER-dated recessions.

exogenous because the estimated feedback coefficient to the cycle, Σ_Y, is small compared to the size of fluctuations in income risk; see Figure E.22 in the appendix for the historical decomposition. The decrease in tax progressivity and the increase in income risk over time allows the model to match the long-run evolution of income inequality even better.

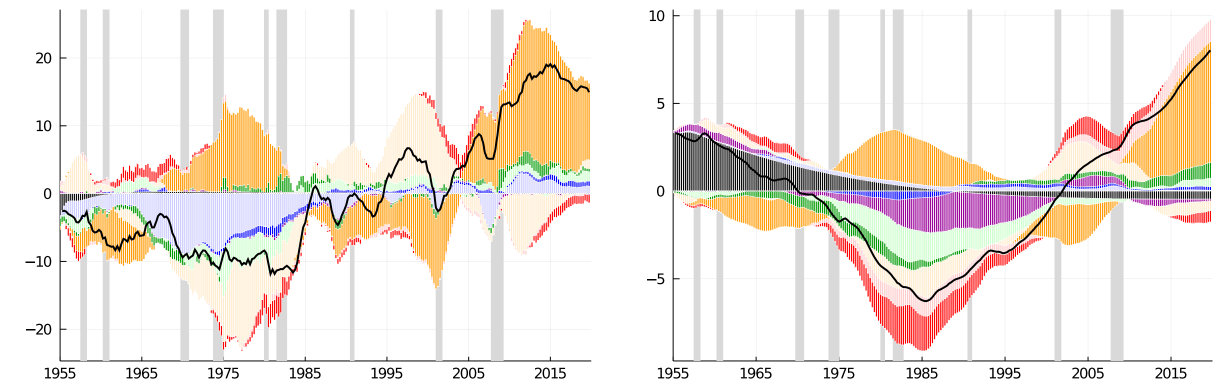
Transitory shocks lead to slow-moving inequality dynamics, because household wealth accumulates past transitory shocks to incomes and returns. In Appendix G, we report and discuss the IRFs of wealth and income inequality for all nine shocks. The business cycle shocks in our model redistribute income across households with different portfolios, and wealth (inequality) therefore partly keeps track of past shocks. To see this, consider a textbook permanent-income-hypothesis model. There, individual wealth follows a random walk in the presence of purely transitory shocks to income.

This matching of the distributional data, on top of the "standard" macroeconomic time series, does not change significantly what we infer about shocks and frictions, as we have seen in the previous section. The next section will provide a more detailed account of the driving forces behind US inequality.

5.1 Historical decompositions of US inequality

To dig into the details of the evolution of inequality, Figure 5(a) plots the historical decomposition of the top 10% income share. The decomposition of income inequality shows that medium-term trends of income inequality primarily result from markup shocks and fluctuations in income risk.

Figure 5: Historical decompositions of US inequality



(a) Top 10% income share (b) Top 10% wealth share

Notes: Historical decomposition of the estimated HANKX+ model. Left panel shows the log-deviations of the top 10% share in pre-tax income and the right panel of the top 10% wealth share. Shaded areas correspond to NBER-dated recessions.

With respect to particular historical episodes, our decomposition suggests the following. Rising wage markups and low idiosyncratic productivity risks are mainly responsible for the decrease in income inequality throughout the 1960s until the 1970s. The 1980s are seen as a period of liberalization through the lens of our model (both in terms of output cycles and inequality). Wage markups fell, which increased income inequality, but this was partly offset by falling price markups. This picture changes throughout the 1990s but most clearly after the dot-com recession. Through the lens of our model, it is larger income risks and sharply increasing price markups after 2000 that best explain aggregate fluctuations and the sharp rise in income inequality these years have witnessed. Interestingly and despite the use of partly different data sources and methodologies, the historical decomposition thus is in line with the evidence by De Loecker and Eeckhout (2020) on the evolution of markups in the

Table 6: Contribution of shocks to US inequality, 1985 to 2019

Shock	Top 10% Income	Top 10% Wealth
TFP, ϵ^Z	-0.79	1.42
Inv.-spec. tech., ϵ^Ψ	0.37	2.37
Price markup, $\epsilon^{\mu Y}$	12.03	4.32
Wage markup, $\epsilon^{\mu W}$	4.43	1.00
Risk premium, ϵ^A	-1.29	1.37
Income risk, ϵ^σ	3.02	-0.14
Monetary policy, ϵ^R	0.63	2.12
Structural deficit, ϵ^G	0.21	1.78
Tax progressivity, τ^P	1.44	0.58
Sum of shocks	19.91	14.18

Notes: The table displays the contribution (in log points) of each shock to the increase in the top 10% share of pre-tax income and the top 10% share of wealth from 1985 to 2019 based on our historical shock decompositions in the HANKX+ model.

US.[34] The sharp increase in price markups after 2000 leads our model to predict that the firm sector has been accumulating retained earnings ever since, in line with the data (see, e.g., Faulkender et al., 2019, Figure 1).

Figure 5(b) shows the historical decomposition of the top 10% wealth share. Wealth inequality fell in the first half of the sample and then increased. The pattern is similar in shape to that of income inequality, but smoother. Of course, income inequality partly translates into wealth inequality, but changes in asset prices play an independent and important role for wealth inequality (in line with Kuhn et al., 2020). Hence, technology shocks but also policy shocks are important to understanding the evolution of US wealth inequality. These shocks affect the expected and surprise return differences between liquid and illiquid assets (liquidity premium), which mimics the U-shape of wealth inequality, see Figure E.23 in the appendix. Since relatively poor households hold a larger fraction of their wealth in liquid assets, technology and policy shocks redistribute wealth directly. A low liquidity premium redistributes to wealth-poor households while a high premium favors wealthy households. In fact, Table 6 shows that the overall increase in wealth inequality since 1985 is explained

[34]There is a growing literature on the rise of markups; see, e.g., Karabarbounis and Neiman (2019), Barkai (2020), Hall (2018), or Kehrig and Vincent (2020).

roughly to an equal extent by three groups of shocks: the technology shocks, the policy shocks, and the markup shocks. This is in stark contrast to the increase in income inequality where the two markup shocks explain more than two-third of the increase.

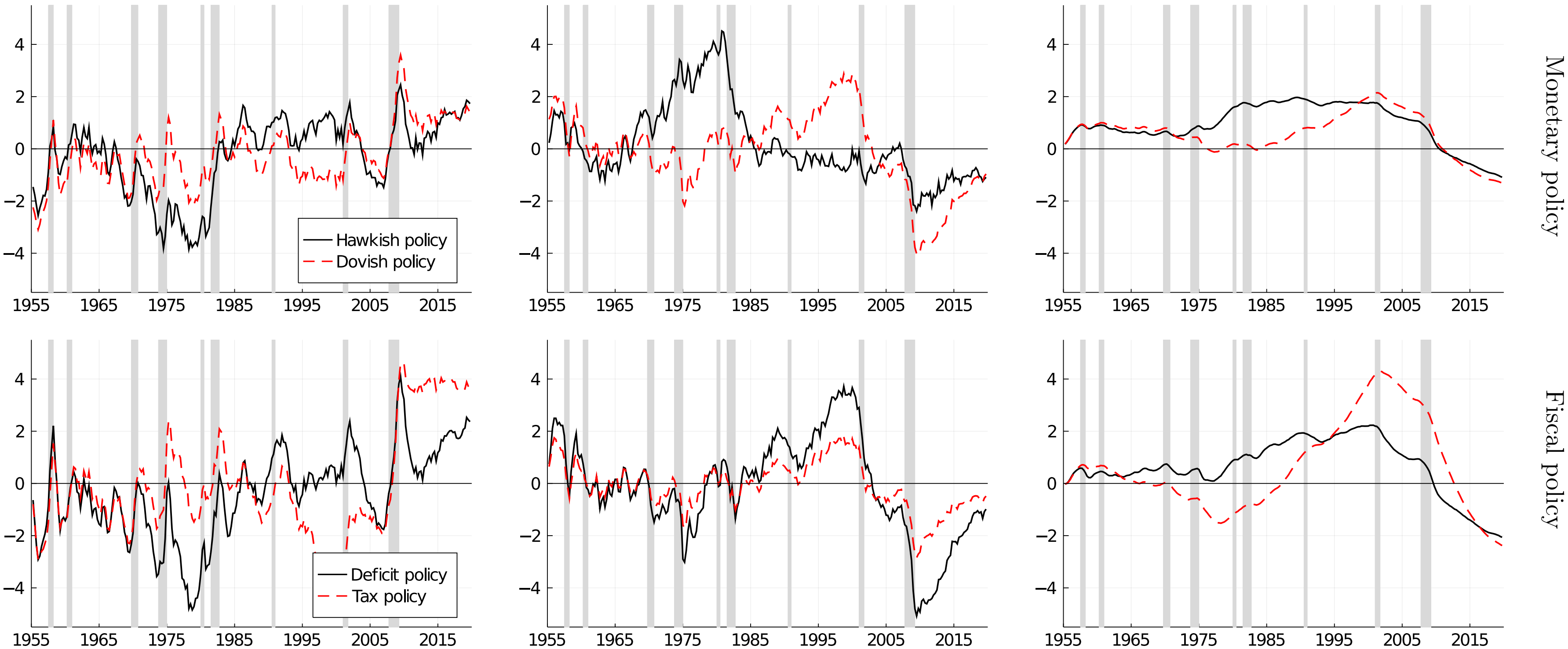
These findings complement the prevailing literature on the evolution of inequality that typically focuses on a few drivers of inequality at a time, such as the variance of income, changes in the tax and transfer system, or differential ex-post returns interacting with complex household portfolio structures (see, e.g., Hubmer et al., 2020; Kaymak and Poschke, 2016; Kuhn et al., 2020). We add to this literature by inspecting these three mechanisms jointly with the business cycle shocks, which we also find to have persistent effects on inequality. Focusing on government policy, we find that, among the three shocks, the decline in tax progressivity is the most important for income inequality, closely followed by monetary policy shocks. Structural deficits play little role in income inequality. However, for wealth inequality, structural deficits and monetary policy shocks are more important than tax progressivity. One reason for the muted effect on wealth inequality is that lower progressivity leads to a higher capital stock through its interaction with portfolio heterogeneity, because wealthy households have a higher propensity to invest in capital; see Luetticke (2020). This reduces dividends and increases wages, which mitigates the effect on inequality.

5.2 Inequality counterfactuals

After looking at policy shocks, we now evaluate the importance of policy rules for the evolution of US inequality. To understand the role of systematic business cycle policies in shaping inequality, we run a set of counterfactual monetary and fiscal policy experiments based on the estimated model. The results of these experiments are displayed in Figure 6. In detail, the figure displays the difference in the evolution of output, income inequality, and wealth inequality between running the estimated shock sequence through the model (HANKX+) with the estimated parameters and with counterfactually set policy parameters.

First, we consider an experiment where the Fed reacts more aggressively to inflation (doubling θ_π); see Figure 6, top panel, black solid line. This creates large output losses after markup shocks because markups adjust toward their targets more quickly, but the same policy stabilizes output very effectively after demand shocks. Given the series of shocks, output would have been lower in the 1970s. This reflects the fact that our model attributes a substantial fraction of the fluctuations of the 1970s to markup (cost-push) shocks. The sharper increase of profit margins is mirrored in higher income inequality in the 1970s. Even though income inequality returns to baseline by the mid 1980s, wealth inequality remains elevated until the early 2000s.

Figure 6: Counterfactual monetary and fiscal policy



(a) Output (b) Top 10% income share (c) Top 10% wealth share

Notes: The panels display the evolution of output, income inequality, and wealth inequality that the model would counterfactually predict had the government policies been different, feeding the same smoothed sequence of shocks (as in Figure 5) through the model. The lines represent the difference (in log points) in the evolution compared to the baseline model. Monetary policy: The solid line corresponds to a setup where we double the inflation response, θ_π. The dashed line reflects the counterfactual where we double the estimated response to output, θ_Y. Fiscal Policy: The solid line corresponds to a setup where we allow for more aggressive deficits, doubling γ_Y and halving γ_B^τ. The dashed line reflects the counterfactual where we double the estimated tax response to output, γ_Y^τ. Shaded areas correspond to NBER-dated recessions.

Second, we consider a dovish policy where we double the monetary policy response to output fluctuations (doubling θ_Y); see Figure 6, top panel, red dashed line. This leads in general to more stable markups and output at the expense of higher inflation volatility; see also Gornemann et al. (2012). It is not fully the mirror image of the hawkish policy we looked at before because this experiment changes the response to fluctuations in output, not inflation. A more dovish policy reduces income inequality in the 1970s and after the Great Recession, while it increases income inequality in the 1990s, where markups were falling. This again translates into a delayed and persistent increase in wealth inequality.

Finally, we consider alternative fiscal policy scenarios; see Figure 6, bottom panel. First, we assume more aggressive deficit policies in response to output (doubling γ_Y and halving γ_B^τ, black solid line). Second, we consider a policy that adjusts taxes more heavily in response to output (doubling γ_Y^τ, red dashed line). That is, we consider a policy that lowers taxes rather than raises government consumption when fighting a recession with a government deficit.

More aggressive stabilization through average tax rates can effectively stabilize the economy, e.g., output is higher in recessions. At the same time, it reduces income inequality in recessions but comes at the cost of higher inequality in booms. Wealth inequality mirrors income inequality with a delay but is more persistent. More aggressive stabilization through fiscal deficits is on average somewhat less effective in stabilizing output compared to tax cuts. What is more, an aggressive deficit policy makes income inequality more volatile relative to stabilization via taxes. Despite this fact, the aggressive deficit policy stabilizes wealth inequality more effectively through emitting more government debt and thereby shrinking the liquidity premium. Bayer et al. (2020) provides a more detailed discussion of the response of the liquidity premium to government debt.

6 Conclusion

How much does inequality matter for the business cycle and vice versa? To shed light on this two-way relationship, this paper estimates a state-of-the-art New-Keynesian business cycle model with household heterogeneity and portfolio choice on macro and micro data. We find that a two-asset incomplete markets model can explain simultaneously the evolution of the US business cycle and US inequality.

Relative to the representative agent literature, our view of the US business cycle puts more emphasis on household consumption-savings and portfolio decisions and thereby strengthens the importance of the demand side of the model. We add to the debate on inequality the importance of business cycle shocks and policies. Our analysis suggests that price markups and the excess return of capital have substantially increased over the last two decades. This has driven down output and has increased income and wealth inequality. We find that business cycle policies have a long-lasting effect on inequality.

These findings suggest that future research on inequality should take business cycles into account, and the study of optimal business cycle policy should take inequality into account. Our findings further suggest exploring the role of shocks that affect household insurance for the business cycle. Including a micro-foundation for income risk, e.g., via search and matching, is of first-order importance to understanding how the business cycle and policies work differently by affecting income risk itself.

References

Ahn, S., Kaplan, G., Moll, B., Winberry, T., and Wolf, C. (2018). When inequality matters for macro and macro matters for inequality. *NBER Macroeconomics Annual*, 32(1):1–75.

An, S. and Schorfheide, F. (2007). Bayesian analysis of DSGE models. *Econometric Reviews*, 26(2-4):113–172.

Auclert, A. (2019). Monetary policy and the redistribution channel. *American Economic Review*, 109(6):2333–67.

Auclert, A., Bardóczy, B., Rognlie, M., and Straub, L. (2019). Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models. NBER Working Paper 26123.

Auclert, A., Rognlie, M., and Straub, L. (2018). The intertemporal Keynesian Cross. NBER Working Paper 25020.

Auclert, A., Rognlie, M., and Straub, L. (2020). Micro jumps, macro humps: Monetary policy and business cycles in an estimated HANK model. NBER Working Paper 26647.

Barkai, S. (2020). Declining labor and capital shares. *Journal of Finance*, forthcoming.

Bayer, C., Born, B., and Luetticke, R. (2020). The liquidity channel of fiscal policy. CEPR Discussion Paper 14883.

Bayer, C. and Luetticke, R. (2020). Solving heterogeneous agent models in discrete time with many idiosyncratic states by perturbation methods. *Quantitative Economics*, 11:1253–1288.

Bayer, C., Luetticke, R., Pham-Dao, L., and Tjaden, V. (2019). Precautionary savings, illiquid assets, and the aggregate consequences of shocks to household income risk. *Econometrica*, 87(1):255–290.

Benabou, R. (2002). Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, 70(2):481–517.

Berger, D., Bocola, L., and Dovis, A. (2020). Imperfect risk sharing and the business cycle. Technical report, Stanford University.

Blundell, R., Pistaferri, L., and Preston, I. (2008). Consumption inequality and partial insurance. *American Economic Review*, 98(5):1887–1921.

Boar, C. and Midrigan, V. (2019). Markups and inequality. NBER Working Paper 25952.

Born, B. and Pfeifer, J. (2014). Policy risk and the business cycle. *Journal of Monetary Economics*, 68:68–85.

Bricker, J., Henriques, A., Krimmel, J., and Sabelhaus, J. (2016). Measuring income and wealth at the top using administrative and survey data. *Brookings Papers on Economic Activity*, 2016(1):261–331.

