NBER WORKING PAPER SERIES

ON THE EFFECTS OF MONETARY POLICY SHOCKS ON INCOME AND CONSUMPTION HETEROGENEITY

Minsu Chang
Frank Schorfheide

Working Paper 32166
http://www.nber.org/papers/w32166

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2024

We thank Xiaohong Chen, Stephanie Ettmeier, Eva Janssens, Chi Hyun Kim, and participants at various seminars and conferences for helpful suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

On the Effects of Monetary Policy Shocks on Income and Consumption Heterogeneity
Minsu Chang and Frank Schorfheide
NBER Working Paper No. 32166
February 2024
JEL No. C11,C32,C52,E32

ABSTRACT

In this paper we use the functional vector autoregression (VAR) framework of Chang, Chen, and Schorfheide (2024) to study the effects of monetary policy shocks (conventional and informational) on the cross-sectional distribution of U.S. earnings (from the Current Population Survey), consumption, and financial income (both from the Consumer Expenditure Survey). We find that a conventional expansionary monetary policy shock reduces earnings inequality, in large part by lifting individuals out of unemployment. There is a weakly positive effect on consumption inequality and no effect on financial income inequality, but credible bands are wide.

Minsu Chang
Georgetown University
Department of Economics
37th St NW & O St NW, ICC 553
Washington DC 20007
minsu.chang@georgetown.edu

Frank Schorfheide
University of Pennsylvania
Department of Economics
The Ronald O. Perelman Center for
Political Science and Economics
133 South 36th Street
Philadelphia, PA 19104-6297
and CEPR
and also NBER
schorf@ssc.upenn.edu

Replication files will be made available at is available at
https://web.sas.upenn.edu/schorf/working-papers/

1 Introduction

Traditionally, the effects of monetary policy interventions have been studied through the lens of models that abstract from micro-level heterogeneity, such as structural vector autoregressions (VARs) specified in terms of macroeconomic aggregates or as representative agent New Keynesian (RANK) models. However, in view of concerns about rising inequalities in advanced economies in the aftermath of the global financial crisis, there is growing interest in the distributional impacts of conventional and unconventional monetary policies. The contribution of this paper is to apply the functional VAR framework of Chang, Chen, and Schorfheide (2024), henceforth CCS, to study the effects of monetary policy shocks (conventional and informational) on the cross-sectional distribution of earnings, consumption, and financial income.

The textbook effect of an expansionary monetary policy shock is a temporary fall in the real interest rate that stimulates demand and increases aggregate output and nominal prices. Kaplan, Moll, and Violante (2018) emphasize that a decomposition into direct and indirect effects is useful for the analysis of the propagation of monetary policy shocks in the presence of household heterogeneity. The direct effect is generated through the consumption Euler equation: an expansionary monetary policy lowers the real rate and creates an incentive for households to consume in the current period rather than to save for future consumption. This channel is inactive for households that are unable to save and consume all of their income even in the absence of the monetary expansion. Indirect effects are generated through general equilibrium mechanisms that alter the income and wealth distribution. For instance, increased labor demand might raise wages and employment, which boosts consumption. On the other hand, a rising price level may generate income losses for recipients of nominal government transfers. Moreover, to the extent that debt contracts are nominal, inflation shifts wealth from lenders to borrowers.[1] The indirect effect on aggregate consumption crucially depends on the households' idiosyncratic marginal propensity to consume (MPC).

The existing empirical literature has considered two related, but distinct questions. First, what is the effect of a monetary policy shock on the cross-sectional distributions of, say, income and consumption, and inequality measures associated with it? Second, how does income or consumption of particular households or groups of households respond to a monetary policy shock? Answers to the first question provide guidance to central banks that are

[1] This channel is due to Fisher (1933) and its quantitative importance has been recently studied in Doepke and Schneider (2006).

concerned about distributional effects of their actions. Answers to the second question shed light on the shock propagation mechanism. In principle, household- or agent-level responses could be aggregated to derive the response of the cross-sectional distribution, but this requires very accurate measurement of unit-level responses, which often is obstructed by data availability and nonlinearities. Our functional VAR approach is designed to provide direct answers to the first question, which is the focus of this paper.

To identify standard and informational monetary policy shocks we use the instrumental variables (interest rate and stock price surprises) proposed by Jarocinski and Karadi (2020) as internal instruments in a structural VAR. The empirical analysis generates the following findings: First, representing the labor earnings distribution obtained from the Current Population Survey (CPS) as a mixture of a continuous part, capturing positive earnings of employed individuals, and a point mass at zero that corresponds to the unemployed individuals, we find that an expansionary monetary policy shock reduces earning inequality because the unemployment rate falls and individuals with previously no earnings receive positive earnings (employment channel). The estimated effects are broadly consistent with the heterogeneous agent New Keynesian model (HANK) with indivisible labor studied by Ma (2021). If we focus solely on the continuous part of the earnings distribution, then the effect on inequality is small and short-lived. Thus, the employment channel dominates.

Second, we generate impulse response functions (IRFs) for the consumption distribution, obtained from the Consumer Expenditure Survey (CEX). Our analysis captures the indirect effect of rising earnings onto consumption, but also the direct effect of a real-interest change on household consumption. We find that an expansionary policy shock slightly increases consumption inequality measures at the posterior median, but posterior credible bands imply a substantial amount of uncertainty. Third, from the estimation of a functional VAR with CEX financial income data, we conclude that the monetary policy contraction has no significant effect of measures of financial income inequality. An important caveat is that the data set misses households with large financial incomes. Fourth, we compute impulse responses to a (negative) information shock which leads to an increase in earnings inequality, mainly due to a rise in unemployment, and leaves measures of consumption inequality largely unaffected, albeit the posterior median responses of the 90-10 ratio and the Gini coefficient are negative.

The functional approach used in this paper relies on the availability of repeated cross-sections of micro-level observations.[2] In each time period t the cross-sectional distribution is

[2]General treatments of functional data analysis are provided in the books by Bosq (2000), Ramsey and

represented by a log probability density function (pdf), which in turn is approximated by a finite-dimensional linear sieve. We use cubic splines as the basis function. The advantage of using log pdfs is that one does not have to impose non-negativity or monotonicity constraints. The un-normalized log pdfs can be coherently propagated from period to period through a linear law of motion and then be normalized ex post so that they integrate to one. We showed in CCS that the estimation of the functional VAR is straightforward: the sieve coefficients are estimated in each period based on cross-sectional data by maximum likelihood. The coefficient estimates are then included as (noisy) observations in a linear state-space model or a VAR, in combination with the macroeconomic aggregates. The resulting time series model can be estimated with standard techniques. Finally, one can compute IRFs for the sieve coefficients, which can be converted into IRFs for the cross-sectional densities and implied summary statistics, such as percentiles or inequality measures.

The functional analysis has several advantages over seemingly simpler approaches such as the direct inclusion of inequality statistics in a standard VAR, e.g., Coibion, Gorodnichenko, Kueng, and Silvia (2017), Furceri, Loungani, and Zdzienicka (2018), or Guerello (2018). First, it provides a single model from which the dynamics of a large set of distributional statistics (percentiles, Gini coefficient, 90-10 ratio, Theil index) can be derived without generating internal inconsistencies such as quantile crossings in forward simulations of the model. Second, CCS provide simulation evidence that the functional approach leads to tighter credible intervals for the same IRFs than VARs that simply stack percentiles or inequality measures in part because the sieve coefficients efficiently summarize the cross-sectional information, which in turn sharpens inference.

As an alternative to stacking cross-sectional distributions or inequality statistics in a VAR, some authors have used an indirect approach of measuring the effect of monetary policy shocks on cross-sectional distributions and inequality statistics derived from them. Examples are Del Canto, Grigsby, Qian, and Walsh (2023), Lenza and Slacalek (2023), and McKay and Wolf (2023). The indirect approach has a weaker data requirement because it can be implemented with a single cross section. The basic idea is to extract information from the micro data about individual-level income and portfolio compositions or consumption

Silverman (2005), and Horvath and Kokoszka (2012). A fundamental model in the functional time series literature is the functional autoregressive model of order one. An extended version of this model forms the core of the law of motion of the states in our framework. Applications of functional data analysis in macroeconometrics are rare but growing. Examples are Diebold and Li (2006), Chang, Kim, and Park (2016), Hu and Park (2017), Meeks and Monti (2019), Inoue and Rossi (2020), Bjornland, Chang, and Cross (2023).

shares, and then combine this information with IRFs for aggregate wages, hours worked, interest rates, asset returns, consumption, and so forth, to simulate micro-level outcomes forward. These, in turn, can be converted into distributional responses. A disadvantages of this approach is its reliance on the assumption that income and portfolio compositions do not change in response to the shock.

The most demanding data and modeling requirements are associated with panel approaches. For many countries, including the U.S., high-quality panel data are not available at a frequency that is suitable to study fluctuations of inequality measures over the business cycle. However, some countries make administrative data available to researchers. For instance, Holm, Paul, and Tischbirek (2021) use administrative panel data from Norway to estimate panel local projections with observed group heterogeneity (defined in terms of liquid asset distribution).[3] Amberg, Jansson, Klein, and Rogantini Picco (2022) and Andersen, Johannesen, Jorgensen, and Peydro (2021) apply a similar approach to Swedish and Danish administrative data, respectively. The panel approach is well suited to compare the IRFs of different groups of individuals, but it is challenging to aggregate those responses to deduce the response of the cross-sectional distribution, because it is difficult to capture the time series properties (non-linearities due to health and family status changes, job losses, job-to-job transitions, promotions) and the full extent of cross-sectional heterogeneity.

Because of the data and modeling challenges, some researchers have considered pseudo panels as an alternative to actual panels. They can be constructed from rotating panels or repeated cross sections and cross-sectional averaging smoothes out some of the nonlinearities in the unit-level histories. For instance, Cloyne, Ferreira, and Surico (2020) aggregate cross-sectional information into three types of individuals: mortgagors, outright owners, and renters. Anderson, Inoue, and Rossi (2016) track the dynamics of quintiles and Mitman, Broer, and Kramer (2022) track ventiles. Pseudo-panels are well suited to examine how different groups respond to policy shocks. But the approach averages over within-group heterogeneity which makes it difficult to convert results into statements of inequality statistics.

Our empirical analysis contributes to the growing body of evidence about the heterogeneous effects of monetary policy shocks, focusing on the response of cross-sectional distributions and inequality statistics derived from it.[4] Using a different methodological approach,

[3]Inference methods for panel local projections are studied in Almuzara and Sancibrian (2023).

[4]There is a large literature that conducts a quantitative analysis of the distributional effects of monetary policy shocks using calibrated or estimated HANK models. Since our quantitative analysis is more data driven than theory driven we do not provide a comprehensive survey of this literature.

our analysis provides further evidence that expansionary monetary shocks reduce labor earnings inequality, see Coibion, Gorodnichenko, Kueng, and Silvia (2017), Furceri, Loungani, and Zdzienicka (2018), Lenza and Slacalek (2023), Del Canto, Grigsby, Qian, and Walsh (2023), Mitman, Broer, and Kramer (2022). A subset of the studies also emphasizes that the effect is driven by the extensive margin, i.e., workers transitioning out of unemployment, which is consistent with our findings.

With respect to consumption inequality, we obtain fairly wide credible intervals that span positive and negative responses and results in the literature are more mixed. This has, in part, to do with the fact that consumption is less well measured than labor earnings. Coibion, Gorodnichenko, Kueng, and Silvia (2017) find that consumption inequality decreases in response to an expansionary monetary policy shock. Cloyne, Ferreira, and Surico (2020) document in their pseudo-panel analysis that when interest rates fall, households with a mortgage increase their spending considerably, while outright home-owners without mortgage debt do not change their expenditure at all. Using an indirect approach, McKay and Wolf (2023) find that the overall consumption response is quite even in the cross section, but there is heterogeneity in regard to the channels that lead to the consumption response. Holm, Paul, and Tischbirek (2021) show that in their Norwegian administrative data the consumption response has the same U-shape as disposable income with the biggest gains accruing at the bottom and top end of the liquid asset distribution. Andersen, Johannesen, Jorgensen, and Peydro (2021) use a very limited measure of consumption, namely car purchases. The higher the household income, the larger the effect of an expansionary monetary policy shock on car purchases.

The remainder of this paper is organized as follows. Section 2 describes the functional VAR setup and Section 3 discusses the estimation and model selection. These two sections review important material from CCS to make this paper self-contained. The empirical results are presented in Sections 4 to 8. We begin by examining the response of the aggregate variables to a monetary policy shock in an aggregate VAR and the three functional VAR specifications considered subsequently, which include an earnings density, a consumption density, and a financial income density, respectively. In Section 8 we examine responses to an informational shock, and Section 9 concludes. An Online Appendix contains supplemental information on the methodology.

2 A Functional VAR for Cross-Sectional Data

To make this paper self-contained, we provide a summary of the functional VAR framework developed in CCS. The variables in the functional model comprise an $n_y \times 1$ vector of macroeconomic aggregates Y_t and a cross-sectional density $p_t(x)$. Rather than working with the densities directly, we take the logarithmic transformation $\ell_t(x) = \ln p_t(x)$. The advantage of using log densities, instead of density functions, cumulative distribution functions (cdfs), or quantiles is that log densities do not have to satisfy monotonicity or non-negativity restrictions. Thus, they can be easily and coherently propagated using a linear law of motion and then *ex post* normalized to integrate to one in each period.

2.1 Sampling and Measurement

We assume that in every period $t = 1, \ldots, T$ an econometrician observes the macroeconomic aggregates Y_t as well as a sample of N_t draws x_{it}, $i = 1, \ldots, N_t$ from the cross-sectional density $p_t(x)$. In practice, N_t is likely to vary from period to period, but for the subsequent exposition it will be notationally convenient to assume that $N_t = N$ for all t. We collect the time t cross-sectional observations in the vector $X_t = [x_{1t}, \ldots, x_{Nt}]'$. The likelihood function for the functional model is constructed under the assumption that the draws x_{it} are independently and identically distributed (iid) over the cross-section and independent over time conditional on $p_t(x)$. The measurement equation for the cross-section observations takes the form

$$x_{it} \mid p_t(x) \quad \sim \quad \text{iid}\; p_t(x) = \frac{\exp\{\ell_t(x)\}}{\int \exp\{\ell_t(\tilde{x})\} d\tilde{x}}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T. \tag{1}$$

We are working with repeated cross sections or rotating panels and remove a common component from the x_{it}. Thus, the assumption of x_{it} being *iid* across i and t is a reasonable approximation. The functional modeling approach does not require the econometrician to make assumptions about the evolution of x_{it} at the level of an individual, a household, or a firm.

2.2 State Transition

The log density ℓ_t in (1) can be viewed as an infinite-dimensional state variable. We assume that Y_t and ℓ_t evolve according to a joint autoregressive law of motion that we express in terms of deviations from a deterministic component $\big(Y_*, \ell_*(x)\big)$. For notational convenience we assume that the deterministic component is time-invariant and can be interpreted as a steady state. This assumption could be easily relaxed by letting (Y_*, ℓ_*) depend on t. Let

$$Y_t = Y_* + \tilde{Y}_t, \quad \ell_t = \ell_* + \tilde{\ell}_t. \tag{2}$$

The deviations from the deterministic component $\big(Y_t, \ell_t(x)\big)$ evolve jointly according to the following linear functional vector autoregressive (fVAR) law of motion:

$$\begin{aligned} \tilde{Y}_t &= B_{yy}\tilde{Y}_{t-1} + \int B_{yl}(\tilde{x})\tilde{\ell}_{t-1}(\tilde{x})d\tilde{x} + u_{y,t} \\ \tilde{\ell}_t(x) &= B_{ly}(x)\tilde{Y}_{t-1} + \int B_{ll}(x,\tilde{x})\tilde{\ell}_{t-1}(\tilde{x})d\tilde{x} + u_{l,t}(x). \end{aligned} \tag{3}$$

Here $u_{y,t}$ is mean-zero random vector with covariance Ω_{yy} and $u_{l,t}(x)$ is a random element in a Hilbert space with covariance function $\Omega_{ll}(x, \tilde{x})$. We denote the covariance function for $u_{y,t}$ and $u_{l,t}(x)$ by $\Omega_{yl}(x)$. For now, (3) should be interpreted as a reduced-form fVAR in which $u_{y,t}$ and $u_{l,t}(x)$ are one-step-ahead forecast errors. For the empirical analysis below we add more lags to the system. (3) can be viewed as the state-transition equation in a functional state-space model.

2.3 Three Simplifications

Equations (1), (2), and (3) define an infinite-dimensional nonlinear state-space model for the observables $\{Y_t, X_t\}_{t=1}^T$. Unfortunately, the estimation of this model is not practical, and we will simplify it in three steps. First, we replace the infinite-dimensional objects by finite-dimensional objects. Second, we turn the nonlinear state-space model into a linear state-space model. Third, we let the measurement error variance tend to zero.

A Finite-Dimensional Nonlinear State-Space Model. We replace $\ell_t(x)$ by a collection of finite-dimensional representations, indexed by the superscript (K). Let

$$\ell_t^{(K)}(x) = \sum_{k=1}^K \alpha_{k,t}\zeta_k(x) = \big[\zeta_1(x), \dots, \zeta_K(x)\big] \cdot \begin{bmatrix} \alpha_{1,t} \\ \vdots \\ \alpha_{K,t} \end{bmatrix} = \zeta'(x)\alpha_t \tag{4}$$

and $\ell_*^{(K)}(x) = \zeta'(x)\alpha_*$. Here $\zeta_1(x), \zeta_2(x), \dots$ is a sequence of basis functions. We dropped the (K) superscripts from the vectors $\zeta(x)$, α_t, and α_* to simplify the notation. We define

$\tilde{\alpha}_t = \alpha_t - \alpha_*$ such that $\tilde{\ell}^{(K)}(x) = \ell_t^{(K)}(x) - \ell_*^{(K)}(x)$. We can now write a K'th order representation of the density of X_t as follows:

$$\begin{aligned} p^{(K)}\left(X_t|\alpha_t\right) &= \exp\left\{N\mathcal{L}^{(K)}(\alpha_t|X_t)\right\}, \\ \mathcal{L}^{(K)}(\alpha_t|X_t) &= \left(\frac{1}{N}\sum_{i=1}^{N}\zeta(x_{it})'\right)\alpha_t - \ln\int \exp\left\{\zeta'(x)\alpha_t\right\}dx. \end{aligned} \tag{5}$$

We represent the kernels $B_{ll}(x, \tilde{x})$ and $B_{yl}(\tilde{x})$, the function $B_{ly}(x)$, and the functional innovation $u_{l,t}(x)$ that appear in the state-transition equation (3) as follows:

$$\begin{aligned} B_{ll}^{(K)}(x, \tilde{x}) &= \zeta'(x)B_{ll}\xi(\tilde{x}), \quad B_{yl}^{(K)}(x) = B_{yl}\xi(\tilde{x}) \\ B_{ly}^{(K)}(x) &= \zeta'(x)B_{ly}, \quad u_{l,t}^{(K)}(x) = \zeta'(x)u_{\alpha,t}, \end{aligned} \tag{6}$$

where $\xi(x)$ is a second $K \times 1$ vector of basis functions and $u_{\alpha,t}$ is a $K \times 1$ vector of innovations. The matrix B_{ll} is of dimension $K \times K$, B_{yl} is of dimension $n_y \times K$, and B_{ly} is of dimension $K \times n_y$. Combining (2), (3), and (6) yields the following vector autoregressive system for the macroeconomic aggregates and the sieve coefficients (omitting K superscripts):

$$\begin{bmatrix} Y_t - Y_* \\ \alpha_t - \alpha_* \end{bmatrix} = \begin{bmatrix} B_{yy} & B_{yl}C_\alpha \\ B_{ly} & B_{ll}C_\alpha \end{bmatrix} \begin{bmatrix} Y_{t-1} - Y_* \\ \alpha_{t-1} - \alpha_* \end{bmatrix} + \begin{bmatrix} u_{y,t} \\ u_{\alpha,t} \end{bmatrix}, \tag{7}$$

where $C_\alpha = \int \xi(\tilde{x})\zeta'(\tilde{x})d\tilde{x}$. Let $u_t' = [u_{y,t}', u_{\alpha,t}']$. We subsequently assume that the innovations are Gaussian:

$$u_t \sim \mathcal{N}(0, \Sigma). \tag{8}$$

The finite-dimensional state-space representation is given by the measurement equation (5) and the state-transition equation (7). To obtain a more compact notation, we define $\widetilde{W}_t = [(Y_t - Y_*)', \alpha_t']'$ and absorb the matrix C_α into a general regression coefficient matrix Φ, which leads to

$$\widetilde{W}_t = \Phi_1\widetilde{W}_{t-1} + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma). \tag{9}$$

A Finite-Dimensional Linear State-Space Model. To avoid the use of a nonlinear filter for the evaluation of the likelihood function of the state-space model, one can "linearize" the measurement equation by taking a second-order Taylor series approximation of $\ln p^{(K)}\left(X_t|\alpha_t\right)$ in (5) around the maximum likelihood estimator (MLE) $\hat{\alpha}_t$. This approximation can be written as a linear Gaussian measurement equation:

$$\hat{\alpha}_t = \alpha_t + N^{-1/2}\eta_t, \quad \eta_t \sim \mathcal{N}\big(0, \hat{V}_t^{-1}\big), \tag{10}$$

where $\hat{V}_t$ is the negative inverse Hessian associated with the log likelihood function evaluated at the MLE. Note that the observations X_t enter the measurement equation indirectly through the MLE $\hat{\alpha}_t$.

A Finite-Dimensional VAR. If N is large relative to K, then the measurement error $N^{-1/2}\eta_t$ is close to zero and $\alpha_t \approx \hat{\alpha}_t$. Thus, one can replace and for the empirical analysis we simply replace α_t in (7) by $\hat{\alpha}_t$ and estimate a VAR in the macroeconomic variables and the estimated sieve coefficients. The estimation can be conveniently implemented in two steps:

1. For each period $t = 1, \ldots, T$ estimate the log-spline density model for X_t by maximizing the log likelihood function in (5). This leads to the sequence $\hat{\alpha}_t$.

2. Estimate a version of the VAR in (9), replacing the "true" sieve coefficients α_t in the definition of $\widetilde{W}_t$ by $\hat{\alpha}_t$.

CCS provide rates at which (N, T, K) are allowed to tend to infinity to ensure that the likelihood functions of the three finite-dimensional models are asymptotically equivalent. In this paper, we are considering an application in which the cross-sectional dimension N is large and we will work with the finite-dimensional VAR approximation.

3 Estimation and Model Selection

For the empirical analysis we add additional lags to the VAR in (9) and we write the model with an intercept rather than in terms of deviations from a mean (or deterministic trend), as it is common in the VAR literature that includes variables in log levels in the vector Y_t. Let $W_t = [Y_t', \alpha_t']'$, or, under the third simplification in Section 2.3 used in the empirical analysis below, $W_t = [Y_t', \hat{\alpha}_t']'$. The reduced-form specification of our VAR takes the form

$$W_t = \sum_{j=1}^{p} \Phi_j W_{t-j} + \Phi_0 + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma). \tag{11}$$

3.1 From Reduced-Form to Structural VAR

To identify the effects of monetary policy shocks we include instruments for these shocks in the definition of Y_t (and hence W_t), assuming that the instruments are ordered first. We partition $W_t = [W_{1t}', W_{2t}']'$, where the $n_1 \times 1$ vector W_{1t} contains the instruments. In our application n_1 equals either one (standard monetary policy shock) or two (standard shock and informational shock). We denote the dimension of W_{2t} by n_2. Similarly, we partition the vector of reduced-form forecast errors $u_t = [u'_{1t}, u'_{2t}]'$.