Broer, T., Hansen, N.-J. H., Krusell, P., and Öberg, E. (2019). The New Keynesian transmission mechanism: A heterogenous-agent perspective. *Review of Economic Studies*, 87(1):77–101.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12(3):383–398.

Castaneda, A., Diaz-Gimenez, J., and Rios-Rull, J.-V. (1998). Exploring the income distribution business cycle dynamics. *Journal of Monetary Economics*, 42(1):93 – 130.

Challe, E. and Ragot, X. (2015). Precautionary saving over the business cycle. *Economic Journal*, 126(590):135–164.

Chang, M., Chen, X., and Schorfheide, F. (2018). Heterogeneity and aggregate fluctuations. Mimeo, University of Pennsylvania.

Chetty, R., Guren, A., Manoli, D., and Weber, A. (2011). Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins. *American Economic Review*, 101(3):471–475.

De Loecker, J. and Eeckhout, J. (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics (forthcoming)*.

Den Haan, W. J., Rendahl, P., and Riegler, M. (2017). Unemployment (fears) and deflationary spirals. *Journal of the European Economic Association*, 16(5):1281–1349.

Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford University Press.

Faulkender, M. W., Hankins, K. W., and Petersen, M. A. (2019). Understanding the rise in corporate cash: Precautionary savings or foreign taxes. *Review of Financial Studies*, 32(9):3299–3334.

Fernández-Villaverde, J. (2010). The econometrics of DSGE models. *SERIEs*, 1(1):3–49.

Ferriere, A. and Navarro, G. (2018). The heterogeneous effects of government spending: It's all about taxes. *FRB International Finance Discussion Paper*, (1237).

Gabaix, X., Lasry, J.-M., Lions, P.-L., and Moll, B. (2016). The dynamics of inequality. *Econometrica*, 84(6):2071–2111.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., editors, *Bayesian Statistics*, volume 4, pages 641–649. Clarendon Press, Oxford.

Geweke, J. (1999). Using simulation methods for Bayesian econometric models: Inference, development, and communication. *Econometric Reviews*, 18(1):1–73.

Gornemann, N., Kuester, K., and Nakajima, M. (2012). Monetary policy with heterogeneous agents. *Philadelphia Fed Working Paper 12-21*.

Greenwood, J., Hercowitz, Z., and Huffman, G. W. (1988). Investment, capacity utilization, and the real business cycle. *American Economic Review*, 78(3):402–417.

Guerrieri, V. and Lorenzoni, G. (2017). Credit crises, precautionary savings, and the liquidity trap. *Quarterly Journal of Economics*, 132(3):1427–1467.

Guvenen, F., Kaplan, G., and Song, J. (2014). How risky are recessions for top earners? *The American Economic Review*, 104(5):148–153.

Hagedorn, M., Manovskii, I., and Mitman, K. (2018). Monetary policy in incomplete market models: Theory and evidence. Mimeo, University of Pennsylvania.

Hagedorn, M., Manovskii, I., and Mitman, K. (2019). The fiscal multiplier. NBER Working Paper 25571.

Hall, R. E. (2018). New evidence on the markup of prices over marginal costs and the role of mega-firms in the US economy. NBER Working Paper 24574.

Heathcote, J., Perri, F., and Violante, G. L. (2010). Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006. *Review of Economic Dynamics*, 13(1):15–51.

Heathcote, J., Storesletten, K., and Violante, G. L. (2017). Optimal tax progressivity: An analytical framework. *Quarterly Journal of Economics*, 132(4):1693–1754.

Hubmer, J., Krusell, P., and Smith Jr, A. A. (2020). Sources of U.S. wealth inequality: Past, present, and future. *NBER Macroeconomics Annual*, 35.

Jaimovich, N. and Rebelo, S. (2009). Can news about the future drive the business cycle? *American Economic Review*, 99(4):1097–1118.

Justiniano, A., Primiceri, G. E., and Tambalotti, A. (2011). Investment shocks and the relative price of investment. *Review of Economic Dynamics*, 14(1):101–121.

Kahle, K. and Stulz, R. M. (2020). Why are corporate payouts so high in the 2000s? NBER Working Paper 26958.

Kaplan, G., Moll, B., and Violante, G. L. (2018). Monetary policy according to HANK. *American Economic Review*, 108(3):697–743.

Kaplan, G. and Violante, G. L. (2014). A model of the consumption response to fiscal stimulus payments. *Econometrica*, 82:1199–1239.

Karabarbounis, L. and Neiman, B. (2019). Accounting for factorless income. *NBER Macroeconomics Annual*, 33(1):167–228.

Kaymak, B. and Poschke, M. (2016). The evolution of wealth inequality over half a century: The role of taxes, transfers and technology. *Journal of Monetary Economics*, 77:1–25.

Kehrig, M. and Vincent, N. (2020). The micro-level anatomy of the labor share decline. *Quarterly Journal of Economics*, forthcoming.

King, R. G., Plosser, C. I., and Rebelo, S. T. (1988). Production, growth and business cycles: I. The basic neoclassical model. *Journal of Monetary Economics*, 21(2):195 – 232.

Klein, P. (2000). Using the generalized Schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control*, 24(10):1405–1423.

Koopman, S. J. and Durbin, J. (2000). Fast filtering and smoothing for multivariate state space models. *Journal of Time Series Analysis*, 21(3):281–296.

Kopczuk, W. (2015). What do we know about the evolution of top wealth shares in the United States? *Journal of Economic Perspectives*, 29(1):47–66.

Kuhn, M., Schularick, M., and Steins, U. I. (2020). Income and wealth inequality in America, 1949-2013. *Journal of Political Economy*, 128(9):3469–3519.

Liu, L. and Plagborg-Møller, M. (2019). Full-information estimation of heterogeneous agent models using macro and micro data. Mimeo, Princeton University.

Luetticke, R. (2020). Transmission of monetary policy with heterogeneity in household portfolios. *American Economic Journal: Macroeconomics*, forthcoming.

McKay, A., Nakamura, E., and Steinsson, J. (2016). The power of forward guidance revisited. *American Economic Review*, 106(10):3133–3158.

McKay, A. and Reis, R. (2016). The role of automatic stabilizers in the US business cycle. *Econometrica*, 84(1):141–194.

McKay, A. and Reis, R. (2020). Optimal automatic stabilizers. *Review of Economic Studies*, forthcoming.

Mertens, K. and Montiel Olea, J. L. (2018). Marginal tax rates and income: New time series evidence. *Quarterly Journal of Economics*, 133(4):1803–1884.

Piketty, T. and Saez, E. (2003). Income inequality in the United States, 1913–1998. *Quarterly Journal of Economics*, 118(1):1–41.

Ravn, M. O. and Sterk, V. (2017). Job uncertainty and deep recessions. *Journal of Monetary Economics*, 90:125–141.

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

Saez, E. (2004). Reported incomes and marginal tax rates, 1960-2000: evidence and policy implications. *Tax policy and the economy*, 18:117–173.

Saez, E. and Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, 131(2):519–578.

Schmitt-Grohé, S. and Uribe, M. (2004). Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control*, 28(4):755–775.

Schmitt-Grohé, S. and Uribe, M. (2012). What's news in business cycles. *Econometrica*, 80(6):2733–2764.

Smets, F. and Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review*, 97(3):586–606.

Sterk, V. and Tenreyro, S. (2018). The transmission of monetary policy through redistributions and durable purchases. *Journal of Monetary Economics*, 99:124–137.

Storesletten, K., Telmer, C. I., and Yaron, A. (2004). Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy*, 112(3):695–717.

Taylor, J. B. (1993). Discretion versus policy rules in practice. In *Carnegie-Rochester Conference Series on Public Policy*, volume 39, pages 195–214. Elsevier.

Wong, A. (2019). Refinancing and the transmission of monetary policy to consumption. Mimeo, Princeton University.

Woodford, M. (1995). Price-level determinacy without control of a monetary aggregate. *Carnegie-Rochester Conference Series on Public Policy*, 43(Supplement C):1 – 46.

Wu, J. C. and Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2-3):253–291.

A Data

A.1 Data for calibration

Mean illiquid assets. Fixed assets (NIPA table 1.1) over quarterly GDP (excluding net exports; see below), averaged over 1954 – 2019.

Mean liquidity. Gross federal debt held by the public as percent of GDP (FYPUGDA188S). Available from 1954 – 2019.

Fraction of borrowers. Taken from the Survey of Consumer Finances (1983 – 2013); see Bayer et al. (2019) for more details.

Average top 10% share of wealth. Source is the World Inequality Database (1954 – 2019).

A.2 Data for estimation

Formally, the vector of observable variables is given by:

$$OBS_t = \begin{bmatrix} \Delta \log(Y_t) \\ \Delta \log(C_t) \\ \Delta \log(I_t) \\ \Delta \log\left(w_t^F\right) \\ \log(N_t) \\ \log\left(R_t^b\right) \\ \log(\pi_t) \\ \log\left(p90p100_t^{wealth}\right) \\ \log\left(p90p100_t^{income}\right) \\ \log(s_t) \\ \log\left(\tau_t^P\right) \end{bmatrix} - \begin{bmatrix} \overline{\Delta \log(Y_t)} \\ \overline{\Delta \log(C_t)} \\ \overline{\Delta \log(I_t)} \\ \overline{\Delta \log\left(w_t^F\right)} \\ \overline{\log(N_t)} \\ \overline{\log(\pi_t)} \\ \overline{\log\left(R_t^b\right)} \\ \overline{\log\left(p90p100_t^{wealth}\right)} \\ \overline{\log\left(p90p100_t^{income}\right)} \\ \overline{\log(s_t)} \\ \overline{\log\left(\tau_t^P\right)} \end{bmatrix}$$

where Δ denotes the temporal difference operator and bars above variables denote time-series averages.

Unless otherwise noted, all series available at quarterly frequency from 1954Q3 to 2019Q4 from the St.Louis FED - FRED database (mnemonics in parentheses).

Output, Y_t. Sum of gross private domestic investment (GPDI), personal consumption expenditures for nondurable goods (PCND), durable goods (PCDG), and services (PCESV), and government consumption expenditures and gross investment (GCE) divided by the GDP deflator (GDPDEF) and the civilian noninstitutional population (CNP16OV).

Consumption, C_t. Sum of personal consumption expenditures for nondurable goods

(PCND), durable goods (PCDG), and services (PCESV) divided by the GDP deflator (GDPDEF) and the civilian noninstitutional population (CNP16OV).

Investment, I_t. Gross private domestic investment (GPDI) divided by the GDP deflator (GDPDEF) and the civilian noninstitutional population (CNP16OV).

Real wage, w_t^F. Hourly compensation in the nonfarm business sector (COMPNFB) divided by the GDP deflator (GDPDEF).

Hours worked, N_t. Nonfarm business hours worked (COMPNFB) divided by the civilian noninstitutional population (CNP16OV).

Inflation, π_t. Computed as the log-difference of the GDP deflator (GDPDEF).

Nominal interest rate, R_t^b. Quarterly average of the effective federal funds rate (FEDFUNDS). From 2009Q1 to 2015Q4, we use the Wu and Xia (2016) shadow federal funds rate.

Wealth inequality, $p90p100_t^{wealth}$. p90p100 of US net personal wealth from the World Inequality Database. Available annually 1954 to 2019.

Income inequality, $p90p100_t^{income}$. p90p100 of US pre-tax national income from the World Inequality Database. Available annually 1954 to 2019.

Idiosyncratic income risk, s_t. We take the estimated time series for the variance of idiosyncratic income from Bayer et al. (2019) who use the Survey of Income and Program Participation. Available from 1983Q1 to 2013Q1.

Tax progressivity, τ_t^P. We follow Ferriere and Navarro (2018) and construct our measure of tax progressivity using the average and average marginal tax rate: P = (AMTR - ATR)/(1 - ATR). For a loglinear tax system, this measure equals the parameter capturing the curvature of the tax function. Available annually 1954 to 2017.

Details on the construction of the tax-progressivity measure

We extend the Mertens and Montiel Olea (2018)-calculations of average (ATR) and average marginal tax rates (AMTR) to the years 2013-2017. First, in constructing the ATR series, we obtain total tax liability for 1929-2017, from the National Income and Product Accounts (NIPA), *Table 3.2*. Federal social insurance contributions, which are added to total tax liability, come from NIPA, *Table 3.6*, line 3 and 21. For total income, we take Piketty and Saez (2003)'s income series, which uses a broader income concept based on adjusted gross income, excluding taxable social security and unemployment insurance benefits.

The AMTR is the sum of the average marginal individual income tax rate (AMIITR) and the average marginal payroll tax rate (AMPRT). We follow Ferriere and Navarro (2018) and use Saez (2004)'s income concept.[35] This income concept includes all income items reported

[35] For a detailed explanation on the construction of the AMTRs, see Appendix A of Mertens and Mon-

on an individual's tax return before deductions and excluding capital gains. Income items include salaries and wages, small business/farm income, partnership and fiduciary income, dividends, interest, rents, royalties and other small items reported as other income. Realized capital gains are excluded in this measure of income.

To construct the AMTR, we first use several tables from the Statistics of Income (SOI) to construct the discrete distributions of adjusted gross income by income brackets needed for the AMIITR. *Table 1.1 All Returns* of the SOI archives contains information on number of returns, adjusted gross income (AGI), and taxable income for different ranges of AGI per return. These ranges define the discretization. Given the distribution is fit for every year and by filing status, *Table 1.2 All Returns: by Marital Status* provides the equivalent table distinguishing by filing status, e.g., married filing jointly or separately, head of household, single, and surviving spouse. *Table 1.3 All Returns: Sources of Income* provides information on how many of these returns reported income from salaries and wages. *Table 1.4 All Returns: Sources of Income, Adjustments, and Tax Items* contains data on taxable income and number of corresponding returns by bracket. *Table 3.3 All Returns: Tax Liability, Tax Credits, and Tax Payments* provides information on how many filed for self-employment and their tax liability. Finally, *Table 3.4* contains the number of returns and adjusted gross income by marginal tax bracket and filing status using.

To construct the Average Marginal Payroll Tax Rate (AMPTR), we collect data from the 2019 Annual Statistical supplement, *Table 2.A3* (columns 1, 2, 3 and 9), to obtain the taxation of labor and self-employed earnings under the Old Age, Survivors and Disability Insurance (OASDI) and Hospital Insurance (HI) programs. The columns respectively cover the number of covered workers and self employed with maximum earnings as well as total taxable earnings. Their difference allows us to calculate the total taxable earnings of covered workers with earnings *below* the maximum. Information on earnings can be found in *Table 4.B* from the same source.

tiel Olea (2018). We follow *method 1* for computing the AMIITRs.

B RANK calibration

Table B.1 shows the steady-state parameterization of the representative agent analogue of the HANK model. We adjust the discount factor to match a capital-to-output ratio of 11.44 (quarterly) and the level of the tax rate to match the ratio of government-spending-to-output (0.2). All other parameters are externally chosen and equal to the parameterizaton of the HANK model.

Table B.1: External/calibrated parameters in RANK (quarterly frequency)

Parameter	Value	Description	Target
Households			
β	0.996	Discount factor	K/Y=11.44
ξ	4.00	Relative risk aversion	Kaplan et al. (2018)
γ	2.00	Inverse of Frisch elasticity	Chetty et al. (2011)
Firms			
α	0.68	Share of labor	62% labor income
δ_0	1.75%	Depreciation rate	7.0% p.a.
$\bar{\eta}$	11	Elasticity of substitution	Price markup 10%
$\bar{\zeta}$	11	Elasticity of substitution	Wage markup 10%
Government			
$\bar{\tau}^L$	0.23	Tax rate level	$G/Y = 20\%$
$\bar{\tau}^P$	0.12	Tax progressivity	
$\bar{R}^b$	1.00	Nominal rate	
$\bar{\pi}$	1.00	Inflation	

C MCMC diagnostics

We estimate each model using five parallel RWMH chains starting from an over-dispersed target distribution after an extensive mode search. After burn-in, 250,000 draws from the posterior distribution are used to compute the posterior statistics. The acceptance rates across chains are between 20% and 30%. Here, we provide Gelman and Rubin (1992) and Geweke (1992) convergence statistics as well as traceplots of individual parameters of the RANK, HANK, HANKX, and HANKX+ models. The Gelman and Rubin (1992) approach is based on comparing the estimated between-chains and within-chain variances for each model parameter. Large differences between these variances indicate non-convergence. Table C.2 reports the Gelman and Rubin potential scale reduction factor (PSRF) and its 97.5% quantile based on five chains. A common rule-of-thumb declares convergence if $PSRF < 1.1$. Geweke (1992) tests the equality of means of the first 10% of draws and the last 50% of draws (after burn-in). If the samples are drawn from the stationary distribution of the chain, the two means are equal and Geweke's statistic has an asymptotically standard normal distribution. Table C.3 reports the Geweke z-score statistic and the p-value for the pooled chains of each parameter. Taking the evidence from Geweke (1992), Gelman and Rubin (1992), and traceplot graphs together, we conclude that our chains have converged.