We now express the forecast errors as functions of structural innovations. First, we let

$$u_{1t} = \Phi^{\epsilon}_{11}\epsilon_{1t}, \quad \epsilon_{1t} \sim \mathcal{N}(0, I_{n_1}) \tag{12}$$

with the restriction that $\Phi^{\epsilon}_{11}\Phi^{\epsilon'}_{11} = \Sigma_{11}$, where Σ_{11} is the partition of Σ that corresponds to u_{1t}. Here ϵ_{1t} are the orthogonalized innovations of the instruments W_{1t}. Second, we assume that u_{2t} is driven by the policy shocks of interest, denoted by $\tilde{\epsilon}_{2.1,t}$, and non-policy shocks $\epsilon_{2.2,t}$. The tilde indicates that the policy shocks are (not yet) normalized. Thus,

$$u_{2t} = \Phi^{\tilde{\epsilon}}_{22.1}\tilde{\epsilon}_{2.1,t} + \Phi^{\epsilon}_{22.2}\epsilon_{2.2,t}, \quad \epsilon_{2.2,t} \sim \mathcal{N}(0, I_{n_2-n_1}). \tag{13}$$

Here $\tilde{\epsilon}_{2.1,t}$ is of dimension $n_1 \times 1$ and $\epsilon_{2.2,t}$ is of dimension $(n_2 - n_1) \times 1$. The object of interest is $\Phi^{\tilde{\epsilon}}_{22.1}$ which is the effect of the policy shocks on u_{2t}, and hence the impact effect on W_{2t}.

The key assumption is that the policy shocks can be expressed as a function of the instrument innovations and an orthogonal component:

$$\tilde{\epsilon}_{2.1,t} = \Gamma_1\epsilon_{1t} + \Gamma_{2.1}\epsilon_{2.1,t}, \quad \epsilon_{2.1,t} \sim \mathcal{N}(0, I_{n_1}), \quad \mathbb{E}[\epsilon_{1t}\epsilon'_{2.1,t}] = 0. \tag{14}$$

This equation resembles the equation $x_i = \gamma z_i + \epsilon_i$ in a linear instrumental variable model with one endogenous regressor, x_i, and one instrument, z_i. We assume that Γ_1 is diagonal. Thus, each element of the instrument innovation vector ϵ_{1t} is connected to one single element of the policy shock vector $\tilde{\epsilon}_{2.1,t}$. The instrument relevance condition corresponds to $\Gamma_1 \neq 0$ and the instrument validity condition is that $\epsilon_{1,t}$ is uncorrelated with the non-policy shocks $\epsilon_{2.2,t}$. To ensure that policy shocks $\tilde{\epsilon}_{2.1,t}$ and non-policy shocks $\epsilon_{2.2,t}$ are uncorrelated, we impose that $\epsilon_{2.1,t}$ and $\epsilon_{2.2,t}$ are uncorrelated. Let $\epsilon_t = [\epsilon'_{1t}, \epsilon'_{2.1,t}, \epsilon'_{2.2,t}]$. We deduce that $\epsilon_t \sim \mathcal{N}(0, I_{n_1+n_2})$.

By combining (13) and (14) we obtain

$$u_{2t} = \Phi^{\tilde{\epsilon}}_{22.1}\Gamma_1\epsilon_{1t} + \Phi^{\tilde{\epsilon}}_{22.1}\Gamma_{2.1}\epsilon_{2.1,t} + \Phi^{\epsilon}_{22.2}\epsilon_{2.2,t}, \tag{15}$$

which implies that the response of W_{2t} to the instrument innovation ϵ_{1t} measures the effect of the policy shock up to the scale normalization Γ_1. Equations (12) and (15) lead to a block-triangular system that determines u_t as a function of ϵ_t. We can write

$$u_t = \Phi^{\epsilon}\epsilon_t, \quad \Phi_{\epsilon} = \begin{bmatrix} \Phi^{\epsilon}_{11} & 0_{n_1 \times n_2} \\ \Phi^{\tilde{\epsilon}}_{22.1}\Gamma_1 & [\Phi^{\tilde{\epsilon}}_{22.1}\Gamma_{2.1} \; \Phi^{\epsilon}_{22.2}] \end{bmatrix}. \tag{16}$$

Let Σ_{tr} be the lower-triangular Cholesky factor of the covariance matrix Σ and Ω be an orthogonal matrix, which has the property that $\Omega\Omega' = \Omega'\Omega = I$. Now factorize $\Phi^{\epsilon} = \Sigma_{tr}\Omega$ which ensures that $\mathbb{E}[u_t u_t'] = \Phi^{\epsilon}\Phi^{\epsilon'} = \Sigma$. Partion Σ and Ω such that Σ_{ij} and Ω_{it} conform with the partions of u_t. The block diagonal structure of (16) implies that $\Omega_{12} = 0$. In turn, Ω_{22} has to be full rank and its columns span $\mathbb{R}^{n_2}$. Thus, the only vector satisfying $\lambda'\Omega_{22} = 0$ is $\lambda = 0$. We deduce that the columns of $[\Omega_{11}' \; \Omega_{21}']'$ can only be orthogonal to the columns of $[0_{n_2 \times n_1} \; \Omega_{22}']'$ if $\Omega_{21} = 0$. Overall, this leads to

$$u_t = \Sigma_{tr}\Omega\epsilon_t, \quad \Omega = \begin{bmatrix} \Omega_{11} & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & \Omega_{22} \end{bmatrix}. \tag{17}$$

Because we are only interested in the response to ϵ_{1t}, the value of Ω_{22} is irrelevant for the subsequent analysis. This implementation of VAR shock identification through instrumental variables has been used, for instance, in Anderson, Inoue, and Rossi (2016) for fiscal policy shocks and Jarocinski and Karadi (2020) for monetary policy shocks. A more detailed theoretical analysis is provided in Plagborg-Møller and Wolf (2021).[5]

3.2 Bayesian Estimation

Because the vector α_t, and hence W_t, can potentially be large, we use the VAR parameterization and prior proposed by Chan (2022). His specification is suitable for high-dimensional settings because it leads to equation-by-equation estimation while allowing for some asymmetry of the prior across equation.

Likelihood Function. The structural VAR by (11) and (17) can be rewritten as follows:

$$AW_t = \sum_{j=1}^{p} B_j W_{t-j} + B_0 + \eta_t, \quad \eta_t = D^{1/2}\Omega\epsilon_t, \tag{18}$$

where D is a diagonal matrix with diagonal elements D_i and A is a lower-triangular matrix with ones on the diagonal. Multiplying both sides of the equality by A^{-1} we deduce that $A^{-1}D^{1/2} = \Sigma_{tr}$ and $A^{-1}B_j = \Phi_j$, $j = 0.\ldots,p$. Note that $\eta_t \sim \mathcal{N}(0, D)$.

Using the lower-triangular structure of the A matrix, define the $(i-1) \times 1$ vectors

$$\mathcal{A}_i = [A_{i,1}, \ldots, A_{i,i-1}], \quad \tilde{W}_{<i,t} = -[W_{1,t}, \ldots, W_{i-1,t}]', i = 2, \ldots, n.$$

[5] Alternative implementations of Bayesian estimation of structural VARs identified by instruments are discussed in Caldara and Herbst (2019) and Arias, Rubio-Ramirez, and Waggoner (2022).

Moreover, let $k_i = k + i - 1$ and define the $k_i \times 1$ vectors

$$Z_{it} = \left[\tilde{W}'_{<i,t}, W'_{t-1}, \ldots, W'_{t-p}, 1\right], \quad \beta_i = \left[\mathcal{A}'_i, B'_{i\cdot,1}, \ldots, B'_{i\cdot,p}, B'_{i\cdot,0}\right]',$$

where $B_{i\cdot,j}$ is the ith row of the matrix B_j. Finally, define W_i to be the $T \times 1$ vector with elements W_{it}, Z_i the $T \times k_i$ matrix with rows Z'_{it}, and η_i the $T \times 1$ vector with elements η_{it}. Then we can write the ith equation in matrix form as

$$W_i = Z_i \beta_i + \eta_i. \tag{19}$$

Because A is lower-triangular with ones on the diagonal, the Jacobian associated with the change-of-variables from η_t to W_t in (18) is equal to one. In turn, the likelihood function for the system is the product of the likelihood functions for each variable i. Let $\beta = (\beta_1, \ldots, \beta_n)$. Then:

$$p(W|\beta, D) \propto \prod_{i=1}^{n} |D_i|^{-1/2} \exp\left\{-\frac{1}{2D_i}(W_i - Z_i\beta_i)'(W_i - Z_i\beta_i)\right\}. \tag{20}$$

As always in structural VAR settings, the rotation matrix Ω does not enter the likelihood function.

Prior Distribution. Chan (2022) proposes a prior distribution that assumes that parameters are independent across equations. This implies that the model can be estimated equation-by-equation, speeding up the Bayesian computations in high-dimensional settings considerably. The prior takes the form

$$p(\beta, D|\lambda) = \prod_{i=1}^{n} p(\beta_i|D_i, \lambda) p(D_i|\lambda), \tag{21}$$

where λ is a vector of hyperparameters. For each pair (β_i, D_i) we use a Normal-Inverse Gamma (NIG) distribution of the form

$$\beta_i|(D_i, \lambda) \sim \mathcal{N}\big(\underline{\beta}_i, D_i \underline{V}_i^{\beta}\big), \quad D_i|\lambda \sim IG\big(\underline{\nu}_i, \underline{S}_i\big). \tag{22}$$

The prior distribution loosely follows that of a Minnesota prior. The mean vectors $\underline{\beta}_i$ are chosen such that the implied prior for the reduced form parameters Φ is centered at univariate random-walk specifications for those elements of Y_t that correspond to log levels of macroeconomic aggregates. For other elements of Y_t and the elements of the α_t vector the priors are centered at zero. We introduce a hyperparameter λ_1 that controls the overall prior precision $[\underline{V}_i^{\beta}]^{-1}$ and a hyperparameter λ_2 that controls the relative precision of coefficients that capture the effect of lagged α_ts on current Y_ts. As $\lambda_2 \longrightarrow \infty$, the cross-sectional density does not Granger-cause the aggregate variables and the system becomes block-triangular. Further details on the specification of $\underline{\beta}_i$, $\underline{V}_i^{\beta}$, $\underline{\nu}_i$, and $\underline{S}_i$, the posterior distribution, and the MDD are provided in the Online Appendix.[6]

If $n_1 = 1$ then Ω_{11} is simply equal to one. For the case $n_1 = 2$, we impose sign restrictions on the impulse response functions, described in Section 4, to set-identify the responses of W_{1t} to ϵ_{1t}. We start from a prior for Ω_{11} that is uniform on the space of orthogonal matrices and then truncate the prior so that the sign restrictions are satisfied. We denote the resulting prior by $p(\Omega_{11}|\beta, D)$.

Posterior Sampling. The conjugate form of the prior implies that the posterior distribution of (β, D) also belongs to the NIG family. Thus, we can generate posterior draws of (β, D) by direct sampling and because both likelihood and prior factorize in terms of (β_i, D_i), $i = 1, \ldots, n$ we can sample the parameters for each equation separately. Because Ω_{11} does not enter the likelihood function, this prior does not get updated and the posterior equals the prior.

For each draw $(\beta^s, D^s, \Omega_{11}^s)$, $s = 1, \ldots, N_{sim}$ we iterate (18) forward to obtain a draw from the posterior distribution of the $W_t = [Y_t', \alpha_t]'$ IRFs. The α_t IRFs are converted into density IRFs using

$$p^{(K)}(x|\alpha_t) = \frac{\exp\{\zeta'(x)\alpha_t\}}{\int \exp\{\zeta'(\tilde{x})\alpha_t\}\, d\tilde{x}}. \tag{23}$$

Based on the draws from the IRF posterior we compute summary statistics that are then plotted in the figures.

Model Selection. Our model for $(Y_{1:T}, X_{1:T})$ depends on the sieve approximation order K, the number of lags p in the VAR (11), and the hyperparameter vector λ for the prior distribution in (21). We select these model features by maximizing the the Bayesian MDD over grid of candidate values. It is important to recognize that the MDD cannot be directly calculated based on the likelihood function for W_t in the reduced-form VAR specification (11) because the definition of W_t depends on the dimensionality k through the sieve coefficient vector a_t. Formally, the MDD for the underlying finite-dimensional nonlinear state-space model is defined as

$$p^{(K)}(Y_{1:T}, X_{1:T}|\lambda) = \int \left(\prod_{t=1}^{T} p^{(K)}(Y_t, X_t|Y_{1:t-1}, X_{1:t-1}, \theta) \right) p^{(K)}(\theta|\lambda) d\theta, \tag{24}$$

where θ is the generic parameter vector of the state-space model.

[6] As in Jarocinski and Karadi (2020), we modify the VAR in (11) to impose that $\Phi_{1\cdot} = 0$.

CCS show that once the measurement error variance in (10) is set to zero, one obtains the MDD approximation

$$\begin{aligned} & p_*^{(K)}(Y_{1:T}, X_{1:T}|\lambda) \\ &= \int \left(\prod_{t=1}^{T} p_G^{(K)}(Y_t, \alpha_t = \hat{\alpha}_t | Y_{t-p:t-1}, \alpha_{t-p:t-1} = \hat{\alpha}_{t-p:t-1}, \theta) \right) p^{(K)}(\theta|\lambda) d\theta \\ &\quad \times \left(\prod_{t=1}^{T} p_{pen}^{(K)}(X_t|\hat{\alpha}_t) \right) \\ &= \left[\int (I \cdot II)\, d\theta \right] \cdot III, \end{aligned} \tag{25}$$

say. Here $\theta = (\beta, D)$, term I is the likelihood function (20) associated with the $W_t = [Y_t', \hat{\alpha}_t']'$ VAR, term II is the prior in (21). The integral $\left[\int I \cdot II d\theta\right]$ is the standard MDD associated with the W_t VAR. Under the conjugate prior that we are using, this expression can be evaluated analytically.

Term III turns the MDD for $W_{1:T} = (Y_{1:T}, \hat{\alpha}_{1:T})$ into a MDD for $(Y_{1:T}, X_{1:T})$. The expression $p_{pen}^{(K)}(X_t|\hat{\alpha}_t)$ that appears in term III is the Laplace approximation

$$p_{pen}^{(K)}(X_t|\hat{\alpha}_t) = \exp\left\{N\mathcal{L}^{(K)}(\hat{\alpha}_t|X_t)\right\} \left(\frac{2\pi}{N}\right)^{K/2} |\hat{V}_t|^{1/2}$$

of the MDD associated with the likelihood function for the cross-sectional observations X_t, ignoring the dynamic aspects of the model. The first part, $\exp\left\{N\mathcal{L}^{(K)}(\hat{\alpha}_t|X_t)\right\}$, is the maximized likelihood function that is non-decreasing in K. Note that the dimension of $\hat{V}_t$ depends on K. The second part, $(2\pi/N)^{K/2}|\hat{V}_t|^{1/2}$, is a penalty for the dimensionality K, and hence avoids overfitting.

3.3 Further Implementation Details

As explained in detail in CCS, a few additional steps are required for the implementation. First, we conduct a preliminary data transformation of the form $x = g(z|\vartheta)$, where z are the raw data and x are the transformed data. We explain in each of the empirical sections, which transformation we apply. Second, we need to choose basis functions $\zeta_1(x), \ldots, \zeta_K(x)$. We use a cubic spline basis and provide further details below. Third, some of the the cross-sectional data are top-coded, which means that the likelihood function in (5) needs to be adjusted accordingly. This adjustment is described in detail in the Online Appendix. Forth, the cross-sectional data are typically not seasonally adjusted, whereas the aggregate data are. If empirically necessary, we project the $\hat{\alpha}_t$ estimates on seasonal dummies which can be interpreted as replacing α_* in (7) by $\alpha_{*,t}$.

Fifth, there might be linear dependencies in the estimated $\hat{\alpha}_t$ vectors. Thus, after subtracting estimates of α_* (or $\alpha_{*,t}$ in case of the seasonal adjustment), we use principal components analysis to compress $\hat{\alpha}_t$, $t = 1, \ldots, T$ into the lower-dimensional vectors $\hat{a}_t$, respectively. This leads to $(\alpha_t - \alpha_{*,t})' = a_t'\Lambda$. We replace $\hat{\alpha}_t$ by $\hat{a}_t$ in the definition of W_t. We treat Λ, which is constructed from the $\hat{\alpha}_t$s, as fixed and use the relationship to adjust the affected model equations. If there are no exact linear dependencies, this step does not lead to a dimensionality reduction but it normalizes the log density coefficients to have unit variance. Sixth, before plotting cross-sectional densities, we undo the change of variables $x = g(z)$ and report densities for the original and not the transformed cross-sectional data.

In addition to estimating functional VARs that include the cross-sectional data through $\hat{\alpha}_t$ in the definition of W_t, we also estimate VARs for Y_t, excluding the cross-sectional data. These VARs are obtained by simply dropping $\hat{\alpha}_t$ from the W_t vector. The structure of the prior is left unchanged.

4 Shock Instruments and Aggregate Responses

We are using two types of data in the Y_t vector: high-frequency instruments for monetary policy shocks and macroeconomic time series. The Y_t data will be combined with three types of cross-sectional data: earnings data in Section 5, consumption data in Section 6, and financial income data in Section 7. In addition to providing details on the series included in Y_t, in this section we also report the impulse response functions (IRFs) of the aggregate variables to a monetary policy shock, based on the various functional VAR specifications studied subsequently.

High-frequency Instruments. The empirical analysis in the main part of the paper is based on instrumental variables taken from Jarocinski and Karadi (2020), who consider two surprise variables that allow them to separate unanticipated changes in monetary policy (monetary policy shocks) from the central bank's revelation of information about the state of the economy that is conveyed through interest rates (information shocks). The variables are surprises in the three-month fed funds futures ($ff4\_hf$) and surprises in the S&P 500 stock market index ($sp500\_hf$). Sign restrictions are used to separate the two shocks of interest. It is assumed that a contractionary monetary policy shock generates an interest rate increase and a drop in stock prices, whereas a positive information shock is associated with an increase in both interest rates and stock prices.[7]

Aggregate Variables. Following Jarocinski and Karadi (2020), all of our empirical models include the following monthly macroeconomic variables: the monthly average of the one-year constant-maturity Treasury yield serves as the monetary policy indicator. The advantage of using a one-year rate is that it remains a valid measure of monetary policy stance also when the federal funds rate is constrained by the zero lower bound (ZLB). The monthly average of the S&P 500 stock price index in log levels. Real GDP and GDP deflator in log levels interpolated to monthly frequency based on Stock and Watson (2010). The excess bond premium (EBP) as indicator of financial conditions. In addition, we include an unemployment rate constructed from micro data. For the functional VARs with cross-sectional consumption data we include aggregate per capita consumption from the National Income and Product Accounts (NIPA). For the functional VAR with the financial income distribution we add the fraction of units with zero financial income.

Sample Periods. The functional VAR with earnings data in Section 5 is estimated based on monthly data, using the sample period 1990:M2 to 2016:M12. This sample has one missing value (the financial market disruption after the 9/11 terrorist attack in 2001:M9), which is replaced with zero. Because of cross-sectional data availability, the estimation results reported in Sections 6 and 7 are based on quarterly data from 1990:Q2 to 2016:Q4. The quarterly aggregate data are generated by time-aggregating the monthly series.

Model Selection. The lag length p, the hyperparameters λ, and the approximation order K for the functional VARs, are chosen to maximize the MDD in (25). Table 1 contains some information about model selection for the empirical analysis. Each row corresponds to a different sieve approximation order K. We report the optimal choice of λ_1, λ_2, and p conditional on K.[8] The last column shows MDD differentials relative to $K = 4$. For example, the preferred specification for the earnings fVAR is $K = 10$ with a lag order of $\hat{p} = 1$. For smaller values of K the selected lag order is larger, because the dimension of W_{t-j} is smaller. For instance, for $K = 4$ with $\hat{p} = 4$ there are $4 \cdot (6 + 4) + 1 = 41$ regressors. For $K = 10$ with $\hat{p} = 1$ there are $1 \cdot (6 + 10) + 1 = 17$ regressors. This number would increase to $4 \cdot (6 + 10) + 1 = 64$ regressors in a four-lag specification. Thus generally, the larger

[7] As a robustness exercise we also considered the instruments from Nakamura and Steinsson (2018). The empirical results turned out to be very similar to the ones reported in this paper.

[8] We use an equally-spaced grid with 31 points for $\ln \lambda_j$, ranging from -10 to 20. We consider lag lengths $p \in \{1, 2, 3, 4\}$.

Table 1: Hyperparameter Estimates and Log MDD Differentials

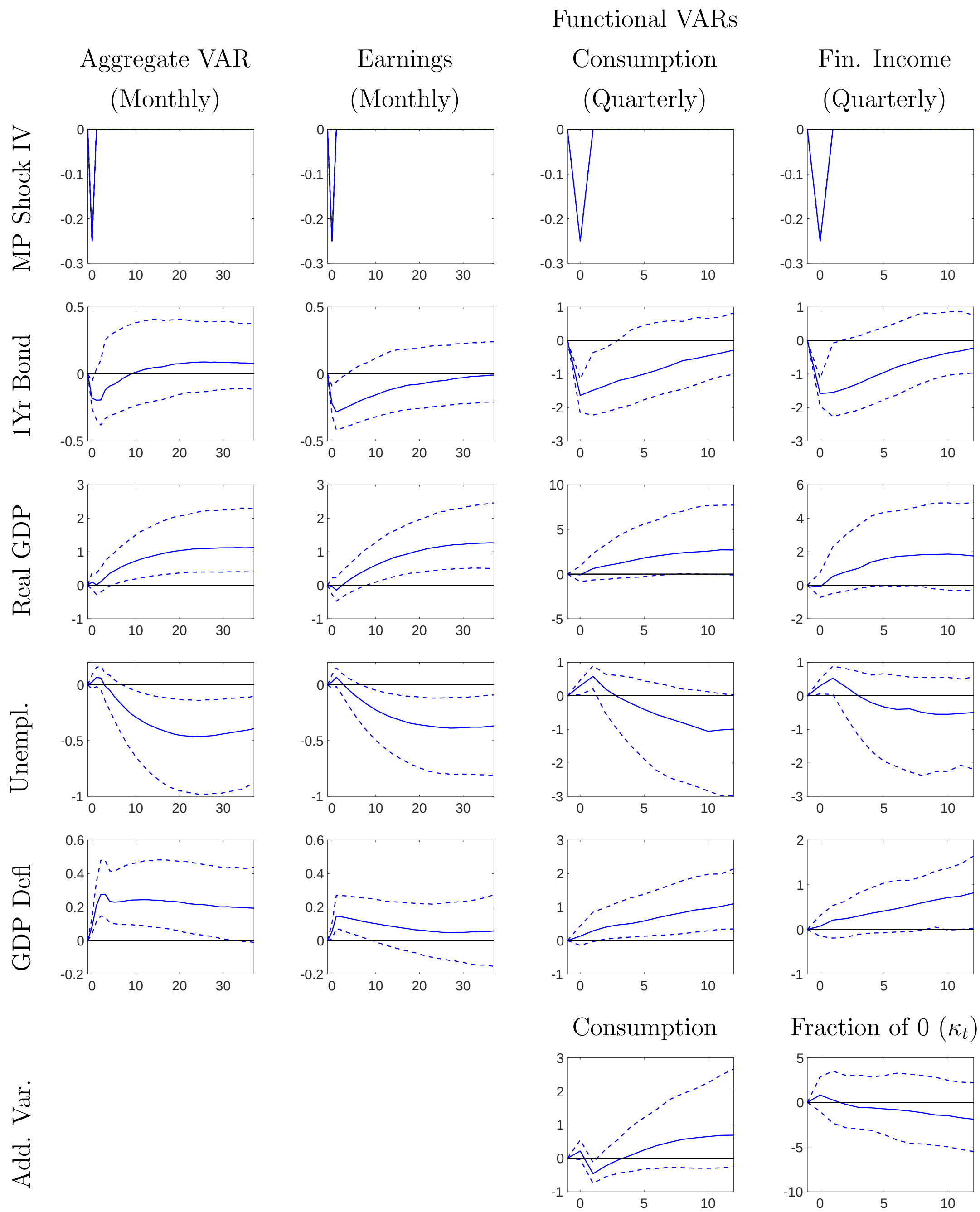
Specification	K	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{p}$	MDD
Earnings fVAR	4	403.43	0.37	4	0
	6	403.43	0.37	2	25,073
	8	1096.63	0.14	3	27,625
	10	*54.60*	*54.60*	*1*	*27,893*
Consumption fVAR	4	403.43	0.14	1	0
	6	*403.43*	*0.37*	*1*	*551.37*
	8	148.41	1.00	1	362.63
	10	54.60	20.09	1	97.41
Financial Inc. fVAR	4	403.43	1.00	1	0
	6	403.43	1.00	1	299.33
	8	*403.43*	*1.00*	*1*	*511.84*
	10	403.43	1.00	1	511.84
Aggregate VAR		148.41		4	

Notes: The log MDD differentials for each model specification are computed with respect to $K = 4, \lambda_1 = \hat{\lambda}_1, \lambda_2 = \hat{\lambda}_2, p = \hat{p}$. For each K we maximized the MDD with respect to λ and p to obtain $\hat{\lambda}_j(K)$ and $\hat{p}(K)$. The $K = 8$ and $K = 10$ entries for the financial income fVAR are identical because a compression step described in the Online Appendix equalizes the effective dimension of the specifications.

the regressor space, the higher the selected prior precision tends to be. The cross-sectional distribution of consumption is smoother than that of earnings. Hence, $K = 6$ suffices and the MDD selects $\hat{p} = 1$ one lag. The consumption and financial income fVARs contain 7 macroeconomic variables and the selected lag order is always $\hat{p} = 1$. For financial income $K = 8$ and $K = 10$ deliver the same MDD because there is a perfect collinearity between elements of the $\hat{\alpha}_t$ sequence that we eliminate. This reduces the effective dimension of the approximation and leads to identical MDDs.