Table C.2: Gelman and Rubin (1992) convergence diagnostics

	RANK		HANK		HANKX		HANKX+	
Parameter	PSRF	97.5%	PSRF	97.5%	PSRF	97.5%	PSRF	97.5%
δ_s	1.001	1.003	1.001	1.003	1.002	1.004	1.001	1.002
ϕ	1.002	1.005	1.002	1.005	1.002	1.005	1.001	1.001
κ	1.002	1.006	1.002	1.006	1.001	1.004	1.004	1.010
κ_w	1.002	1.004	1.002	1.004	1.002	1.005	1.003	1.006
ρ_R	1.003	1.008	1.003	1.008	1.003	1.006	1.001	1.002
σ_R	1.003	1.009	1.003	1.009	1.002	1.003	1.002	1.005
θ_π	1.001	1.003	1.001	1.003	1.002	1.005	1.001	1.004
θ_Y	1.003	1.007	1.003	1.007	1.001	1.002	1.001	1.003
γ_B	1.007	1.016	1.007	1.016	1.008	1.021	1.002	1.005
γ_π	1.001	1.002	1.001	1.002	1.004	1.011	1.002	1.004
γ_Y	1.001	1.002	1.001	1.002	1.011	1.028	1.002	1.005
ρ_D	1.007	1.011	1.007	1.011	1.002	1.004	1.003	1.007
σ_D	1.002	1.004	1.002	1.004	1.002	1.004	1.002	1.006
ρ_τ	1.004	1.011	1.004	1.011	1.008	1.022	1.001	1.001
γ_B^τ	1.009	1.023	1.009	1.023	1.011	1.028	1.003	1.008
γ_Y^τ	1.001	1.001	1.001	1.001	1.009	1.020	1.003	1.007
ρ_P	–	–	–	–	–	–	1.002	1.004
σ_P	–	–	–	–	–	–	1.001	1.003
ρ_A	1.003	1.006	1.003	1.006	1.003	1.006	1.006	1.015
σ_A	1.002	1.004	1.002	1.004	1.001	1.001	1.004	1.010
ρ_Z	1.002	1.005	1.002	1.005	1.001	1.003	1.005	1.010
σ_Z	1.001	1.002	1.001	1.002	1.001	1.003	1.004	1.011
ρ_Ψ	1.001	1.002	1.001	1.002	1.001	1.002	1.002	1.004
σ_Ψ	1.001	1.002	1.001	1.002	1.001	1.002	1.001	1.002
ρ_μ	1.000	1.001	1.000	1.001	1.001	1.001	1.001	1.003
σ_μ	1.001	1.003	1.001	1.003	1.002	1.004	1.004	1.012
$\rho_{\mu w}$	1.001	1.003	1.001	1.003	1.003	1.006	1.002	1.003
$\sigma_{\mu w}$	1.001	1.004	1.001	1.004	1.002	1.002	1.002	1.004
ρ_s	–	–	–	–	–	–	1.004	1.011
σ_s	–	–	–	–	–	–	1.000	1.001
Σ_Y	–	–	–	–	–	–	1.003	1.008
ω^F	–	–	–	–	–	–	1.005	1.012
ω^U	–	–	–	–	–	–	1.004	1.007
σ_{W10}^{me}	–	–	–	–	1.001	1.003	1.001	1.002
σ_{I10}^{me}	–	–	–	–	1.002	1.003	1.008	1.021
$\sigma_{\tau P}^{me}$	–	–	–	–	–	–	1.102	1.233
σ_s^{me}	–	–	–	–	–	–	1.007	1.016

Note: Gelman and Rubin (1992) potential scale reduction factor (PSRF) and 97.5% quantile. A common rule-of-thumb declares convergence if $PSRF < 1.1$.

Table C.3: Geweke (1992) convergence diagnostics

	RANK		HANK		HANKX		HANKX+	
Parameter	z-score	p-value	z-score	p-value	z-score	p-value	z-score	p-value
δ_s	0.723	0.470	-1.069	0.285	0.757	0.449	0.755	0.450
ϕ	-0.688	0.491	-0.952	0.341	-0.154	0.878	1.414	0.157
κ	-0.853	0.394	0.472	0.637	0.552	0.581	-0.722	0.470
κ_w	0.659	0.510	0.280	0.779	2.227	0.026	-1.639	0.101
ρ_R	-0.877	0.381	0.202	0.840	-0.54	0.589	0.141	0.888
σ_R	-0.309	0.757	-2.224	0.026	1.399	0.162	-0.19	0.850
θ_π	-0.322	0.747	-0.323	0.746	0.408	0.683	0.560	0.576
θ_Y	-0.602	0.547	-1.765	0.078	1.027	0.304	1.019	0.308
γ_B	-0.349	0.727	-0.793	0.428	0.723	0.470	-1.40	0.162
γ_π	-0.687	0.492	0.025	0.980	0.241	0.809	-0.071	0.943
γ_Y	-0.277	0.782	-0.057	0.954	-1.489	0.137	0.265	0.791
ρ_G	-1.289	0.197	-1.079	0.280	0.493	0.622	0.966	0.334
σ_G	1.447	0.148	0.800	0.424	0.055	0.957	-1.378	0.168
ρ_τ	0.918	0.359	0.197	0.844	-0.475	0.635	0.252	0.801
γ_B^τ	-0.354	0.723	-0.68	0.496	0.968	0.333	-1.215	0.225
γ_Y^τ	0.337	0.736	0.585	0.558	1.063	0.288	0.145	0.885
ρ_P	–	–	–	–	–	–	1.498	0.134
σ_P	–	–	–	–	–	–	-2.072	0.038
ρ_A	1.075	0.282	0.078	0.938	-1.172	0.241	-0.503	0.615
σ_A	-1.106	0.269	-0.085	0.933	1.314	0.189	1.588	0.112
ρ_Z	1.976	0.048	-0.262	0.793	-0.644	0.520	0.657	0.511
σ_Z	0.503	0.615	-0.332	0.740	0.165	0.869	-0.474	0.635
ρ_Ψ	0.713	0.476	0.014	0.989	-1.005	0.315	0.390	0.697
σ_Ψ	-0.279	0.780	-0.827	0.408	-0.283	0.777	1.922	0.055
ρ_μ	0.379	0.705	-0.644	0.520	-0.375	0.708	0.573	0.567
σ_μ	1.252	0.210	-0.504	0.614	0.474	0.635	-0.548	0.584
$\rho_{\mu w}$	0.984	0.325	0.217	0.828	0.371	0.710	-0.581	0.561
$\sigma_{\mu w}$	-0.385	0.700	0.736	0.462	-1.179	0.238	0.718	0.473
ρ_s	–	–	–	–	–	–	-1.638	0.101
σ_s	–	–	–	–	–	–	1.550	0.121
Σ_Y	–	–	–	–	–	–	-0.551	0.581
ω^F	–	–	–	–	–	–	-0.325	0.745
ω^U	–	–	–	–	–	–	-0.351	0.725
σ_{W10}^{me}	–	–	–	–	1.518	0.129	0.278	0.781
σ_{I10}^{me}	–	–	–	–	1.075	0.282	1.296	0.195
$\sigma_{\tau P}^{me}$	–	–	–	–	–	–	0.541	0.588
σ_s^{me}	–	–	–	–	–	–	1.254	0.210

Note: Geweke (1992) equality of means test of the first 10% vs. the last 50% of draws. Failure to reject the null of equal means indicates convergence.