Response of Aggregate Variables to a Monetary Policy Shock. Figure 1 depicts impulse responses of the aggregate variables to an unanticipated monetary policy shock. We show results for four empirical models: a VAR that only includes aggregate variables, and three functional VAR specifications that include cross-sectional information on labor earnings, consumption, and financial income, respectively. The responses are normalized such that the surprise reduction in the three-month federal funds rate is 25 basis points (bp). In the first two columns the time period for the IRFs is a month. In columns three and four the period is one quarter.

Figure 1: Responses of Aggregate Variables to Monetary Policy Shock



Notes: Responses to a 25 basis point monetary policy shock based on Jarocinski and Karadi (2020). The Aggregate VAR uses aggregate variables only. In addition, the functional VARs use cross-sectional data on earnings, consumption, or financial income. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. The plots depict 10th (dashed), 50th (solid), and 90th (dashed) percentiles of the posterior. GDP defl. and real GDP responses are percentage deviations from the steady state, whereas the other responses are absolute percentages. x-axis horizon is months in columns 1 and 2 and quarters in columns 3 and 4.

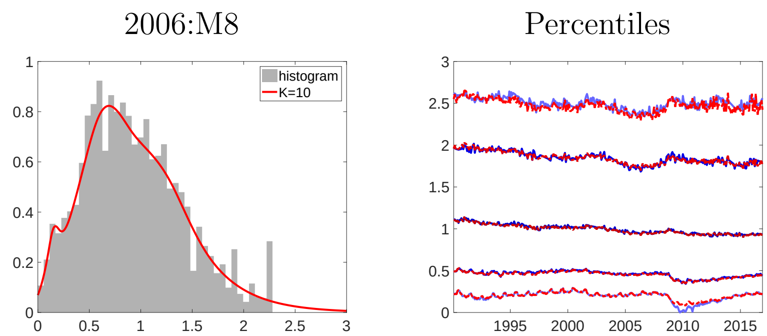
The responses in the first two columns are very similar to each other and also quantitatively similar to the responses reported in Figure 2 of Jarocinski and Karadi (2020) after the latter are scaled by minus five. The presence of the cross-sectional variables has no substantial effect on the impulse response inference. The main difference is that the GDP deflator response in the functional VAR (earnings) is slightly weaker. The width of the bands depends on the number of lags and the hyperparameters selected by the MDD. Recall from Table 1 that the results for the aggregate VAR are based on four lags, whereas the functional VAR (earnings) only uses one lag. At the posterior median, the one-year bond rate moves approximately one-for-one with the federal funds rate surprise and then reverts back to steady state. The expansionary monetary policy leads to an increase of real GDP by about 1.2% and a reduction of the unemployment rate by 0.3 percentages after three years at the posterior median. Prices increase by about 0.25% upon impact in the aggregate VAR and 0.15% in the functional VAR and slightly fall subsequently.

The responses in columns 3 and 4 of Figure 1 are based on functional VARs estimated on quarterly data. Therefore, the x-axis now refers to quarters. Qualitatively, the aggregate responses from the quarterly models are similar to the responses from the monthly models. Quantitatively, there is a difference in magnitude. The reason is that a 25 basis point surprise at monthly frequency is only roughly a third of a 25 basis point surprise over an entire quarter. In calculations not reported in the paper we time aggregated the responses from the monthly earnings VAR to quarterly frequency and rescaled the IRFs such that the instrument moves by 25 basis points over the quarter. After this re-scaling, the magnitude of the aggregate responses from the VARs estimated with monthly and quarterly data, respectively, were indeed very similar. The last row of the figure contains responses for fVAR-specific variables, namely, consumption per capita and the fraction of households with zero financial income, respectively.

5 Earnings Response to a Monetary Policy Shock

The first part of the empirical analysis examines the response of labor earnings to a monetary policy shock. The micro data are discussed in Section 5.1, the IRF results are presented in Section 5.2, and Section 5.3 relates the empirical findings to the existing literature.

Figure 2: Fit of the Estimated Densities



Notes: Left panel: the continuous part of the density integrates to one minus the unemployment rate. Right panel: percentiles (10, 20, 50, 80, 90) computed from the estimated densities (red) and directly from the cross-sectional observations (blue).

5.1 Micro Data on Earnings and Density Representation

The micro-level earnings data are constructed in the same way as in the application in CCS. Weekly earnings (PRERNWA) are obtained from the monthly Current Population Survey (CPS) through the website of the National Bureau of Economic Research (NBER) and scaled to annual earnings by multiplying with 52. Based on the CPS variable PREXPLF "Experienced Labor Force Employment" we construct an employment indicator which is one if the individual is employed and zero otherwise. This indicator is used to compute an aggregate unemployment rate which we include in the Y_t vector; see Section 4.

We standardize individual-level earnings by (2/3) of nominal per-capita GDP and apply the inverse hyperbolic sine transformation, which is given by

$$x = g(z|\theta) = \frac{\ln(\theta z + (\theta^2 z^2 + 1)^{1/2})}{\theta} = \frac{\sinh^{-1}(\theta z)}{\theta}, \quad z = \frac{\text{Earnings}}{(2/3) \cdot \text{per-capita GDP}} \tag{26}$$

with $\theta = 1$. For small values of z the function is approximately equal to z and for large values of z it is equal to $\ln(z) + \ln(2)$. Below we will refer to x as transformed data and to z as original data. We assume that the transformed earnings are located on the interval $[0, \bar{x}]$ and use a cubic spline as basis functions. We construct the spline from $x = \bar{x}$ to $x = 0$, using a linear element for the right tail:

$$\zeta_K(x) = \max\{\bar{x} - x,\, 0\}, \quad \zeta_k(x) = \big[\max\{x_{k-1} - x,\, 0\}\big]^3 \text{ for } k = K-1, \ldots, 1. \tag{27}$$

In the left panel of Figure 2 we plot the estimated density for the non-transformed earnings data z_{it}, see (26), for the period 2006:M8. We overlay a histogram constructed from the raw data. We regard the earnings distribution as a mixture of a point mass at zero and a continuous part. The point mass equals the unemployment rate and the continuous part integrates to one minus the unemployment rate. The density estimate smoothes out the histogram and distributes the point mass at the top coded value into the right tail of the distribution. In the right panel we overlay empirical percentiles with percentiles computed from the estimated densities to show that the two measures are very similar.

5.2 Earnings Distribution Response

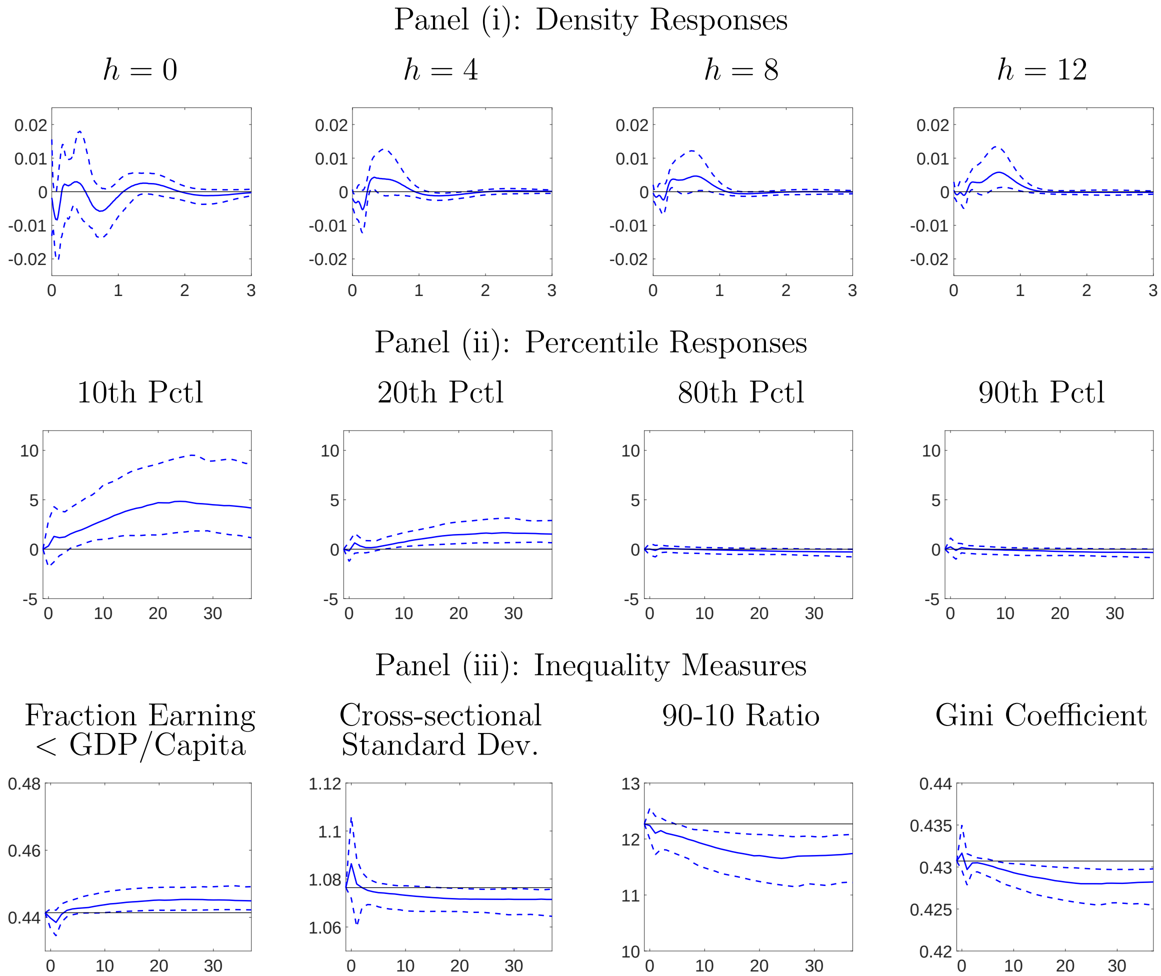
Panel (i) of Figure 3 depicts the response of the continuous part of the earnings distribution to a monetary policy shock. The panels show the difference between the steady state earnings density and the shocked density for $h = 0$ (impact of the shock), $h = 4$, $h = 8$, and $h = 12$. The x-axis in these plots correspond to the level of earnings. Recall that a value of one means that the earnings of the individual are equal to 2/3 (approximately the labor share) of GDP per capita. As mentioned previously, the earnings densities are normalized to integrate to one minus the unemployment rate. Because the unemployment rate drops in response to an expansionary monetary policy shock, the probability mass increases along the density response, relative to the steady state density.

Earnings above 2 are essentially not affected by the monetary policy intervention. The IRFs provide some evidence that the probability mass of individuals earning between 0.5 and 1 times GDP per capita drops and the mass of individuals earning between 1 and 2 times GDP increases upon impact. However, the 80% bands are wide and the sign of the responses is mostly ambiguous. The density differential for earnings between 1 and 2 reverts quickly to zero, whereas the differential for earnings between 0.5 and 1 becomes positive in the medium run before the earnings density reverts back to its steady state in the long run.