Figure C.1: MCMC draws of RANK model

Figure C.2: MCMC draws of RANK model

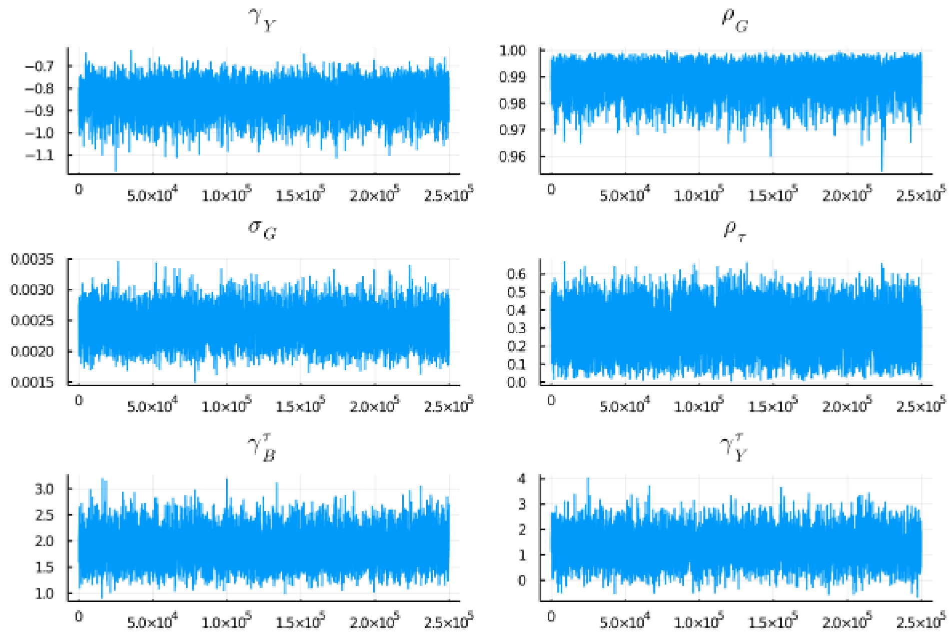


Figure C.3: MCMC draws of HANK model

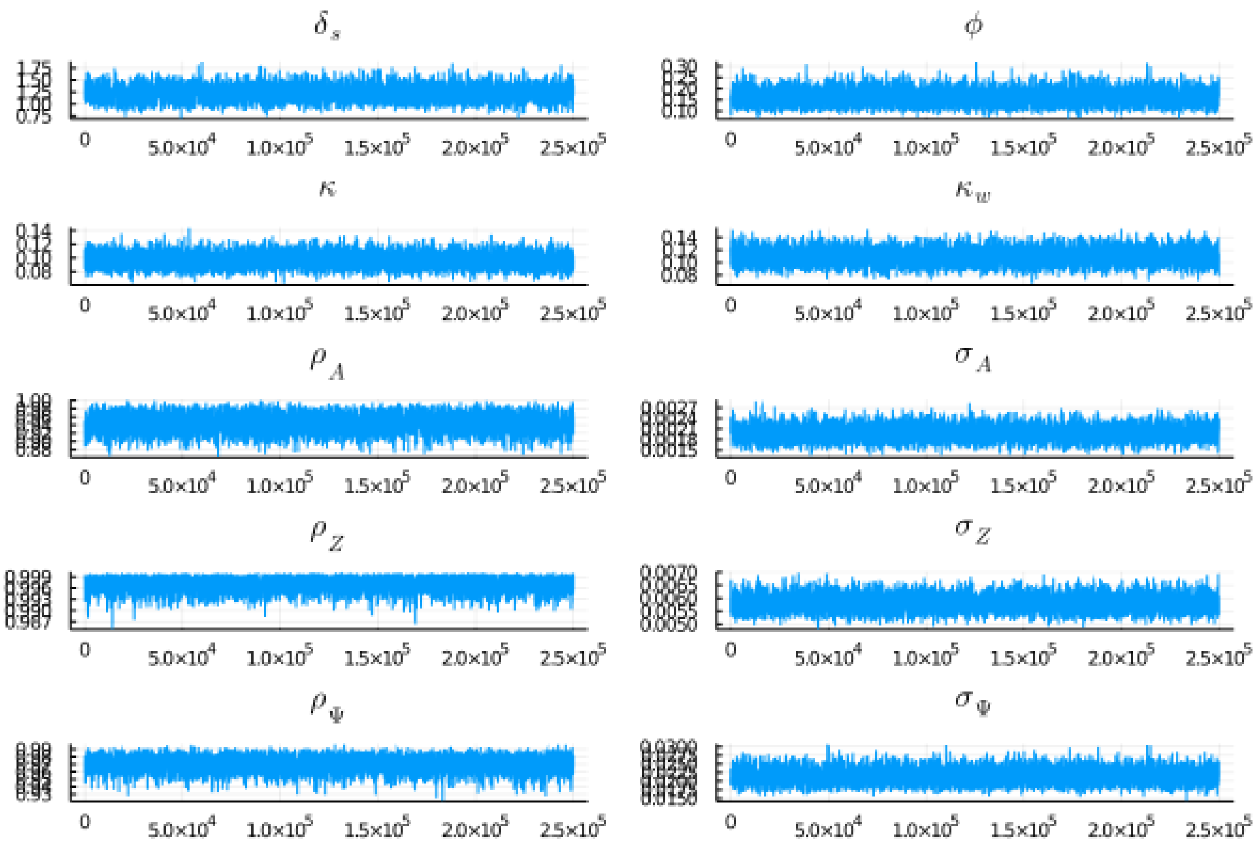


Figure C.4: MCMC draws of HANK model

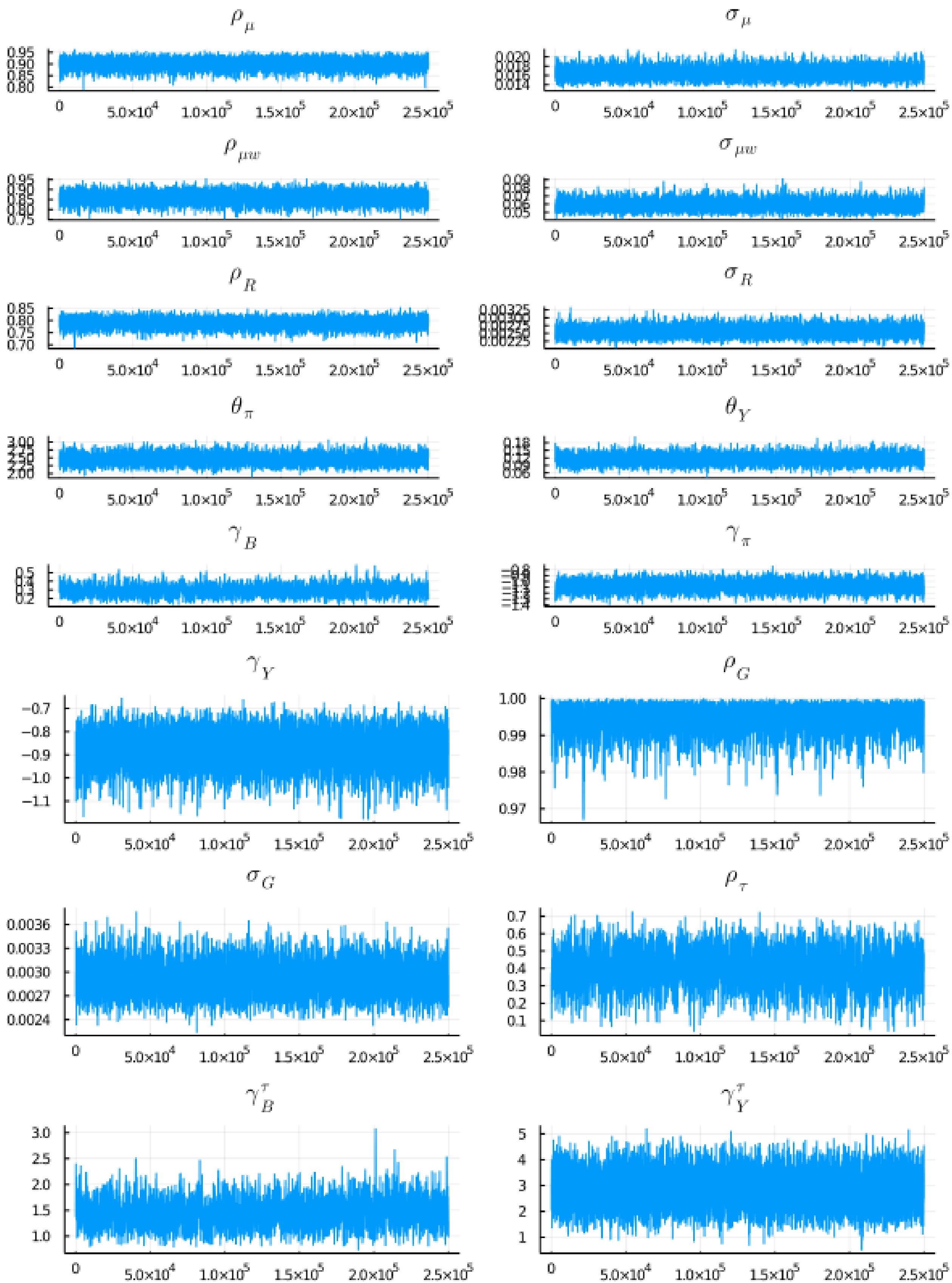


Figure C.5: MCMC draws of HANKX model

Figure C.6: MCMC draws of HANKX model

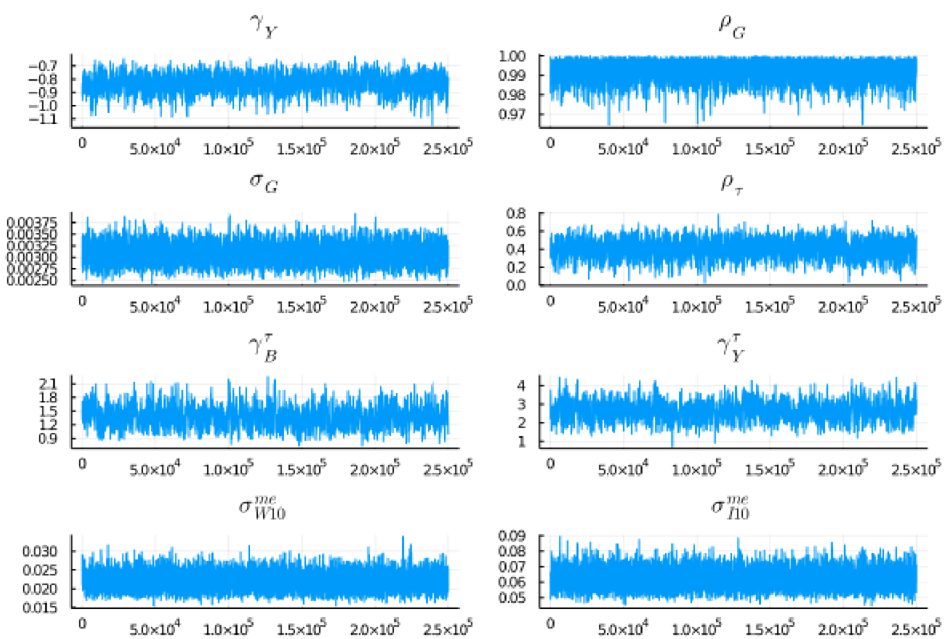
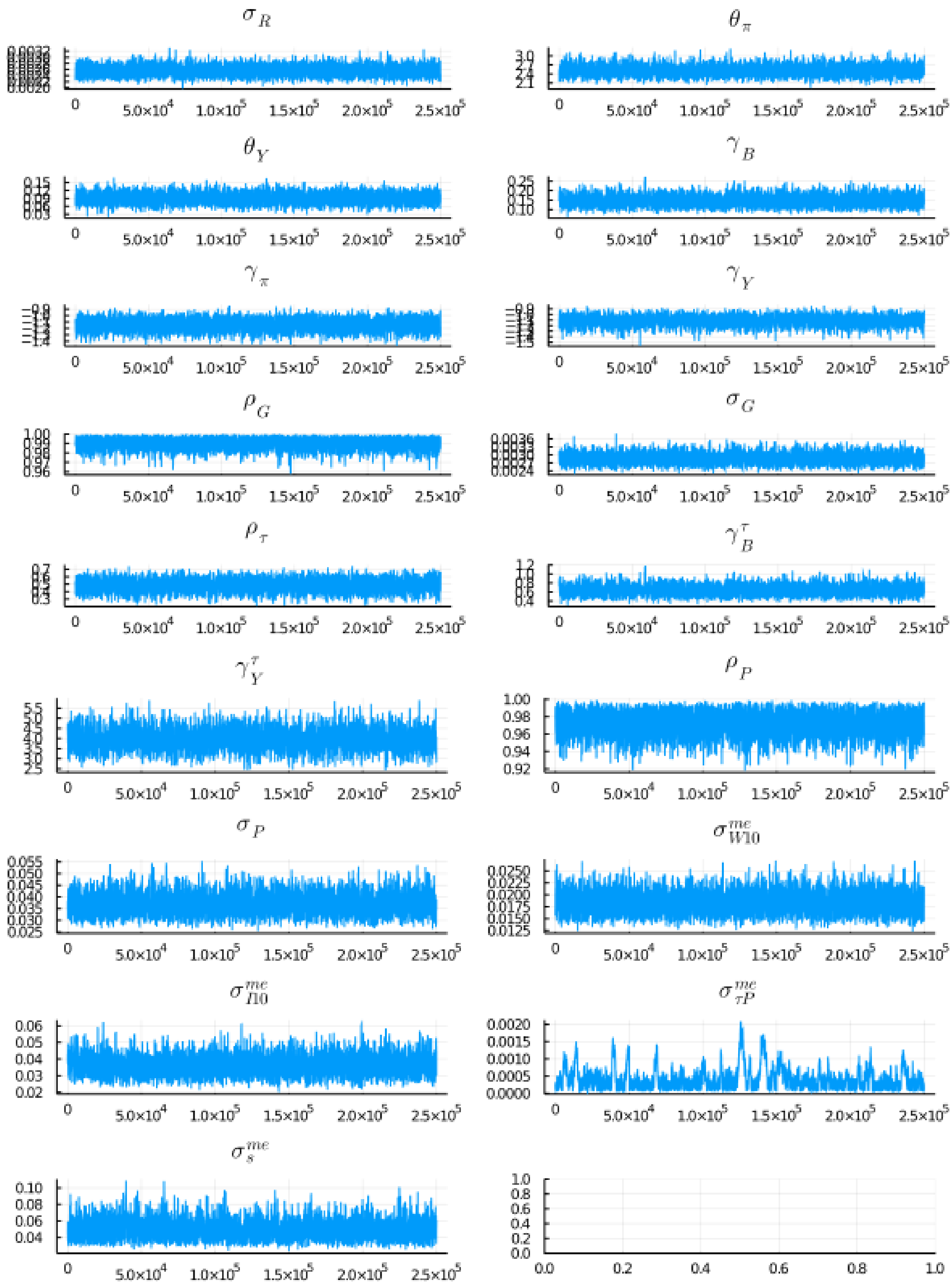


Figure C.7: MCMC draws of HANKX+ model

Figure C.8: MCMC draws of HANKX+ model

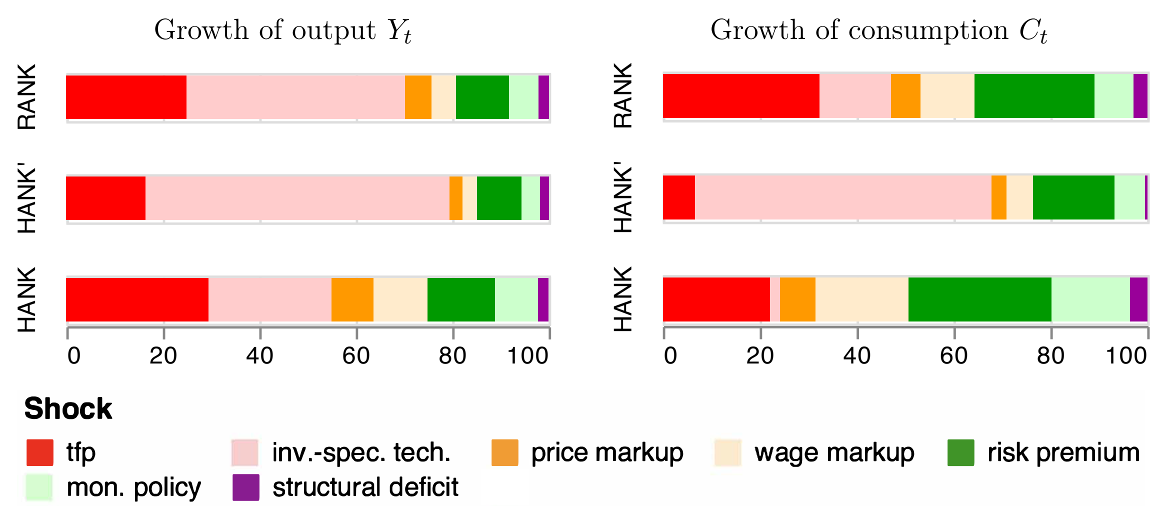


D Inspecting model mechanisms

D.1 HANK with RANK parameterization

Figure D.9 shows the variance decompositions of output and consumption growth for the estimated HANK and RANK models and in addition for the HANK model under the RANK parameter estimates for the aggregate shocks and frictions (HANK'). Under the same parameterization, the difference between the RANK and HANK' model becomes larger. For example, investment-specific technology shocks are now substantially more important in HANK' than in the RANK model. The opposite is true once the HANK model is estimated, in which case investment-specific technology shocks are slightly less important in HANK than in RANK. Hence, the estimation plays an important role in making the models more similar by adjusting the estimated shocks and frictions.

Figure D.9: Variance decompositions: HANK with RANK estimates

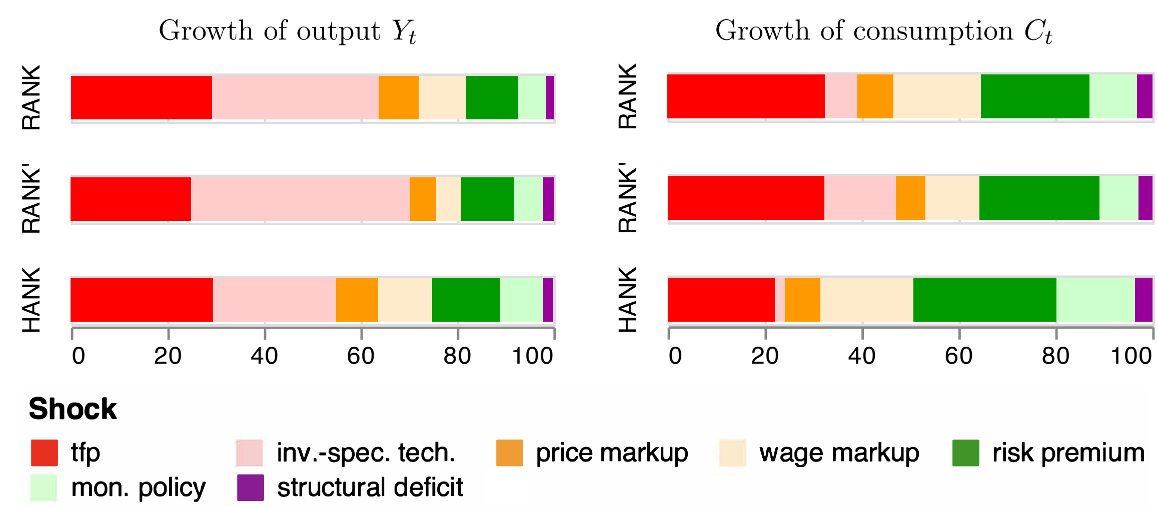


Notes: Conditional variance decompositions at a 4-quarter forecast horizon for 1) the estimated RANK model, 2) the HANK model with RANK parameter estimates (denoted HANK'), and 3) the estimated HANK model.

D.2 RANK with HANK steady-state parameters

In the main text, we calibrate the RANK model to match the ratios of capital-to-output and government-spending-to-output. In Figure D.10, we compare the estimation results of the RANK model with the same steady-state parameters as in the HANK model (denoted RANK'). As in the main text, investment-specific technology shocks are less important in HANK than in RANK, while shocks to the risk premium and monetary policy become more important for understanding output and consumption growth.

Figure D.10: Variance decompositions: RANK with HANK steady-state parameters

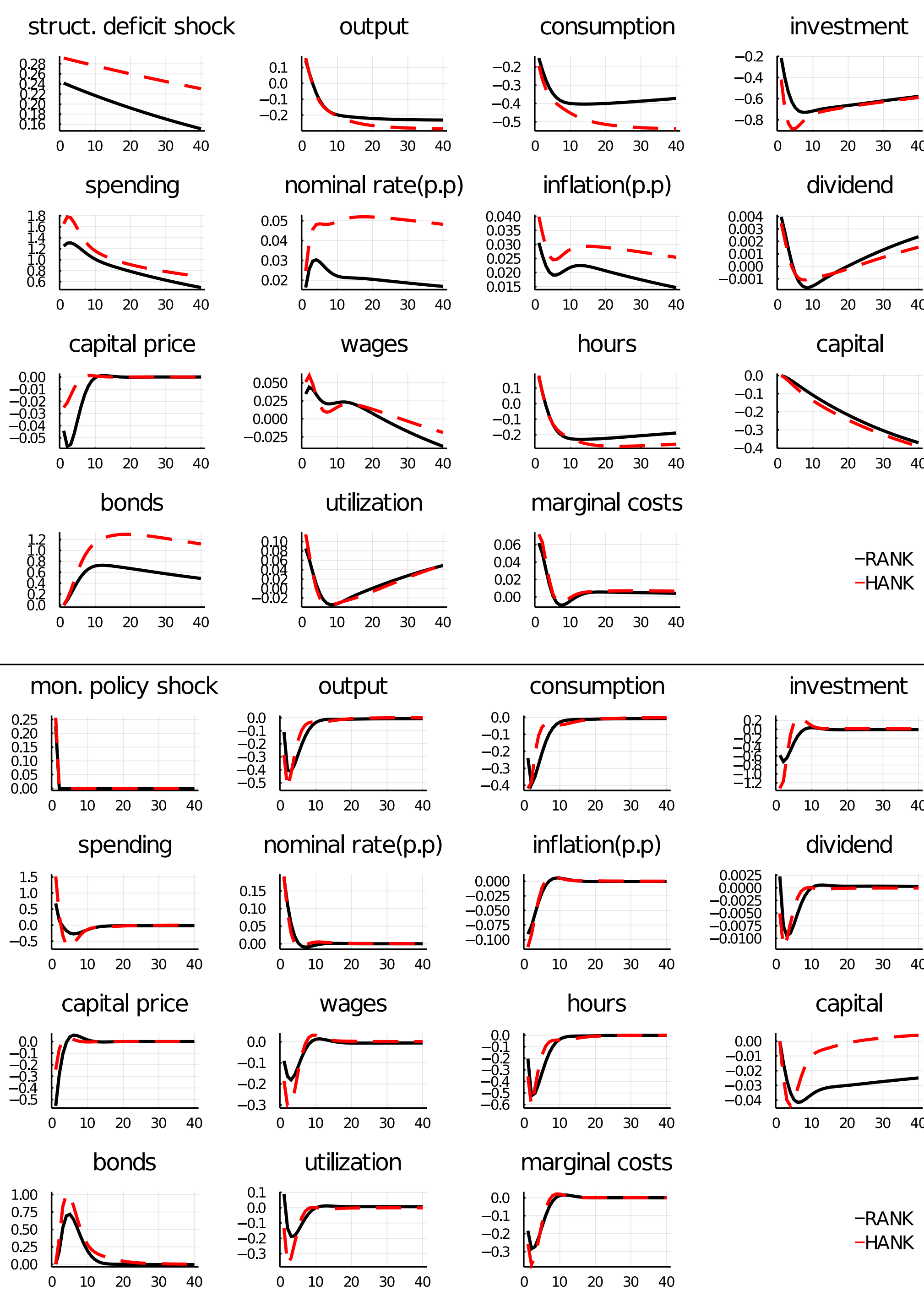


Notes: Conditional variance decompositions at a 4-quarter forecast horizon for 1) the estimated RANK model, 2) the estimated RANK model with HANK steady-state parameterization (denoted RANK'), 3) the estimated HANK model.

D.3 Impulse Responses of Estimated RANK vs. HANK

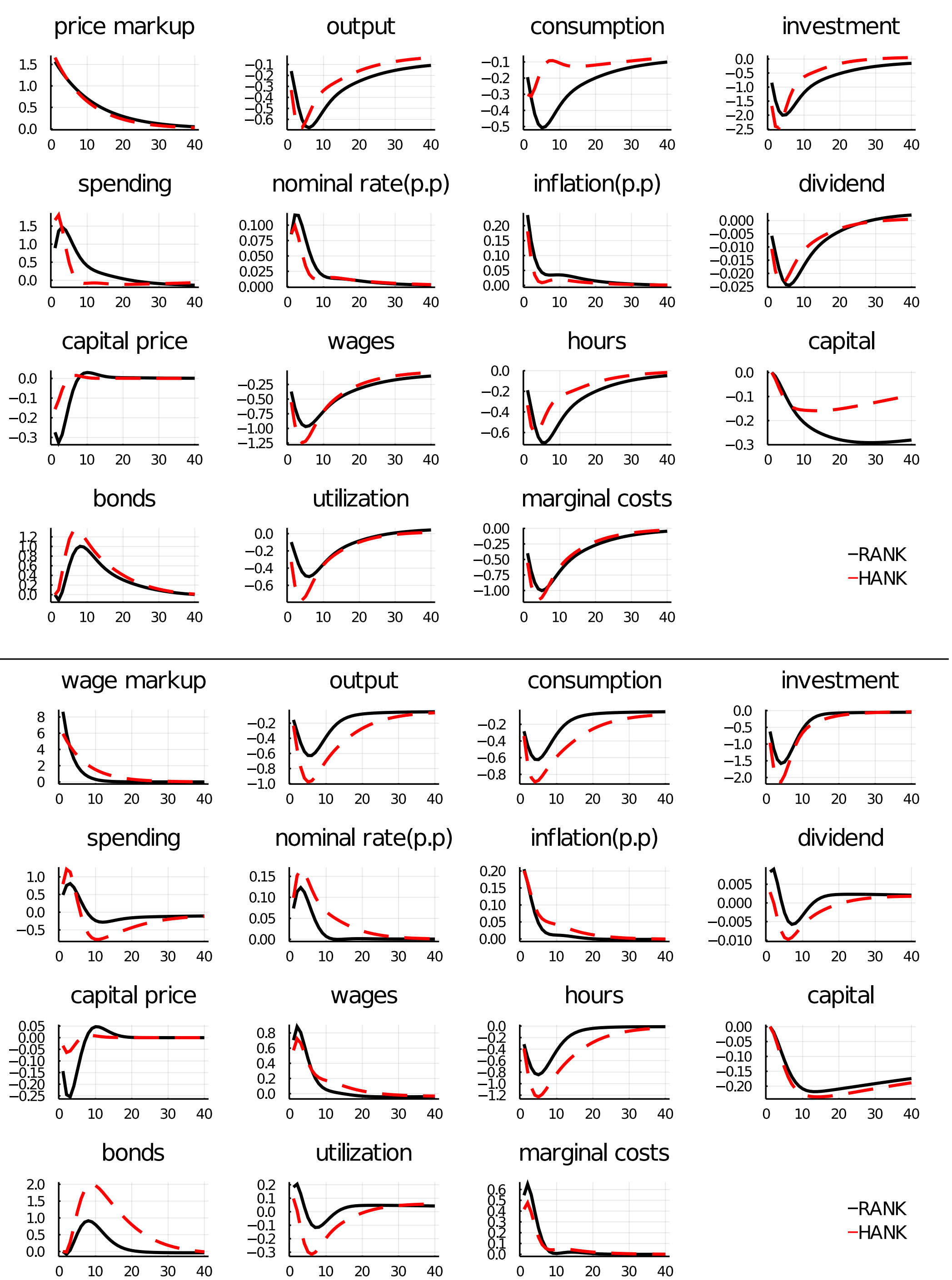
Figures D.11, D.12, D.13, and D.14 plot the impulse response functions for the estimated HANK and RANK model. The first panel on the top left corner of each figure shows the shock and the remaining panels show the response of aggregate variables.

Figure D.11: IRFs to structural deficit and monetary policy shocks



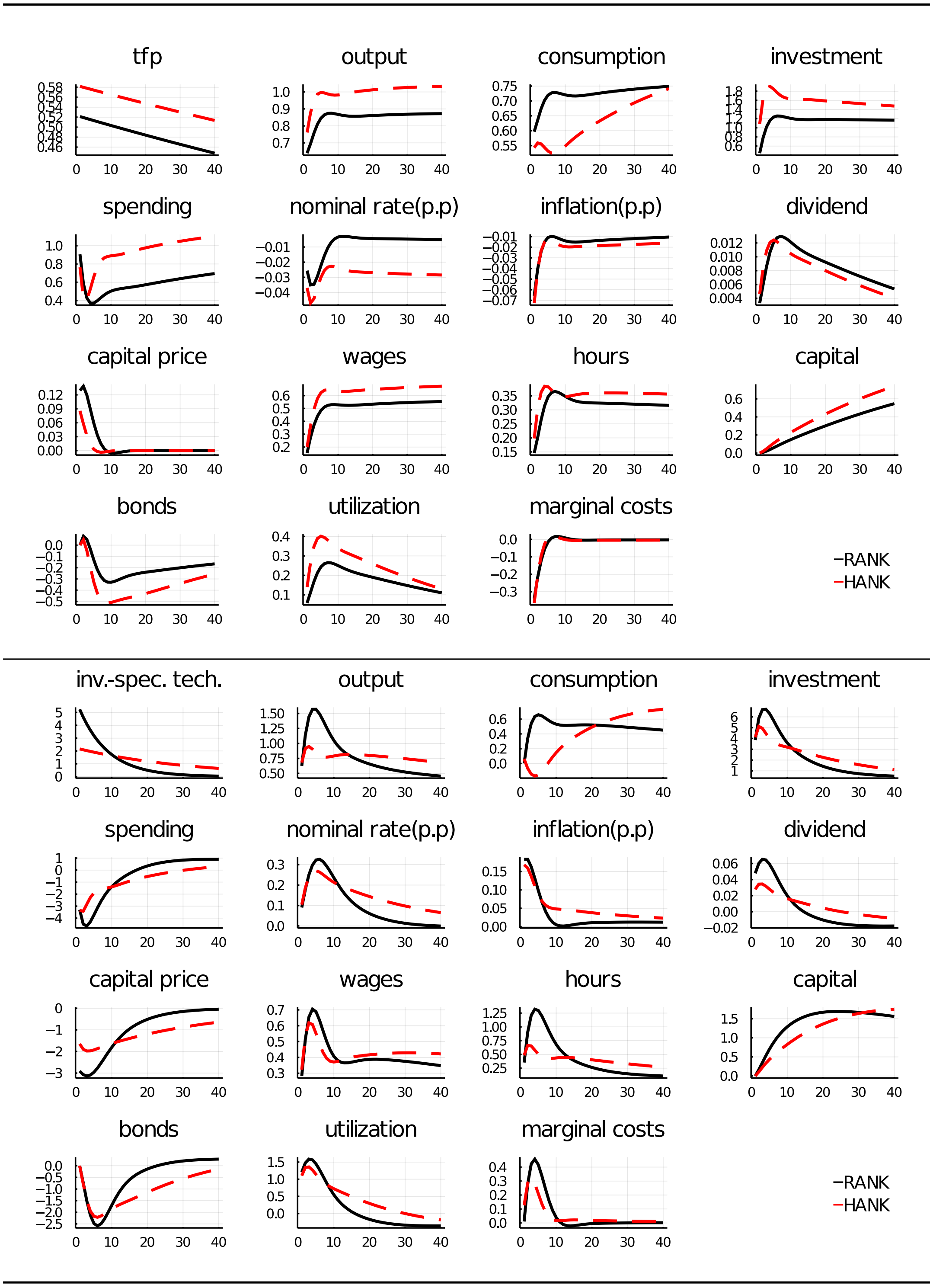
Notes: Top: IRF to a structural deficit shock. Bottom: IRF to a monetary policy shock. Black solid line: RANK. Red dashed line: HANK.

Figure D.12: IRFs to markup shocks



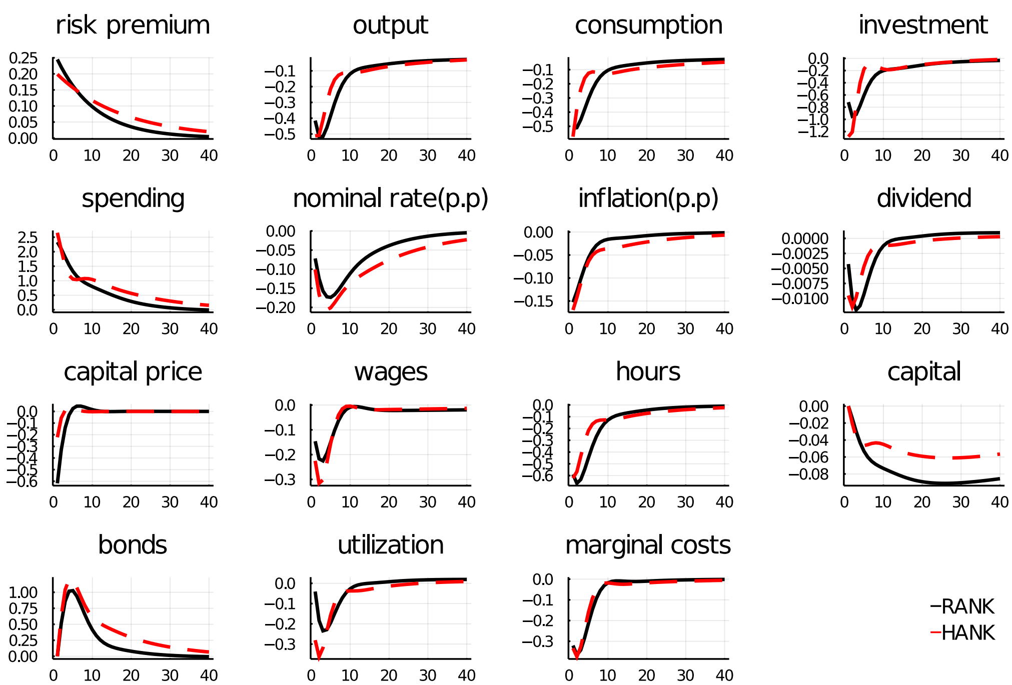
Notes: Top: IRF to a price-markup shock. Bottom: IRF to a wage-markup shock. Black solid line: RANK. Red dashed line: HANK.

Figure D.13: IRFs to technology shocks



Notes: Top: IRF to a TFP shock. Bottom: IRF to an MEI shock. Black solid line: RANK. Red dashed line: HANK.

Figure D.14: IRFs to risk premium shock

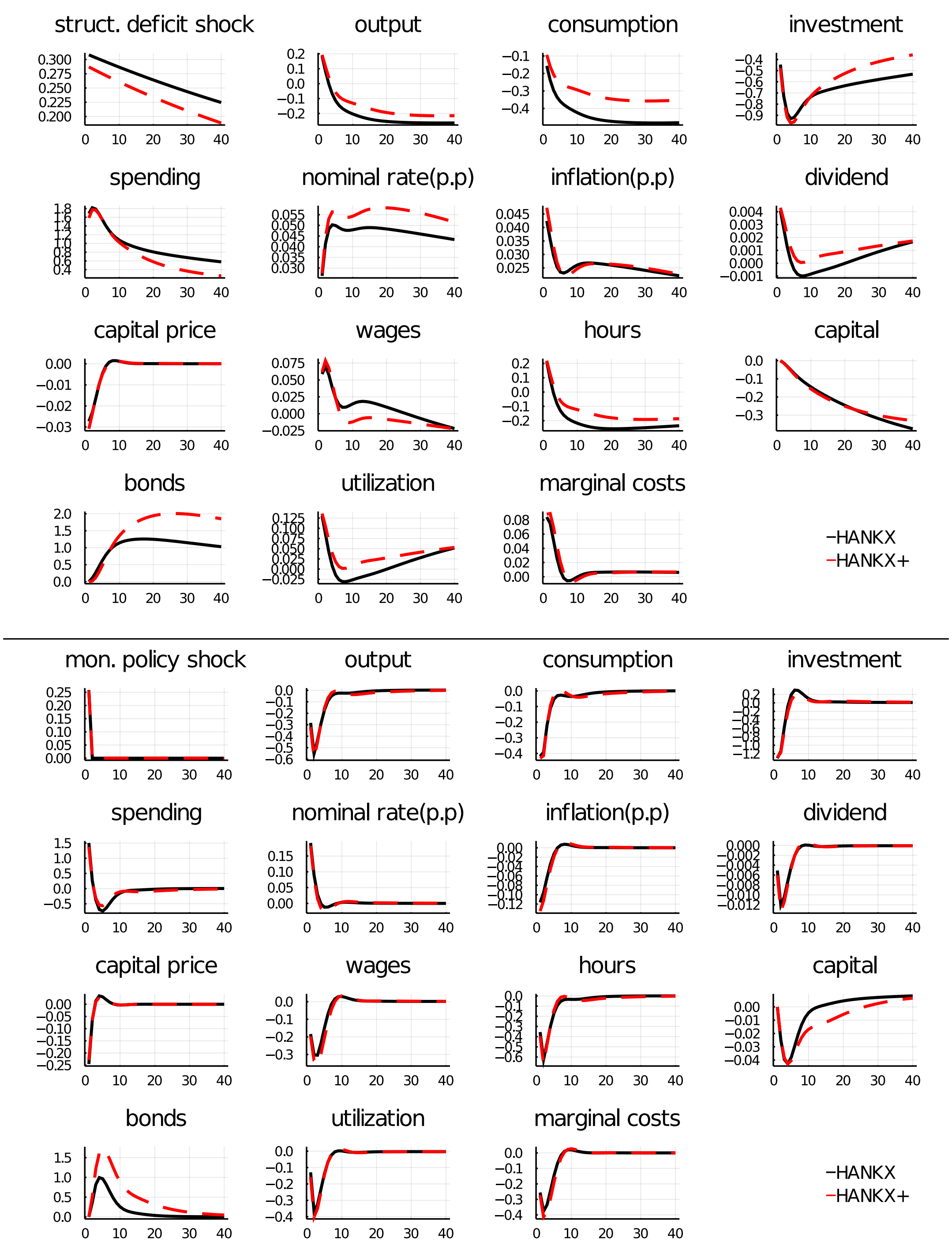


Notes: IRF to a risk premium shock. Black solid line: RANK. Red dashed line: HANK.

D.4 Impulse Responses of Estimated HANKX vs. HANKX+

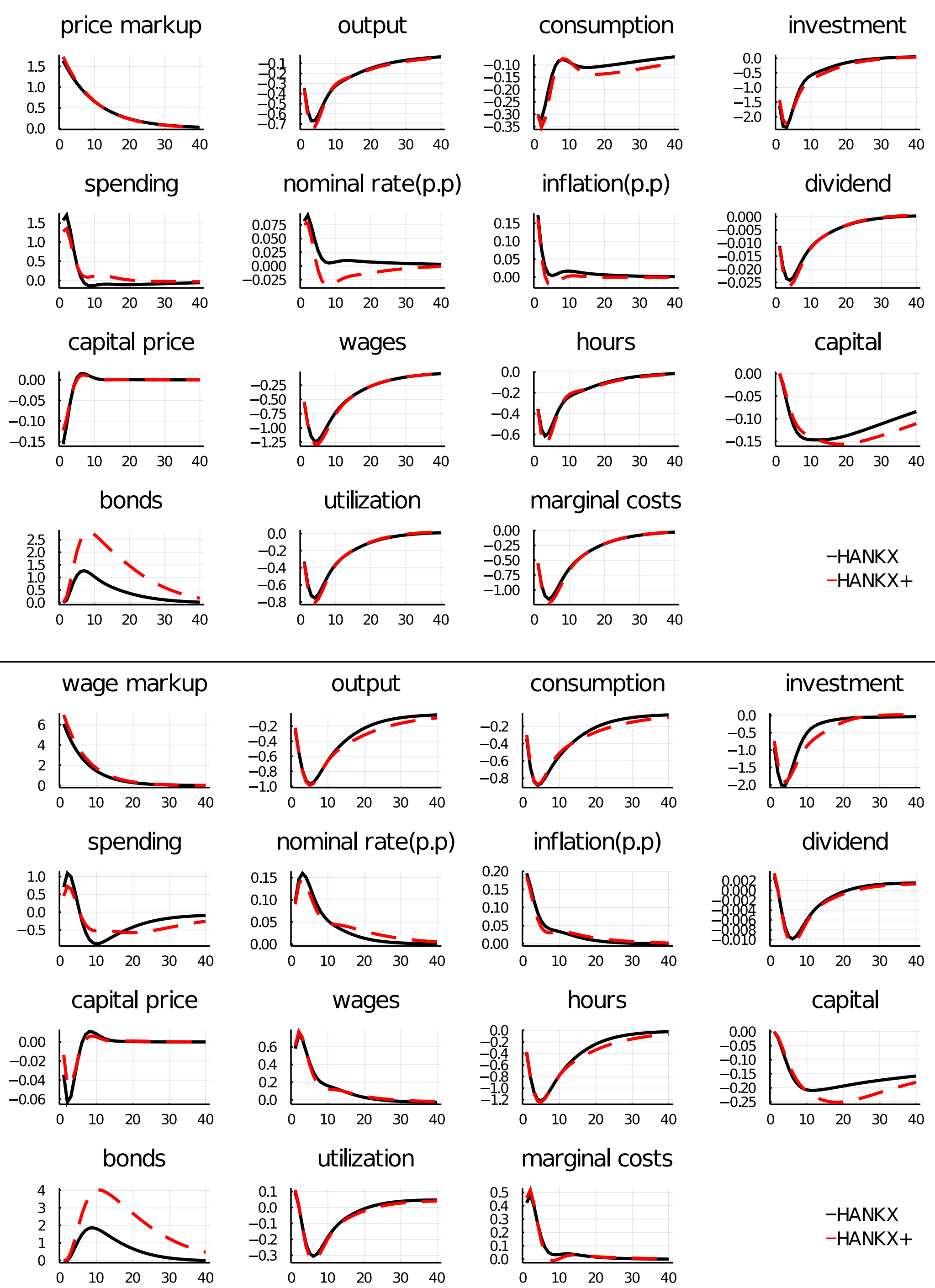
Figures D.15, D.16, D.17, D.18, and D.19 plot the impulse response functions for the estimated HANKX and HANKX+ model. The first panel on the top left corner of each figure shows the shock and the remaining panels show the response of aggregate variables.