The density response can be converted into IRFs for statistics derived from the earnings distribution. Panel (ii) of Figure 3 shows the responses of the percentiles of the earnings distribution as a function of the horizon h. The percentile responses account for a point mass of zero labor earnings that corresponds to the number of individuals that are unemployed. They are reported as percentage changes relative to the base level. For instance, suppose in steady state the earnings level is 0.2 times the labor share of GDP per capita at the 10th percentile and after the shock earnings rise to 0.21. This corresponds to a 5% increase.

According to the plots in Panel (ii), in percentage terms, the monetary policy shock has the largest impact on the earnings distribution at the 10th percentile, capturing in part the

Figure 3: Response of Earnings to Monetary Policy Shock



Notes: Responses to a 25bp monetary policy shock. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. The plots depict 10th (dashed), 50th (solid), and 90th (dashed) percentiles of the posterior distribution. Panel (i): earnings level is on the x-axis. As distributional responses we depict differences between the shocked and the steady state cross-sectional density (continuous part, normalized to $1 - UR_t$) of earnings / (2/3) GDP per capita at various horizons. Panels (ii) and (iii): horizon h is on the x-axis. The percentile responses are computed from distribution of actual earnings and account for the pointmass at zero.

individuals moving from unemployment into employment. The posterior median response of the 10th percentile ranges from about 0 to 5%. For the 20th percentile the response ranges from 0 to 1 percent and for the 80th and 90th percentiles the responses are essentially zero.

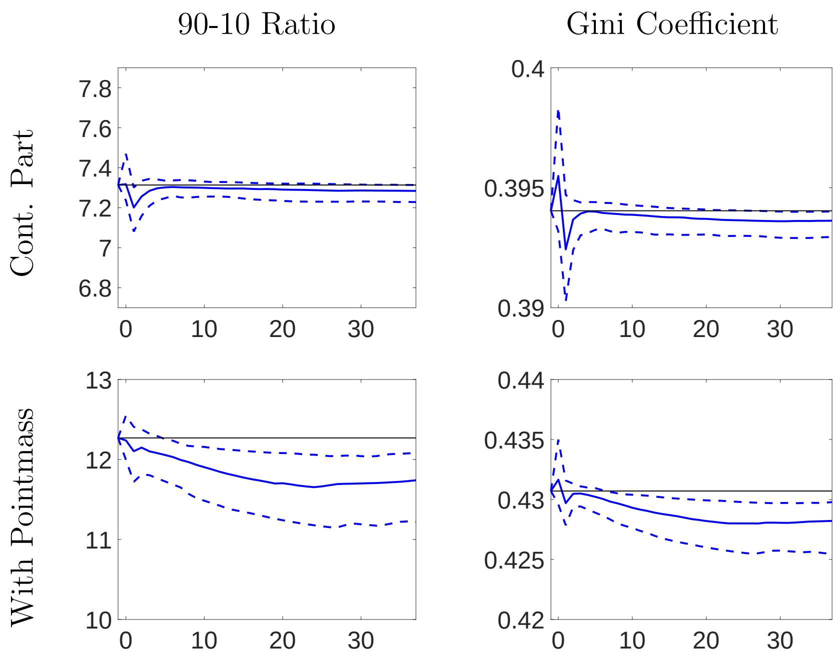
We proceed by computing four measures of earnings inequality from the cross-sectional densities (accounting for the pointmass at zero): the fraction of individuals earning less than the labor share of GDP, the Gini coefficient, the ratio of the 90th and the 10th percentile of the income distribution (90-10 ratio), and the cross-sectional standard deviation. Impulse responses for the inequality measures are depicted in Panel (iii) of Figure 3. At the posterior median the fraction of individuals earning less than two-thirds of GDP per capita slightly rises, from 44% to 44.3% after 36 months. The cross-sectional standard deviation increases, from 1.08 to 1.09, upon impact of the monetary policy shock. However, the 80% credible bands are wide, leaving the signs of the initial response ambiguous. Subsequently, the standard deviation drops below its initial level, to about 1.07. The IRFs imply that the Gini coefficient and the 90-10 ratio fall in response to the expansionary monetary policy shock. At the posterior median, the 90-10 ratio drops from 12.27 to 11.76 after 36 months. The path of the Gini coefficient resembles the path of the 90-10 ratio, falling from 0.431 to 0.428. Thus, overall the expansionary monetary policy shock lowers income inequality by raising incomes at the 10th and 20th percentiles.

5.3 Discussion

Our empirical analysis focuses, in the terminology of Kaplan, Moll, and Violante (2018), on indirect effects of household heterogeneity on the propagation of monetary policy interventions. An expansionary monetary policy lowers the real interest rate temporarily, which creates a disincentive to save, and stimulates economic activity in the current period. Labor demand rises and earnings increase. As we have shown in Section 5.2, this increase is most pronounced at the 10th percentile, which is consistent with the notion that low-productivity workers move out of unemployment. In turn inequality as measured through the 90-10 ratio and the Gini coefficient falls.

This result mirrors the finding in regard to the response of the earnings distribution to a technology shock reported in CCS. It is also broadly consistent with a heterogeneous agent model with indivisible labor supply as in Chang and Kim (2006). This class of models generates a negative correlation between idiosyncratic productivity and reservation wage. In turn, low-skill workers enter the labor market during booms, when the demand for labor is sufficiently high such that the wage per efficiency unit exceeds their reservation wage. At this point the labor earnings switch from zero to a positive value, which reduces labor earnings inequality. Ma (2021) incorporates this mechanism into a HANK model and shows that in his model an expansionary monetary policy shock raises wages and more low productivity individuals are starting to work, which raises earnings in the left tail of the distribution. Under his calibration the Gini coefficient for labor earnings (on a scale from 0 to 1) drops

Figure 4: Responses of Inequality Measures: With and Without Pointmass



Notes: Responses to a 25 basis point monetary policy shock. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. The plot depicts 10th (dashed), 50th (solid), and 90th (dashed) percentiles of the posterior distribution. The inequality measures in the top row are computed from the continuous part of the distribution of actual earnings, not assigning labor earnings of zero for the unemployed. The bottom row includes the pointmass at zero in the computation of the inequality statistics.

by approximately 0.001 upon impact.[9] In our estimated VAR the drop is between 0.001 and 0.003, which is very similar.

As in Ma (2021)'s HANK model, the reduction in the earnings inequality is mainly driven by the fall in unemployment. We recomputed the response of the earnings distribution and the derived inequality measures by excluding the pointmass at zero and normalizing the continuous part of the earnings density to one. The results are plotted in Figure 4. The comparsion of the IRFs in the top row (no pointmass at zero for the unemployed) to the IRFs in the bottom row (which are identical to the ones previously shown in Figure 3), shows that the effect of monetary policy shocks on earnings inequality is mostly driven by individuals switching between unemployment and employment.

There are a number of earlier studies that examined the effect of monetary policy shocks on earnings inequality. Some studies typically included the inequality measures directly into

[9]See Figure 3 in Ma (2021). He considers a 100 bp monetary policy shock and measures the Gini coefficient on a scale from 0 to 100. Thus, $-0.4/(4 \cdot 100) = -0.001$.

a VAR or linear projections. For instance, Coibion, Gorodnichenko, Kueng, and Silvia (2017) report IRFs to a 100 bp increase in the monetary policy rate, estimated on U.S. data. They find that in the medium run the Gini coefficient on earnings rises by about 0.0025. Adjusting for the different shock size, their estimate is slightly smaller than ours. Furceri, Loungani, and Zdzienicka (2018) consider a panel of 32 advanced and emerging market economies. They report an estimate (converted into our scale) of 0.005, which is larger, but in the same order of magnitude as our estimates. Other studies did not aggregate their results into the response of inequality statistics, but they also find that monetary policy shocks predominately affect the left tail of the earnings distribution, and that the effects are driven to a large extent by transitions into and out of employment, e.g., Amberg, Jansson, Klein, and Rogantini Picco (2022), Mitman, Broer, and Kramer (2022), and Lenza and Slacalek (2023).

6 Consumption Response to Monetary Policy

In the second part of the empirical analysis we estimate the response of the cross-sectional consumption distribution to a monetary policy shock. The micro data are discussed in Section 6.1, and the IRF results are summarized in Section 6.2, and Section 6.3 relates the empirical findings to the existing literature. As mentioned in Section 4 the vector of aggregate variables Y_t now also includes aggregate personal consumption expenditures per capita series.

6.1 Micro Data on Consumption and Density Representation

The consumption data are obtained from the Consumer Expenditure Survey (CEX), conducted by the Bureau of Labor Statistics. Although interviews about consumption expenditures are conducted at monthly frequency, we aggregate to quarterly frequency to smooth out unit-level expenditures and reduce reporting errors. Each consumption unit (CU) provides expenditure data for three consecutive months. If a CU was surveyed in 1990Q1, consumption responses could refer to one of three possible combination of months: (i) Oct. 1989, Nov. 1989, Dec. 1989; (ii) Nov. 1989, Dec. 1989, Jan. 1990; or (iii) Dec. 1989, Jan. 1990, Feb. 1990. To convert this information into quarterly expenditure data, we add the three monthly values and assign the sum to the quarter that covers at least two of the months for which responses were obtained. Lastly, we use the count of males and females over 16 years old within a CU to compute per capita expenditures. Following this approach, we compile cross-sectional observations for the periods 1990:Q1 to 2016:Q4.

We construct four measures of consumption: (i) consumption of nondurable goods includes food and beverages, clothing and footware, gasoline and other fuel, personal care, readings, and tobacco. (ii) Consumption of services encompasses child-care spending, hospital and nursery services, household utilities and energy, recreation services, financial services, accommodations, telecommunication, transportation. (iii) Total consumption encompasses consumption of nondurables and services and also covers spending on items such as recreational goods, furniture and furnishing, jewelry and watches, housing and rent (or imputed rental value), health care and insurance, education, cash support for college students, vehicle purchases. (iv) Finally, we define consumption of durables as total consumption minus consumption of nondurables and services.[10]

The micro-level consumption data are scaled by aggregate nominal consumption from NIPA, divided by the population over 16 years old. It is well known that CEX consumption per capita does not aggregate to NIPA consumption per capita. Thus, we assume that only a fraction $\tilde{c}$ of total individual consumption c^*_{it} is reported to the CEX: $c_{it} = \tilde{c}c^*_{it}$. We compute separate constants for each consumption category: total (T), non-durables (N), services (S), and durables (D):

$$\tilde{c}^s = \frac{1}{T}\sum_{t=1}^{T} \text{median}(c^s_{1t}, \ldots, c^s_{Nt})/C_t, \quad s \in \{T, N, S, D\}. \tag{28}$$