Figure D.15: IRFs to structural deficit and monetary policy shocks



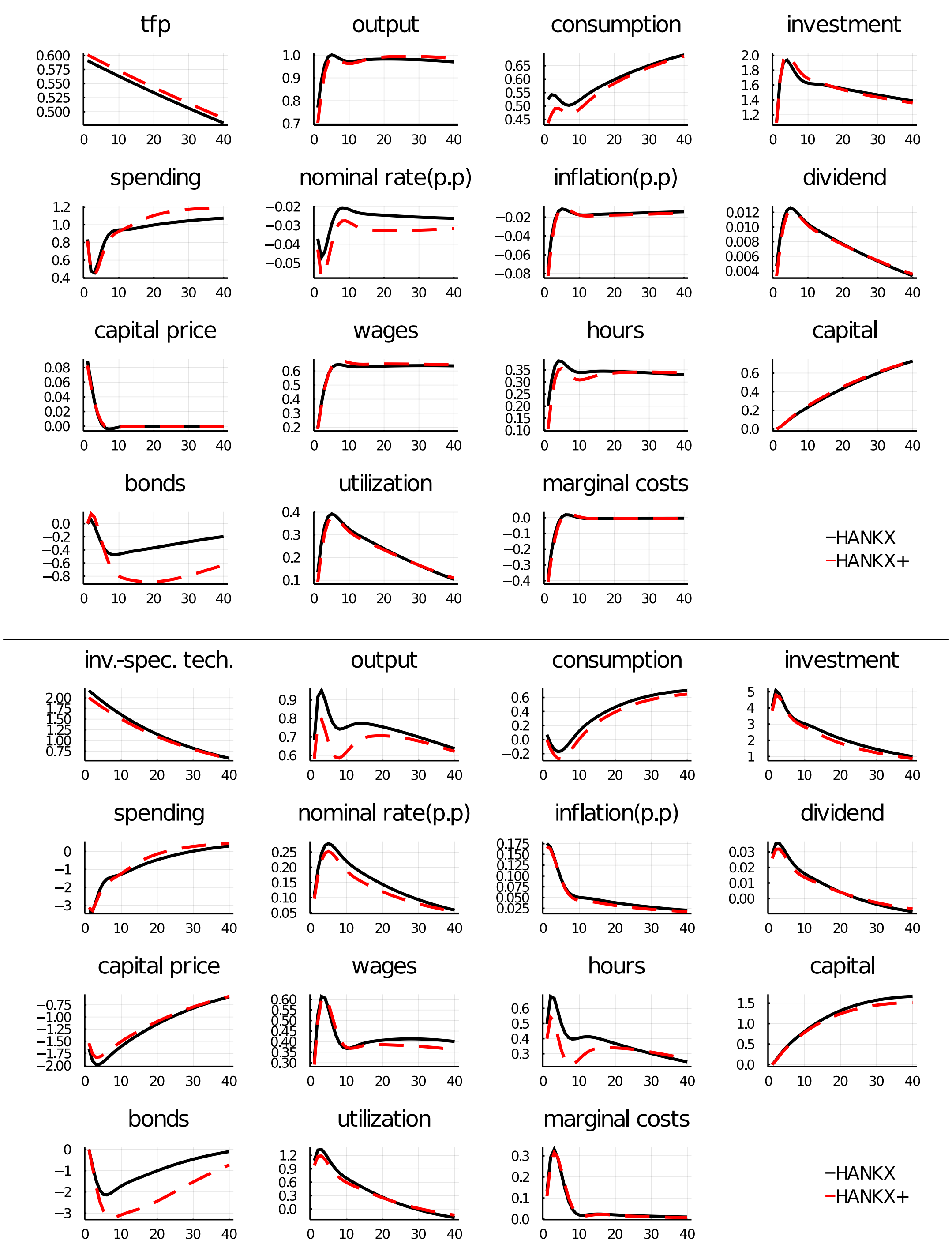
Notes: Top: IRF to a structural deficit shock. Bottom: IRF to a monetary policy shock. Black solid line: HANKX. Red dashed line: HANKX+.

Figure D.16: IRFs to markup shocks



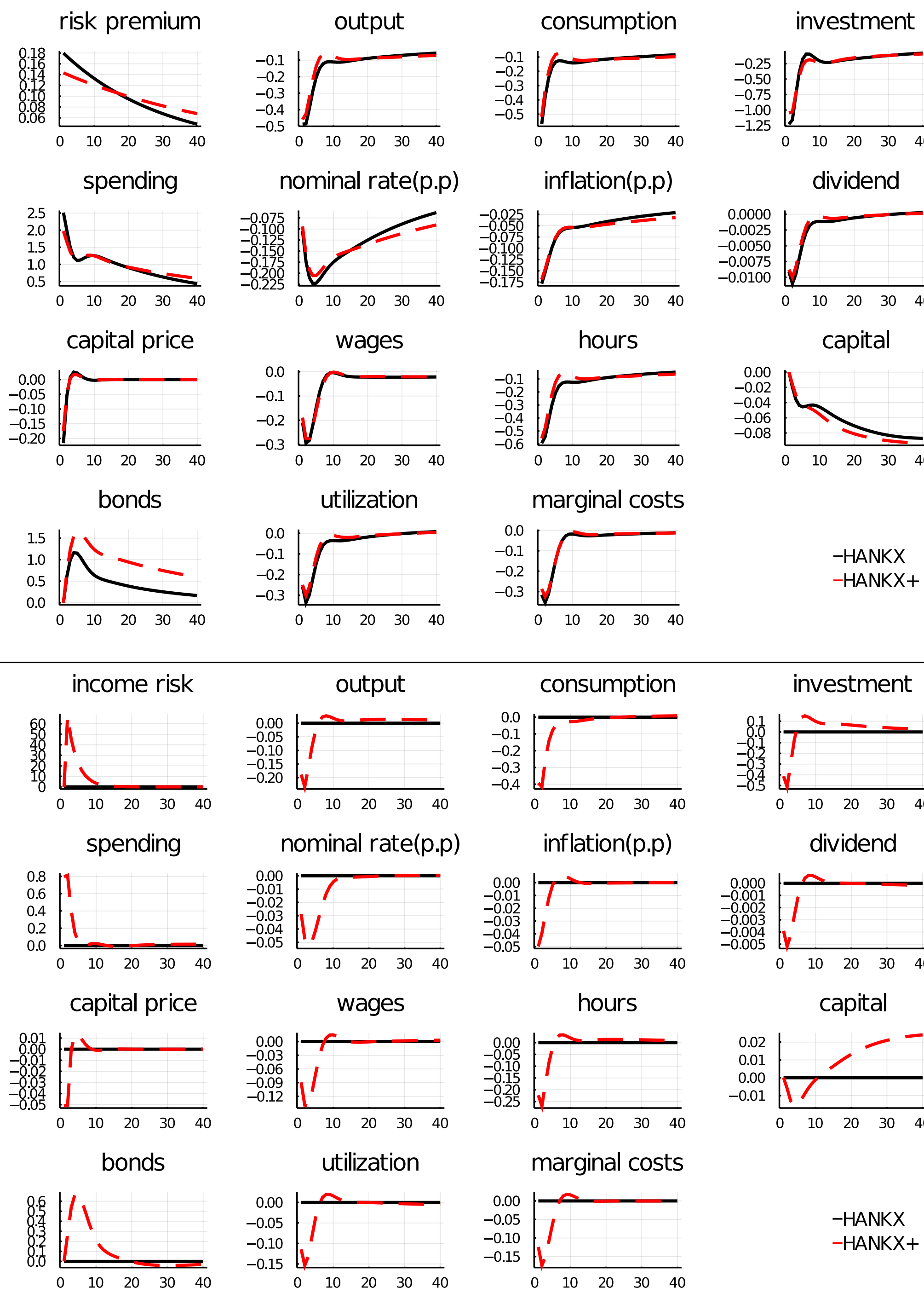
Notes: Top: IRF to a price-markup shock. Bottom: IRF to a wage-markup shock. Black solid line: HANKX. Red dashed line: HANKX+.

Figure D.17: IRFs to technology shocks



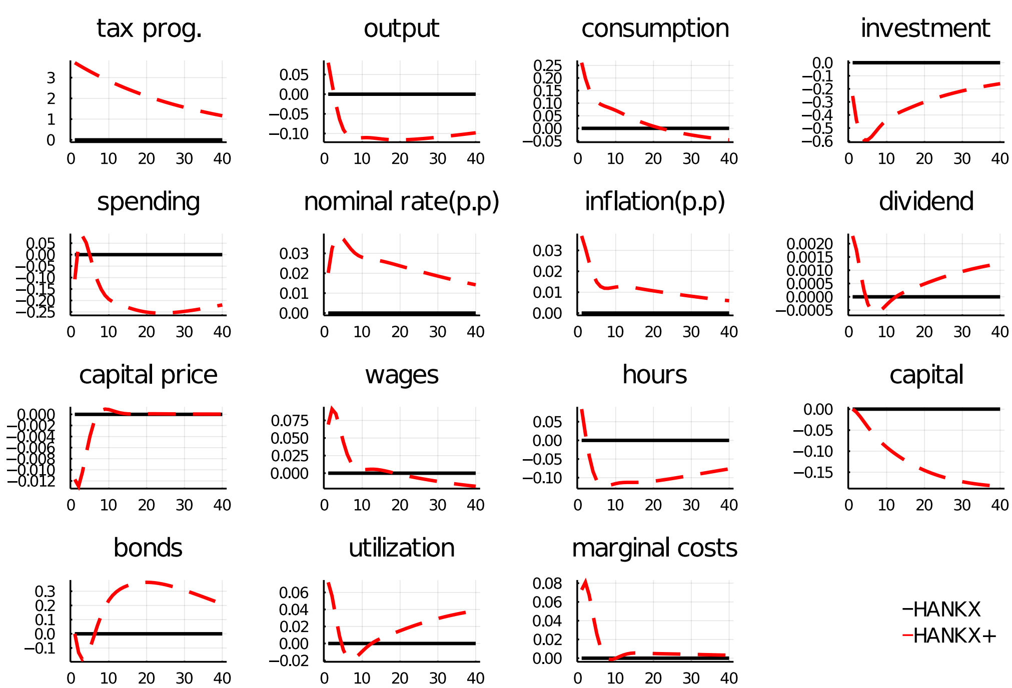
Notes: Top: IRF to a TFP shock. Bottom: IRF to an MEI shock. Black solid line: HANKX. Red dashed line: HANKX+.

Figure D.18: IRFs to risk premium and income risk shocks



Notes: Top: IRF to a risk premium shock. Bottom: IRF to an income risk shock. Black solid line: HANKX. Red dashed line: HANKX+.

Figure D.19: IRFs to a tax progressivity shock

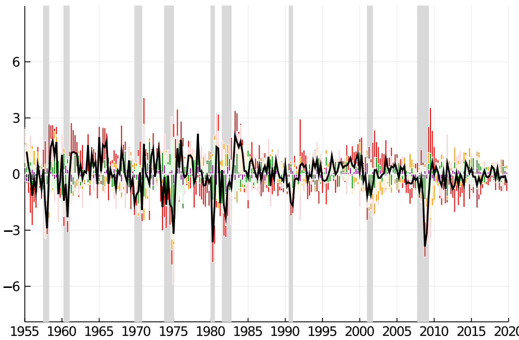


Notes: IRF to a tax progressivity shock. Black solid line: HANKX. Red dashed line: HANKX+.

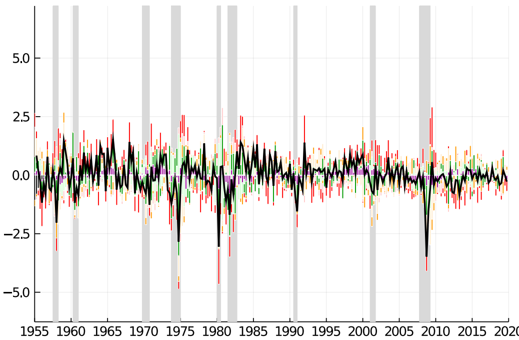
E Historical decompositions of observables: HANK and HANKX+

Figure E.20 shows the historical decomposition of all observables for the estimation of the HANK model. Figures E.21 and E.22 show the historical decomposition of all observables for the estimation of the HANKX+ model. shows the historical decomposition of non-observed variables markups and liquidity premium in the HANKX+ model.

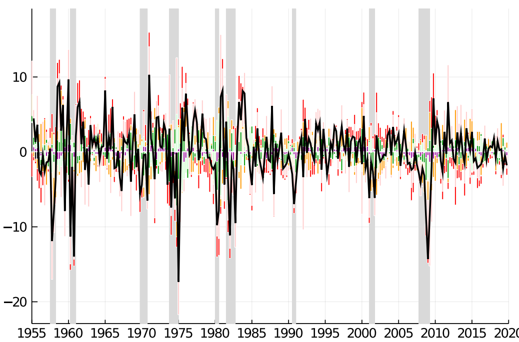
Figure E.20: Historical decompositions of all observables in HANK



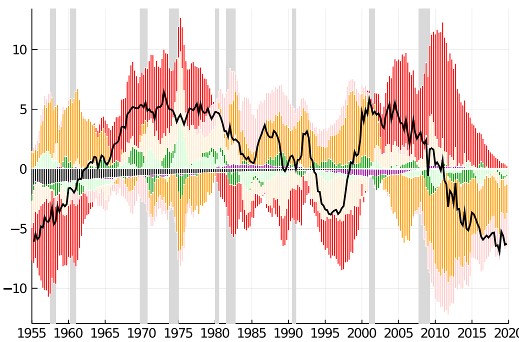
(a) Output growth



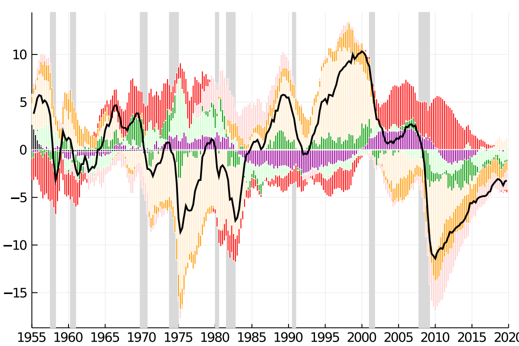
(b) Consumption growth



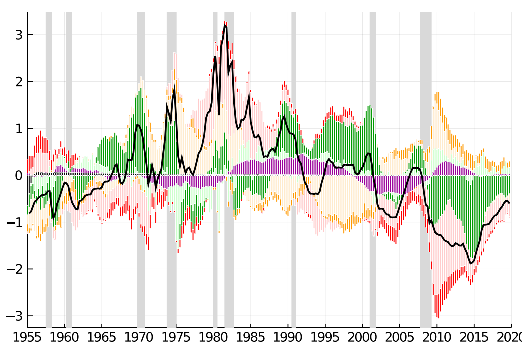
(c) Investment growth



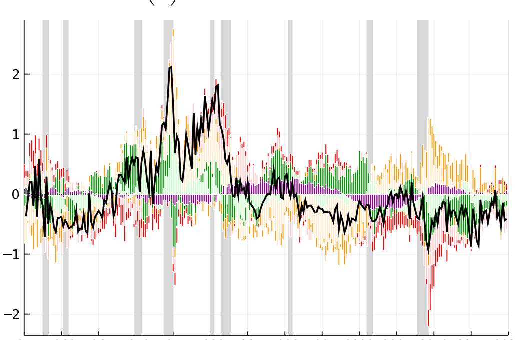
(d) Real wage



(e) Hours worked



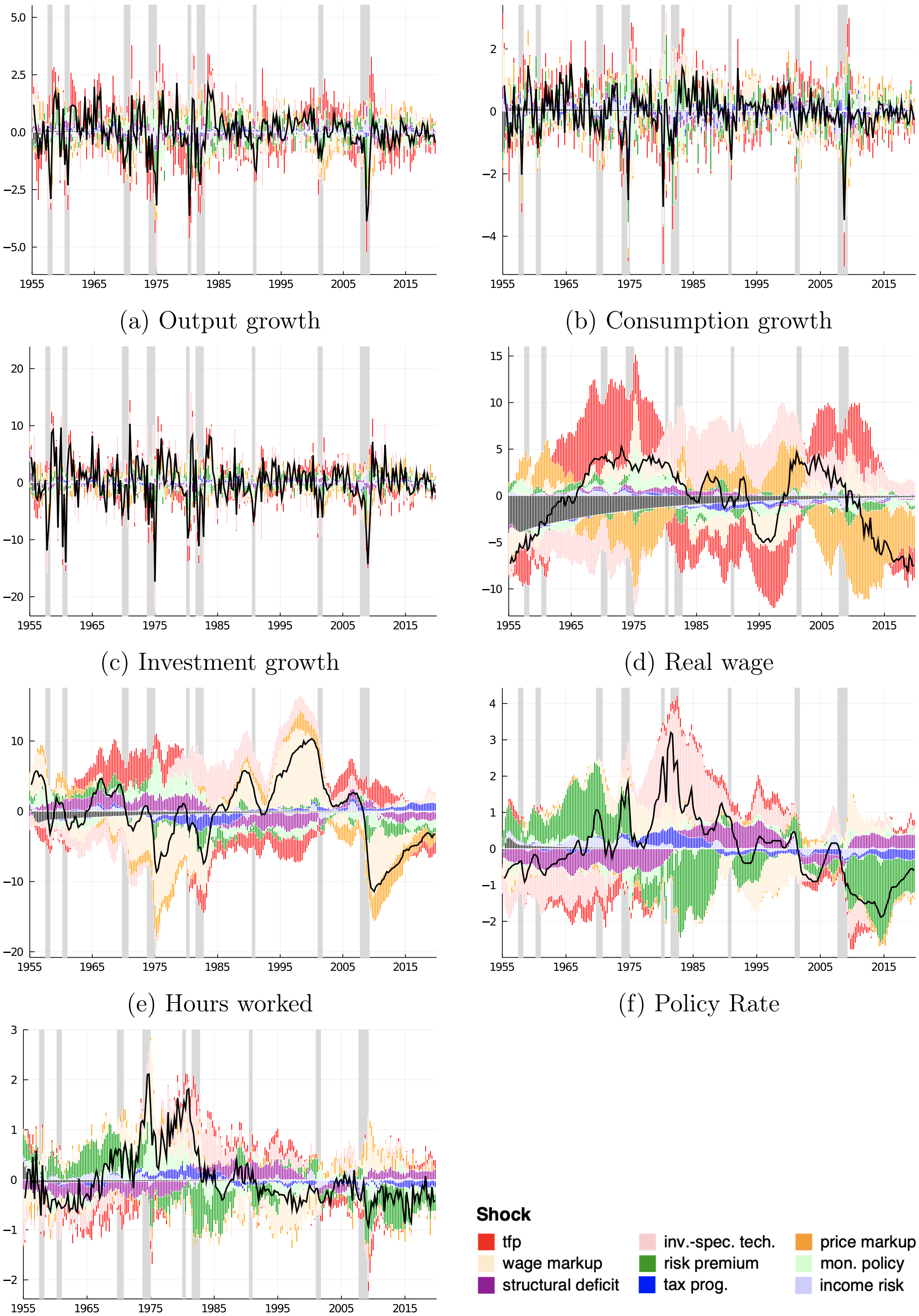
(f) Policy Rate



(e) Inflation

Notes: Historical decompositions of all observables in HANK. Shaded areas correspond to NBER-dated recessions.

Figure E.21: Historical decompositions of all observables in HANKX+



(a) Output growth (b) Consumption growth

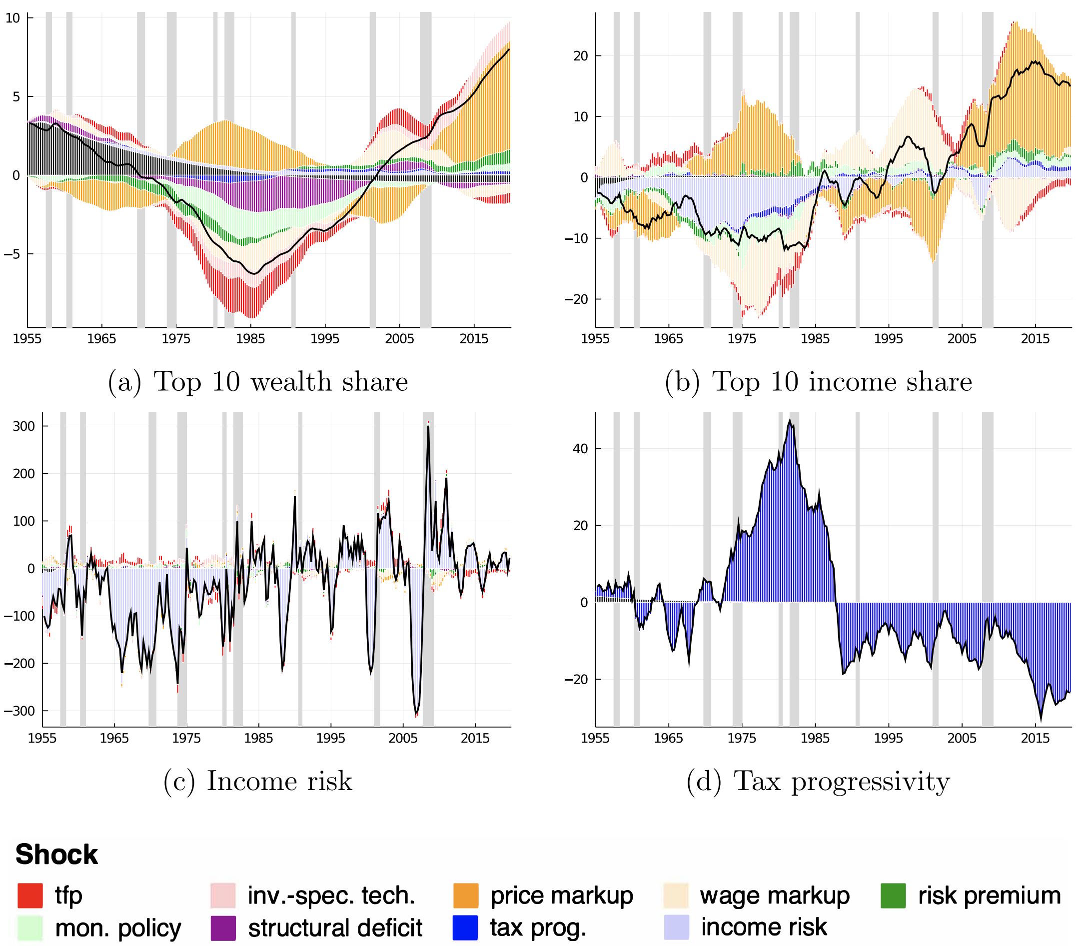
(c) Investment growth (d) Real wage

(e) Hours worked (f) Policy Rate

(e) Inflation

Notes: Historical decompositions of all observables in HANKX+. Shaded areas correspond to NBER-dated recessions.

Figure E.22: Historical decompositions of all observables in HANKX+ ctd.



(a) Top 10 wealth share

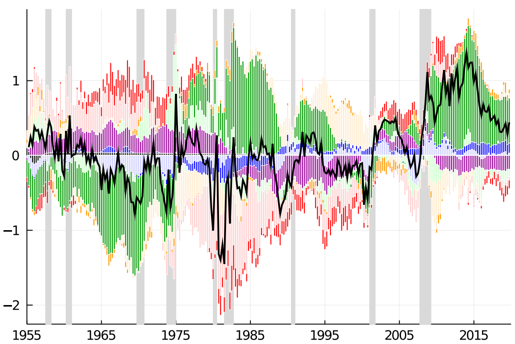
(b) Top 10 income share

(c) Income risk

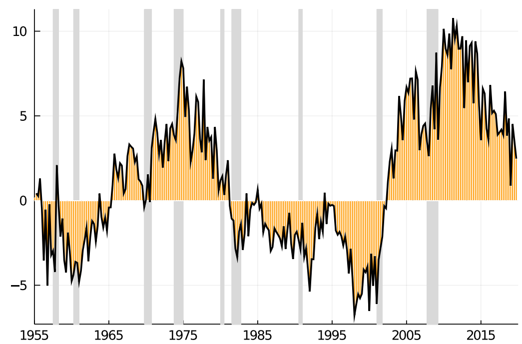
(d) Tax progressivity

Notes: Historical decompositions of all observables in HANKX+. Shaded areas correspond to NBER-dated recessions.

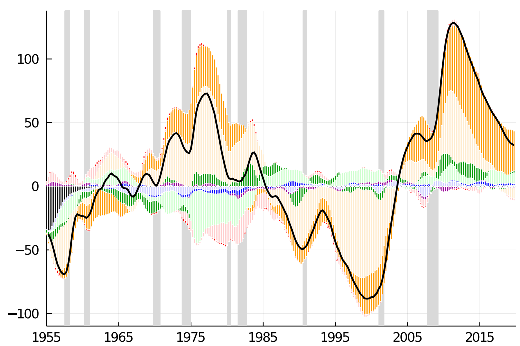
Figure E.23: Further historical decompositions of non-observed variables in HANKX+



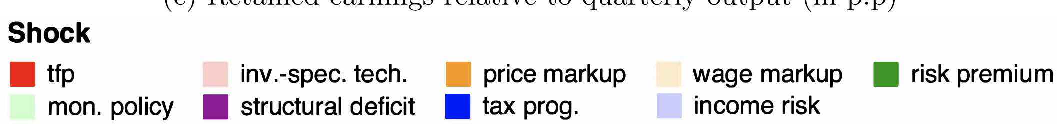
(a) Liquidity premium



(b) Price markups



(c) Retained earnings relative to quarterly output (in p.p)



Notes: Historical decompositions in HANKX+. Shaded areas correspond to NBER-dated recessions.

F Further model variants

Table F.4: Prior and posterior distributions of estimated parameters

Parameter		Prior		Posterior		
	Distribution	Mean	Std. Dev.	HANK (RA2)	HANK (KPR)	HANKX (ret. earn.)
			Frictions			
δ_s	Gamma	5.00	2.00	1.888	1.855	1.398
				(1.527, 2.310)	(1.349, 2.389)	(1.182, 1.629)
ϕ	Gamma	4.00	2.00	0.096	0.173	0.144
				(0.062, 0.137)	(0.126, 0.229)	(0.106, 0.188)
κ	Gamma	0.10	0.01	0.097	0.079	0.099
				(0.083, 0.113)	(0.069, 0.090)	(0.085, 0.114)
κ_w	Gamma	0.10	0.01	0.111	0.110	0.101
				(0.093, 0.129)	(0.091, 0.130)	(0.086, 0.118)
			Debt and monetary policy rules			
ρ_R	Beta	0.50	0.20	0.762	0.738	0.780
				(0.730, 0.791)	(0.705, 0.770)	(0.749, 0.810)
σ_R	Inv.-Gamma	0.10	2.00	0.258	0.237	0.266
				(0.235, 0.284)	(0.217, 0.257)	(0.241, 0.292)
θ_π	Normal	1.70	0.30	2.285	1.711	2.504
				(2.089, 2.499)	(1.572, 1.864)	(2.275, 2.750)
θ_Y	Normal	0.13	0.05	0.109	0.133	0.100
				(0.084, 0.136)	(0.113, 0.154)	(0.073, 0.127)
γ_B	Gamma	0.10	0.08	0.341	0.056	0.174
				(0.252, 0.441)	(0.026, 0.090)	(0.131, 0.221)
γ_π	Normal	0.00	1.00	-1.137	-1.375	-1.153
				(-1.281, -0.998)	(-1.531, -1.218)	(-1.286, -1.021)
γ_Y	Normal	0.00	1.00	-0.937	-0.479	-1.137
				(-1.067, -0.821)	(-0.54, -0.422)	(-1.281, -1.008)
ρ_D	Beta	0.50	0.20	0.983	0.972	0.994
				(0.965, 0.995)	(0.952, 0.989)	(0.987, 0.998)
σ_D	Inv.-Gamma	0.10	2.00	0.315	0.421	0.295
				(0.283, 0.350)	(0.380, 0.467)	(0.267, 0.325)
			Tax rules			
ρ_τ	Beta	0.50	0.20	0.335	0.351	0.489
				(0.181, 0.486)	(0.116, 0.646)	(0.375, 0.600)
γ_B^τ	Normal	0.00	1.00	1.667	-4.029	0.779
				(1.284, 2.099)	(-4.711, -3.427)	(0.602, 0.985)
γ_Y^τ	Normal	0.00	1.00	3.116	-0.002	4.252
				(2.153, 4.180)	(-0.155, 0.142)	(3.516, 5.032)
ρ_P	Beta	0.50	0.20	–	–	–
σ_P	Inv.-Gamma	0.10	2.00	–	–	–
			Structural shocks			
ρ_A	Beta	0.50	0.20	0.929	0.962	0.984
				(0.890, 0.966)	(0.946, 0.976)	(0.968, 0.995)
σ_A	Inv.-Gamma	0.10	2.00	0.185	0.159	0.155
				(0.153, 0.219)	(0.140, 0.180)	(0.137, 0.175)

Table F.4: Prior and posterior distributions of estimated parameters - continued

Parameter	Prior			Posterior		
	Distribution	Mean	Std. Dev.	RANK	HANK	HANKX
ρ_Z	Beta	0.50	0.20	0.994	0.913	0.994
				(0.989, 0.998)	(0.904, 0.922)	(0.990, 0.998)
σ_Z	Inv.-Gamma	0.10	2.00	0.596	1.215	0.599
				(0.554, 0.641)	(1.124, 1.312)	(0.557, 0.645)
ρ_Ψ	Beta	0.50	0.20	0.966	0.920	0.973
				(0.943, 0.985)	(0.880, 0.956)	(0.962, 0.983)
σ_Ψ	Inv.-Gamma	0.10	2.00	1.881	1.410	2.047
				(1.621, 2.181)	(1.208, 1.632)	(1.795, 2.330)
ρ_μ	Beta	0.50	0.20	0.876	0.958	0.894
				(0.842, 0.907)	(0.941, 0.974)	(0.858, 0.925)
σ_μ	Inv.-Gamma	0.10	2.00	1.736	0.642	1.700
				(1.545, 1.953)	(0.549, 0.750)	(1.515, 1.908)
$\rho_{\mu w}$	Beta	0.50	0.20	0.848	0.666	0.865
				(0.797, 0.897)	(0.607, 0.721)	(0.825, 0.900)
$\sigma_{\mu w}$	Inv.-Gamma	0.10	2.00	5.866	5.680	6.955
				(4.968, 6.919)	(4.559, 7.017)	(6.029, 8.054)
Income risk process						
ρ_s	Beta	0.50	0.20	–	–	–
σ_s	Inv.-Gamma	1.00	2.00	–	–	–
Σ_Y	Normal	1.00	100.00	–	–	–
Retained earnings processes						
ω^F	Uniform	0.50	0.29	–	–	0.221
						(0.113, 0.394)
ω^U	Uniform	0.50	0.29	–	–	0.203
						(0.164, 0.252)
Measurement errors						
σ^{me}_{W10}	Inv.-Gamma	0.05	0.01	–	–	1.930
						(1.645, 2.260)
σ^{me}_{I10}	Inv.-Gamma	0.05	0.01	–	–	5.139
						(4.415, 5.989)
$\sigma^{me}_{\tau P}$	Inv.-Gamma	0.05	0.01	–	–	–
σ^{me}_{s}	Inv.-Gamma	0.05	0.01	–	–	–

Notes: The standard deviations of the shocks and measurement errors have been transformed into percentages by multiplying by 100. HANK (RA2) denotes HANK model with risk aversion 2 instead of 4 (see Appendix F.1), HANK (KPR) denotes HANK model with KPR instead of GHH preferences (see Appendix F.2). HANKX (ret. earn.) is an intermediate variant where we only add retained earnings to the HANKX model (see Section 5).

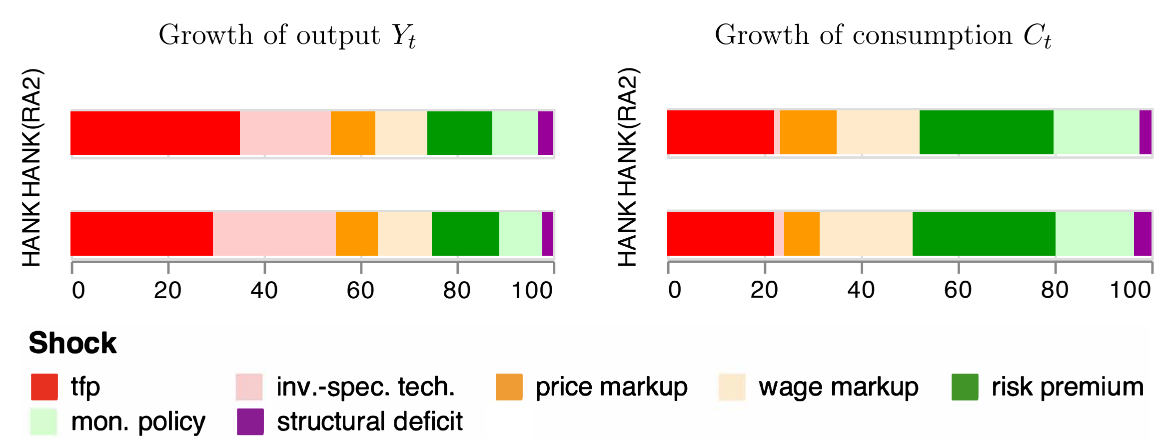
F.1 Alternative risk aversion

We also estimate the model with a coefficient of relative risk aversion of 2 instead of 4. This requires a recalibration of the steady state to match the same targets as listed in Table 3. In particular, we adjust the discount factor, the asset market participation frequency, the fraction of entrepreneurs, and the borrowing penalty. The re-calibration yields $\beta = 0.9912$, $\lambda = 0.065$, $\zeta = 1/2500$, and $\bar{R} = 0.016$.

Figure F.24 shows the variance decomposition of output and consumption growth for the baseline calibration (HANK) and for the new calibration with risk aversion 2 (HANK(RA2)). The decompositions are very similar. Investment-specific technology shocks become even somewhat less important relative to the baseline.