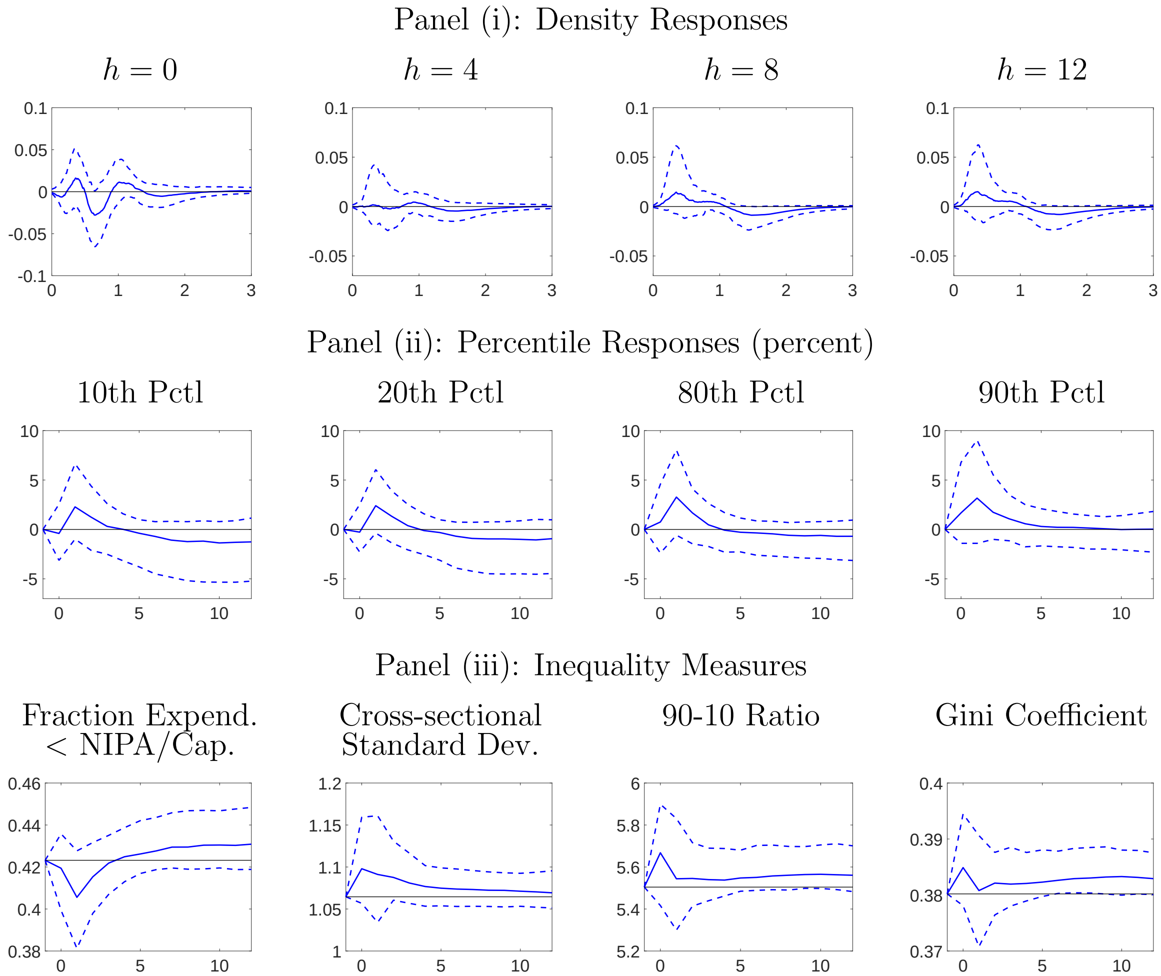
The scaling constants are $\tilde{c}^T = 0.46$, $\tilde{c}^D = 0.02$, $\tilde{c}^S = 0.07$, $\tilde{c}^N = 0.05$ and capture both the underreporting and the fraction of the components N, S, D as part of total consumption. We then define $z^s_{it} = c^s_{it}/(\tilde{c}^s \cdot C_t)$ such that a value of one means that the unit approximately consumes at the aggregate per capital level. Finally, as for the earnings data, we apply the inverse hyperbolic sine transformation to obtain x_{it}.

6.2 Consumption Distribution Response

Figure 5 summarizes the response of the cross-sectional consumption distribution to a monetary policy shock. Panel (i) shows density differentials, Panel (ii) depicts the responses

[10]Starting point of the definition of consumption and its components is Coibion, Gorodnichenko, Kueng, and Silvia (2017), but then we made some adjustments to add items that were excluded from their analysis. Detailed information on the data construction can be obtained from our replication files.

Figure 5: Response of Consumption to Monetary Policy Shock



Notes: Responses to a 25 basis point monetary policy shock. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. The plots depict 10th (dashed), 50th (solid), and 90th (dashed) percentiles of the posterior distribution. Panel (i): we depict differences between the shocked and the steady state cross-sectional density of consumption expenditures/aggregate consumption per capita at various horizons. Panels (ii, iii): The x-axes refer to quarters. The responses are computed from distribution of consumption expenditures/aggregate consumption per capita. Percentile responses are in percent from steady state value.

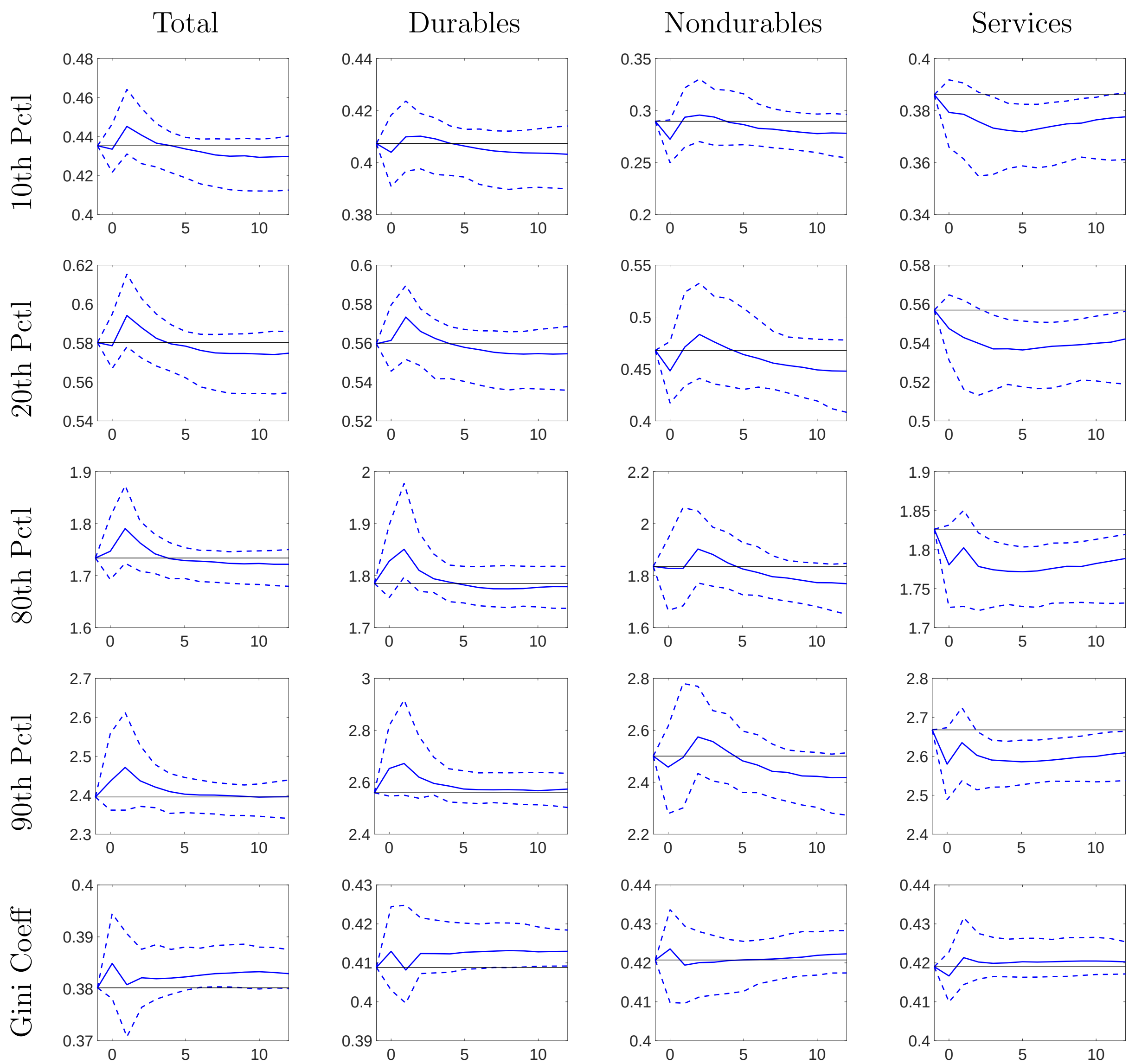
of percentiles of the consumption distribution, and Panel (iii) documents the response of inequality measures. The differentials in Panel (i) are with respect to a steady state distribution with a median of 1.0 (which corresponds to NIPA per capita consumption). While the credible bands span zero, at horizons $h = 8$ and $h = 12$, according to the pointwise posterior median estimates, probability mass shifts from units consuming more than scaled per-capita consumption to those that are consuming less.

The percentile responses reported in Panel (ii) are in percentage deviations from percentiles of the steady state distribution. For instance, a 5% increase at the median would imply that consumption rises from 1 to 1.05. Based on the posterior median the 10th and 20th percentiles rise above their steady state values at $h = 1$, before they revert back to zero four quarters after impact. Over the subsequent quarters the responses stay 0.9% (20th percentile) to 1.3% (10th percentile) below their baseline values. The responses of the 80th and 90th percentiles are similar in shape but quantitatively slightly larger, which is most apparent from the credible bands. The responses of normalized unit-level consumption could be converted into dollar values of individual level consumption by combining the responses in Panel (ii) with the positive response of aggregate consumption in Figure 1.

Panel (iii) of Figure 5 depicts the implied responses of our four inequality measures. The posterior median response of the fraction of households consuming less than the aggregate per capita amount falls initially, but stays above its steady state value from horizon $h = 5$ onwards. The other three inequality measures – the cross-sectional standard deviation, the 90-10 ratio, and the Gini coefficient peak upon impact ($h = 0$) and stay above their steady state levels subsequently. Thus, according to the posterior median estimates, consumption inequality generally increases in response to an interest rate cut. However, there is a considerable amount of uncertainty as the credible bands cover both positive and negative values.

In Figure 6 we compare the responses of the distribution of consumption components. The first column of the figure reproduces the percentile and Gini coefficient responses for total consumption from Figure 5, except that we now report the percentile responses in levels rather than in percentage changes relative to the steady state distribution. The remaining columns show the responses of durables, nondurables, and services. Because of the normalization constants defined in (28), the steady state levels for the percentiles of the consumption components are roughly the same. The responses of the inequality statistics associated with the consumption of durables look very similar to the responses of total consumption. The response of the distribution of nondurables and services, on the other hand, is quite different. Even at the posterior median, the Gini coefficient response is essentially zero. The posterior medians of the percentile responses of nondurable consumption oscillate at short horizons and then fall below the steady state level. The service percentiles drop upon impact below their steady state values and remain low for the subsequent ten quarters, indicating a shift from services to durables. Recall, however, from Figure 1 that total per capita consumption, used to normalize unit-level consumption, also rises, which mitigates the drop in service consumption.

Figure 6: Responses of Consumption Distribution By Components



Notes: Responses to a 25 basis point monetary policy shock. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. The plots depict 10th (dashed), 50th (solid), and 90th (dashed) percentiles of the posterior distribution. The x-axes refer to quarters. The responses are computed from the distribution of consumption expenditures/aggregate consumption per capita. Responses in the first column are identical to those in Figure 5.

6.3 Discussion

The slight increase in consumption inequality may appear puzzling in view of the decrease in earnings inequality documented in Section 5. The consumption response is determined by the income response in combination with the marginal propensity to consume (MPC) and low-income households, i.e., those that experience a strong rise in labor earnings, tend to be those with high MPCs. However, in Section 5 we only considered labor income. Wealthy households typically also have a significant amount of unearned income, which may be directly affected by the monetary policy intervention, e.g., rising stock and bond prices in response to an interest rate cut. For instance, using Swedish administrative data Amberg, Jansson, Klein, and Rogantini Picco (2022) find that high-income individuals experience a substantial rise of capital income, more so than other individuals. This channel is also present in a calibrated DSGE model by Lee (2021), who finds that consumption inequality rises in his model in response to an expansionary monetary policy shock because it boosts profits and equity prices which benefits wealthy households with high level of consumption.

7 Financial Income Response to Monetary Policy

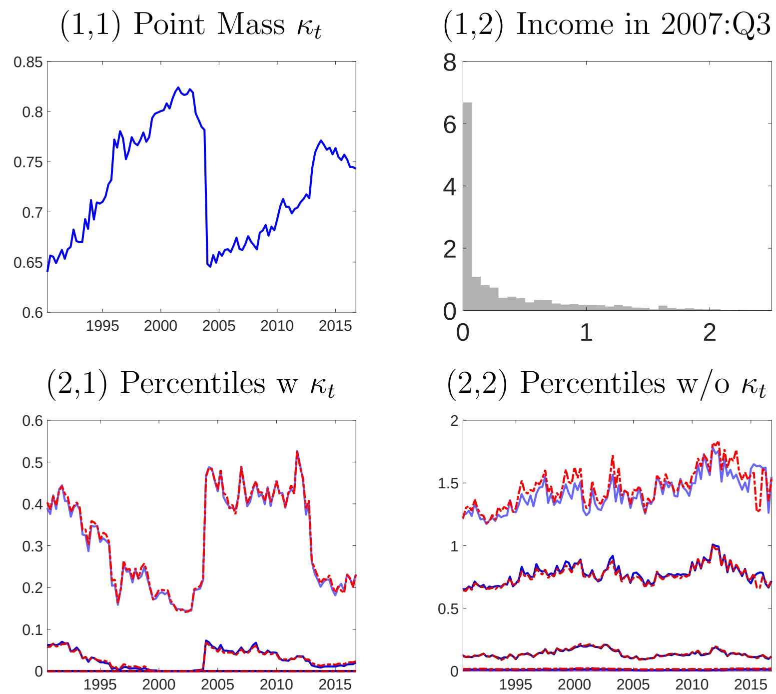
In view of the findings in the previous two sections we now turn to the analysis of the response of the financial income distribution, as measured by the CEX, to a monetary policy shock.