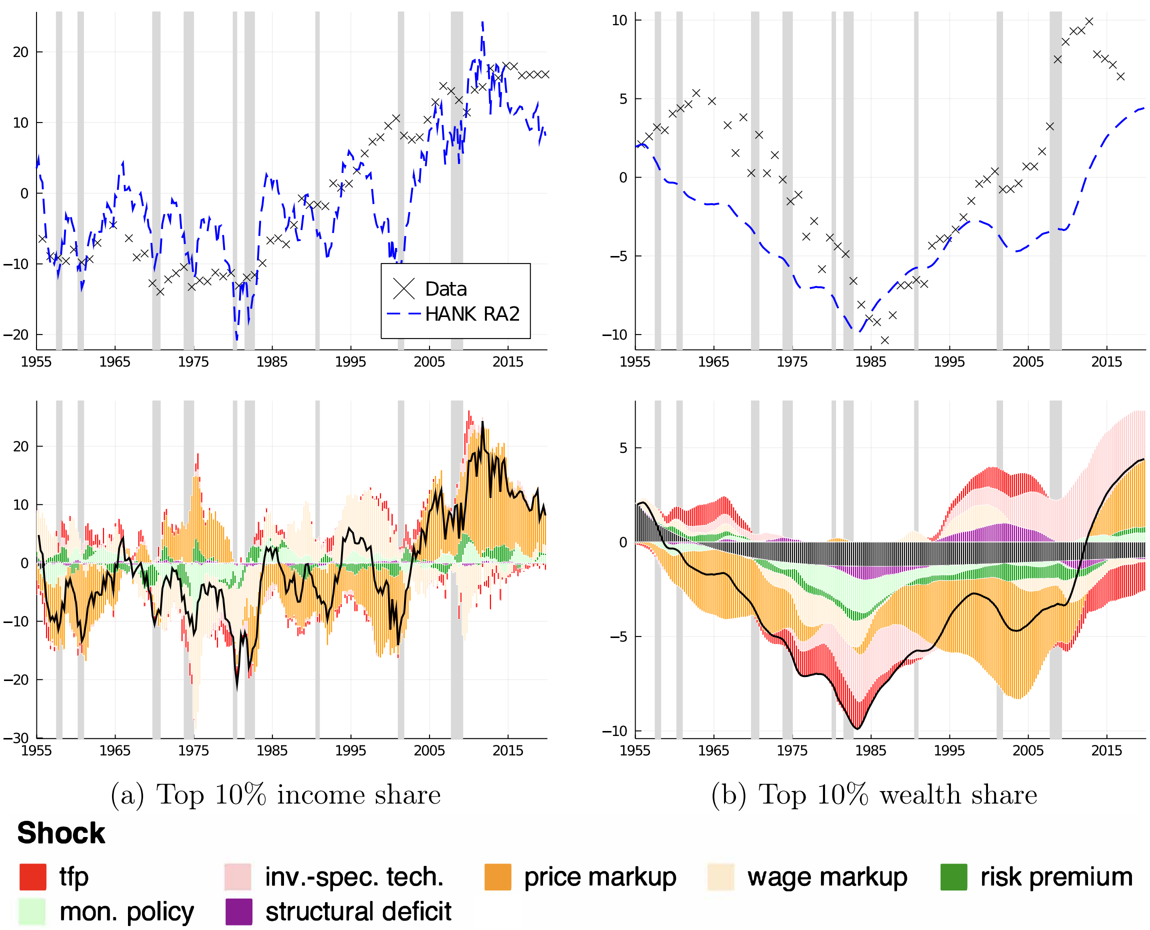
Figure F.25 shows the historical decomposition of US inequality for the estimation only on aggregate data. Our finding of a U-shaped evolution of inequality in line with the US experience is robust to changes in risk aversion.

Figure F.24: Variance decompositions: Output and consumption growth (risk aversion 2)



Notes: Conditional variance decompositions at a 4-quarter forecast horizon for the estimated HANK model with a coefficient of relative risk aversion of 2 and 4 [w/o inequality data].

Figure F.25: US inequality – data vs. model with risk aversion 2



(a) Top 10% income share

(b) Top 10% wealth share

Notes: Top row: Data (crosses) correspond to log-deviations of the annual observations of the share of pre-tax income and wealth held by the top 10% in each distribution in the US taken from the World Inequality Database. Bottom row: Historical decomposition based on the estimated HANK model with risk aversion 2. Shaded areas correspond to NBER-dated recessions.

F.2 KPR preferences

We also estimate the model with preferences that are additively separable in consumption and leisure. The felicity function u now reads:

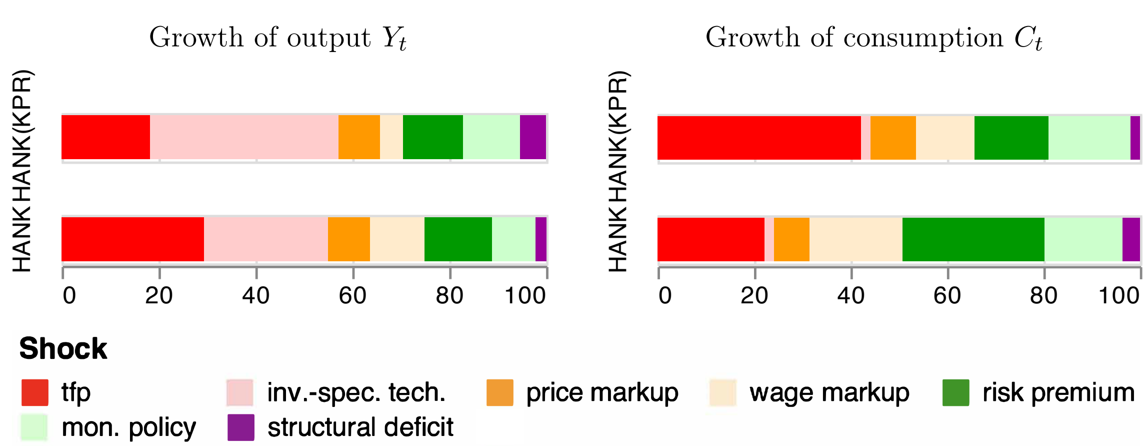
$$u(c_{it}, n_{it}) = \frac{c_{it}^{1-\xi} - 1}{1-\xi} - \Gamma \frac{n_{it}^{1+\gamma} - 1}{1+\gamma},$$

with risk aversion parameter $\xi > 0$ and inverse Frisch elasticity $\gamma > 0$. The first-order condition for labor supply is:

$$n_{it} = \left[\frac{1}{\Gamma} u'(c)(1-\tau_t^P)(1-\tau_t^L)(wh_{it})^{(1-\tau_t^P)}\right]^{\left(\frac{1}{\gamma+\tau_t^P}\right)}.$$

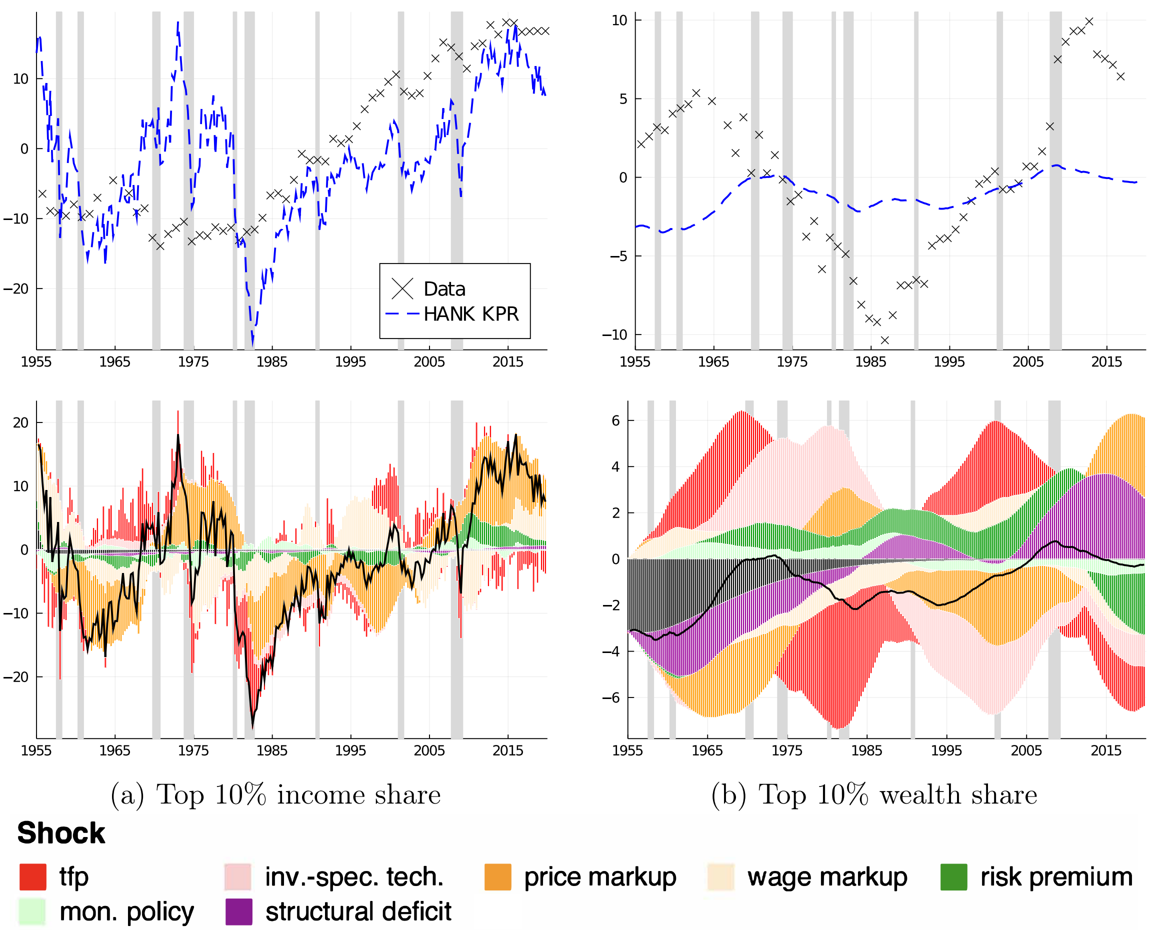
We recalibrate the steady state to match the capital-to-output ratio, the bonds-to-output ratio, the fraction of borrowers, and the wealth held by the top 10% of households as reported in Figure 3. This yields a discount factor of $\beta = 0.987$, a portfolio adjustment probability of $\lambda = 14\%$, a borrowing penalty of $\bar{R} = 3.78\%$, and a probability of becoming an entrepreneur of 0.04%.

Figure F.26: Variance decompositions: Output and consumption growth (KPR preferences)



Notes: Conditional variance decompositions at a 4-quarter forecast horizon for the estimated HANK model with KPR preferences and the baseline HANK model with GHH preferences [w/o inequality data].

Figure F.27: US inequality – data vs. model with KPR preferences



(a) Top 10% income share

(b) Top 10% wealth share

Notes: Top row: Data (crosses) correspond to log-deviations of the annual observations of the share of pre-tax income and wealth held by the top 10% in each distribution in the US taken from the World Inequality Database. Bottom row: Historical decomposition based on the estimated HANK model with KPR preferences. Shaded areas correspond to NBER-dated recessions.

G Further results on inequality dynamics

Why is the model able to explain the slow-moving inequality dynamics? The slow-moving inequality dynamics are not unique to our model but are known in the incomplete markets literature. Typically, this literature studies transitions from steady state to steady state, which can take a very long time. Gabaix et al. (2016), for example, find these transitions to be very slow-moving. These slow dynamics imply that business cycle shocks can accumulate in terms of inequality and an incomplete markets economy that is continuously hit by business cycle shocks generates inequality dynamics in line with the data.

Figure G.28 presents the estimated impulse responses of inequality to all nine shocks based on the HANKX+ model. The general picture is that income inequality shows the most transitory movements. Wealth inequality moves most persistently after all shocks because it is driven by accumulation decisions. Consumption, driven by both income and wealth, shows both short- and long-run dynamics.

Monetary policy and risk premium shocks both increase inequality through a rise in price markups that benefits mostly the (wealthy) entrepreneurs. The effects on income and consumption inequality are rather transitory, while wealth inequality increases persistently.

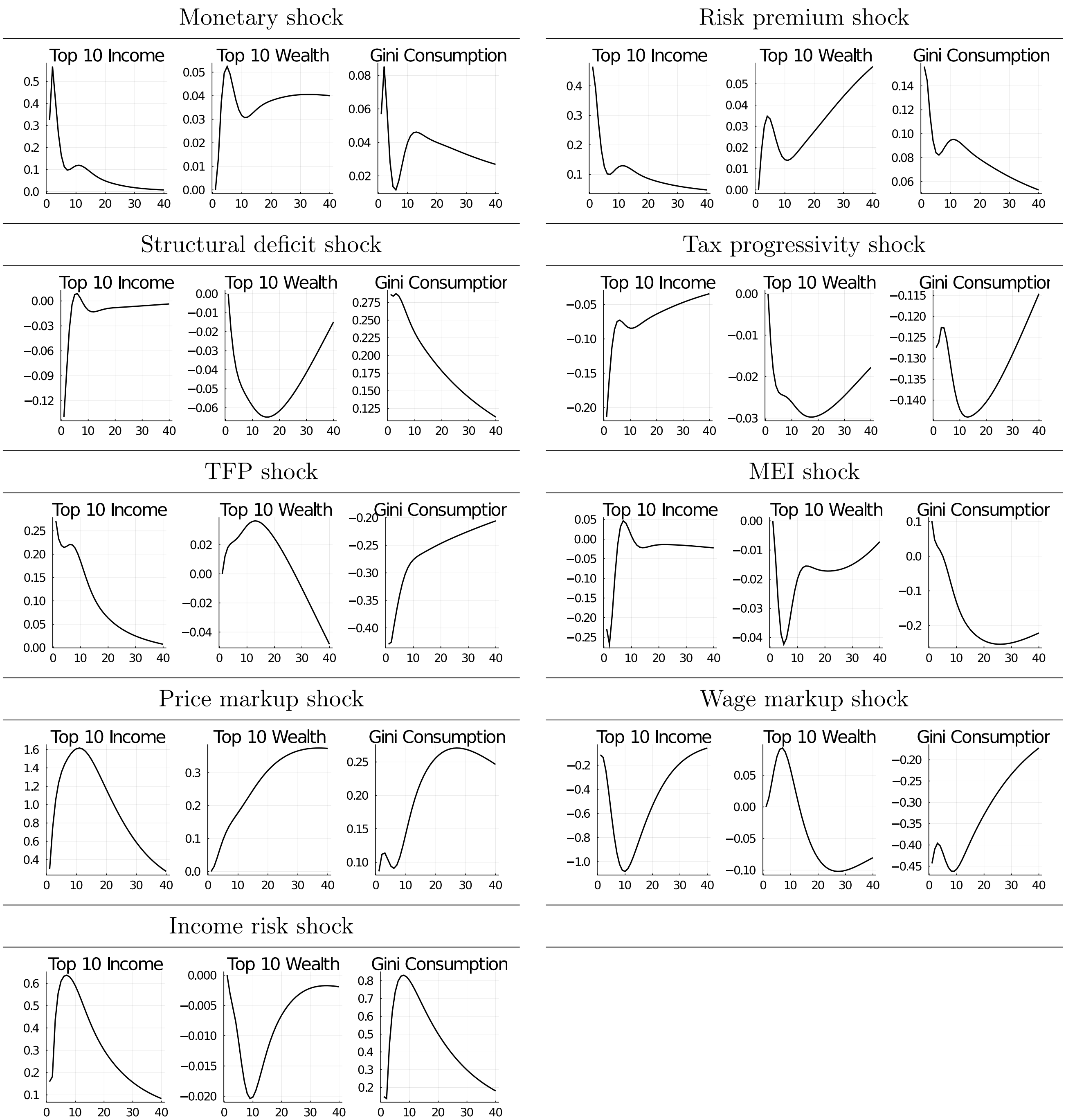
Turning to fiscal policy, increasing the structural deficit lowers income and wealth inequality but increases consumption inequality. On the tax side, a shock to the progressiveness of the tax schedule causes persistently lower income, consumption, and wealth inequality.

Positive technology shocks lead to a persistent decline in consumption inequality. Higher investment-specific technology lowers the price of capital and hence lowers wealth and income inequality. The TFP shock, by contrast, increases income and wealth inequality by raising productivity and therefore the returns to capital held by the wealthy households.

Price and wage markup shocks both have persistent effects on income, consumption, and wealth inequality; however, they differ markedly in the sign of their effects. While price markups increase inequality because they raise profits that go to entrepreneurs, a rise in the target wage markup lowers inequality by increasing wage compression.

Higher income risk leads to a quick increase in income and consumption inequality as poorer households over-proportionally increase their liquid asset holdings for insurance purposes. Wealth inequality falls because poor households react strongly by accumulating extra, mostly liquid, assets.

Figure G.28: Impulse responses of inequality



Notes: The figures display the impulse responses of income, consumption, and wealth inequality in response to the shocks labeled above. Parameter estimates from HANKX+. See main text for further details.