Micro Data. The cross-sectional densities of household level financial income are constructed from the public-use income data of the CEX.[11] We use the CEX family interview files from 1990:Q1 to 2016:Q4, defining financial income as the sum of the following components: (i) amount earned as interest on savings accounts or bonds (intearnx), (ii) amount of regular income earned from dividends, royalties, estates, or trusts (finincx), (iii) amount received from pensions or annuities from private companies, military, or government (pensionx), (iv) amount of net income or loss received from roomers or boarders (inclossa), (v) amount of net income or loss received from payments from other rental properties (inclossb). If a component of income is missing, we set it to zero.

The survey asks consumption units (CU) about their financial income in the aforementioned categories. Ideally, we would like to know the income over the past three months

[11]The CPS, used in Section 5, has accurate information on wages and salaries, but misses more than one third of investment income; see Brady and Brass (2021). Alternative sources for financial income are the Survey of Consumer Finances (SCF) and the Panel Study of Income Dynamics (PSID). Both SCF and PSID are rich data sources, but only available at low frequency (mostly tri-annual for the SCF and bi-annual for PSID), which makes them unsuitable to study the link between income inequality and business cycle fluctuations.

Figure 7: Cross-Sectional Distribution of Financial Income in the CEX



Notes: (1,1): κ_t is the fraction of households with financial income of $\underline{\text{x}} = 0.00143$ or less. (1,2) Histogram of income in 2007:Q3 conditional on $x > \underline{\text{x}}$. Panels (2,1) and (2,2): percentiles (10, 20, 50, 80, 90) computed from the estimated densities (red) and directly from the cross-sectional observations (blue).

prior to the interview. Unfortunately, the survey asks for the income over a twelve-month period which leads us to adopt the following convention: if a CU is surveyed in 2001:Q1, we record the income response for 2000:Q4. Implementing a more careful mixed-frequency analysis requires assumptions about the unit-level income processes and is beyond the scope of this paper. Individual-level income is calculated by dividing the household level income by the number of CU members over 16 years old. We scale the financial income data using the (approximate) capital share of nominal GDP per capita, which we take to be one-third of its quarterly value.

We define the pointmass κ_t as the fraction of individuals earning less than $\underline{\text{x}} = 0.0014$.[12] The time series of κ_t is plotted in Panel (1,1) of Figure 7 and fluctuates between 0.65 and 0.82 over the sample period. We regard the fraction of CUs κ_t as not having any substantial financial income. This includes the 3% of households that report small negative

[12]Given the prevalence of households with modest yet positive financial income, we use $\underline{x}$ as a threshold for households lacking substantial financial income. $\underline{x}$ represents the 10th percentile of pooled financial income data standardized by the capital share of GDP across the sample periods.

financial income. There is a substantial drop of κ_t in 2004, which is caused by a change in the imputation approach for unreported financial income. As discussed in Section 4, κ_t is included as an aggregate variable in the functional VAR specifications. We compute separate means for the pre- and post-2004 κ_t values and include the demeaned series in the vector of aggregate variables.

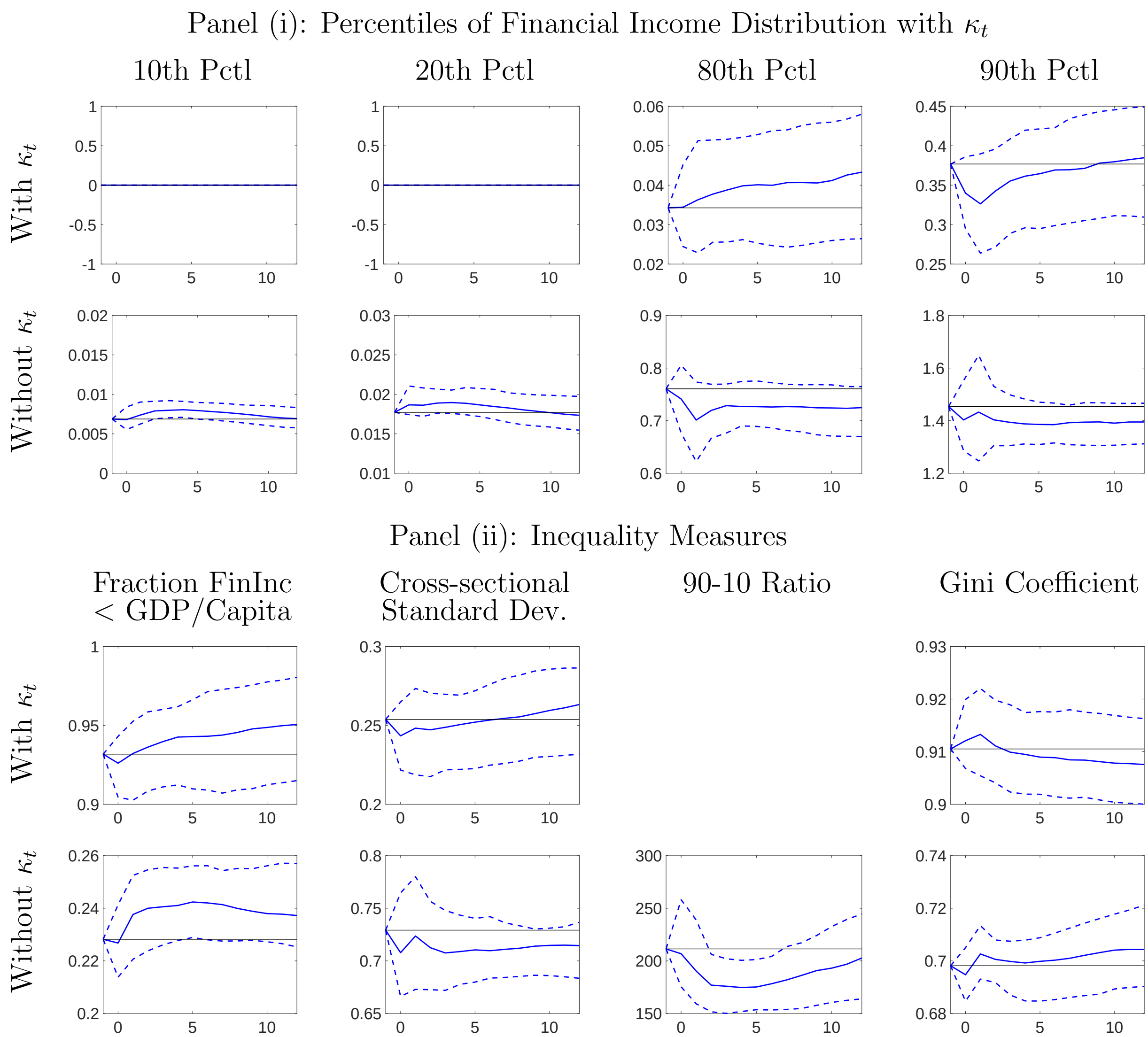
In Panel (1,2) we plot a the histogram for financial income in 2007:Q3, conditional on being larger than $\underline{x}$. Despite excluding households with an income less than 0.0014 of the capital share of per capita GDP, there is still a large mass near zero. In the bottom row of Figure 7 we show percentiles (10, 20, 50, 80, 90) of the financial income distribution. One set of percentiles is computed from the estimated densities and the other one directly from the cross-sectional observations. Both of them are essentially identical. We loosely treat κ_t as a pointmass at zero, just like we treated the unemployment rate in Section 5 as a pointmass of zero (labor) earnings. If κ_t is included in the calculation of the percentiles, then the 10th, 20th, and 50th percentiles are zero; see Panel (2,1). In Panel (2,2) we show the percentiles for the distribution of x conditional on $x \geq \underline{x}$ which are by virtue of the conditioning larger than the ones plotted in Panel (2,1).

Financial Income Distribution Response. Impulse responses for the financial income distribution are plotted in Figure 8. Panel (i) summarizes the responses of the percentiles, and Panel (ii) shows the responses of the inequality measures. For both panels, we compute the cross-sectional statistics with and without κ_t. In view of Panel (2,1) of Figure 7 it is not surprising that the 10th and 20th percentiles do not respond to the monetary policy shock because they are zero throughout the sample. Even if κ_t is excluded from the percentile calculations, the responses are essentially zero. There is more movement in the 80th and 90th percentiles. In three out of four cases the posterior mean response is negative but the bands are wide and include zero.

The fraction of households with financial income less than the (capital share of) GDP per capita tends to increase. The cross-sectional standard deviation and the Gini coefficient essentially don't respond to the monetary policy shock, regardless of whether the κ_t pointmass is included in the calculation or not. The 90-10 ratio is not defined if κ_t is included in the computation of the percentiles, because the 10th percentile is zero. Conditional on $x \geq \underline{x}$ the 90-10 ratio drops, which is due to the fall in the 90th percentile, shown in Panel (i).

We conjecture that the reason for not finding an unambiguous increase in financial income

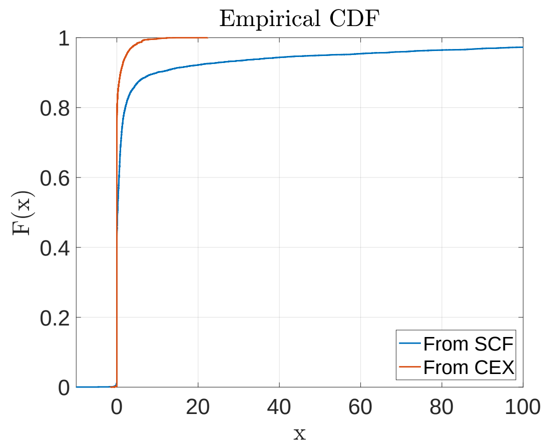
Figure 8: Response of Financial Income to Monetary Policy Shock



Notes: Responses to a 25bp monetary policy shock. The system is in steady state at $h = -1$ and the shock occurs at $h = 0$. The plots depict 10th (dashed), 50th (solid), and 90th (dashed) percentiles of the posterior distribution. horizon h is on the x-axis. Panels (i) and (ii): the percentile responses and inequality measures are computed from distribution of financial income and account for the pointmass at zero.

in the right tail of the distribution, as documented in other studies, is that the CEX excludes many of the wealthy households that are present in other studies. We plot the empirical cumulative distribution function (cdf) of financial income for the year 2012 in Figure 9. The graph overlays cdfs obtained from our CEX data and from the more comprehensive SCF. Relative to the SCF, the CEX misses the top-10 percent of households with the highest

Figure 9: Financial Income Distribution in 2012: CEX vs. SCF



Notes: The figure depicts empirical cdfs, including negative values for financial income. Financial income from both sources is standardized by 1/3 (approximate capital share) of GDP per capita.

financial income, who may benefit from the interest cut. The households included in our study are likely to experience lower interest on their savings account without benefiting from stock market gains or dividend payouts.

Despite the CEX missing high financial income households, it does cover the same set of households that underly our analysis of the consumption distribution response. Thus, the slight increase in consumption inequality shown in Figure 5 may be due to other factors. First, even wealthier households with a relatively high level of consumption may be close to their borrowing constraint (wealthy hand-to-mouth consumers) and have a relatively high marginal propensity to consume which could explain the larger response in the right tail of the distribution. Second, the direct effect of monetary policy on consumption may be heterogeneous, whereby high consumption/income households' consumption decisions are more interest rate sensitive.

8 Response to an Informational Shock

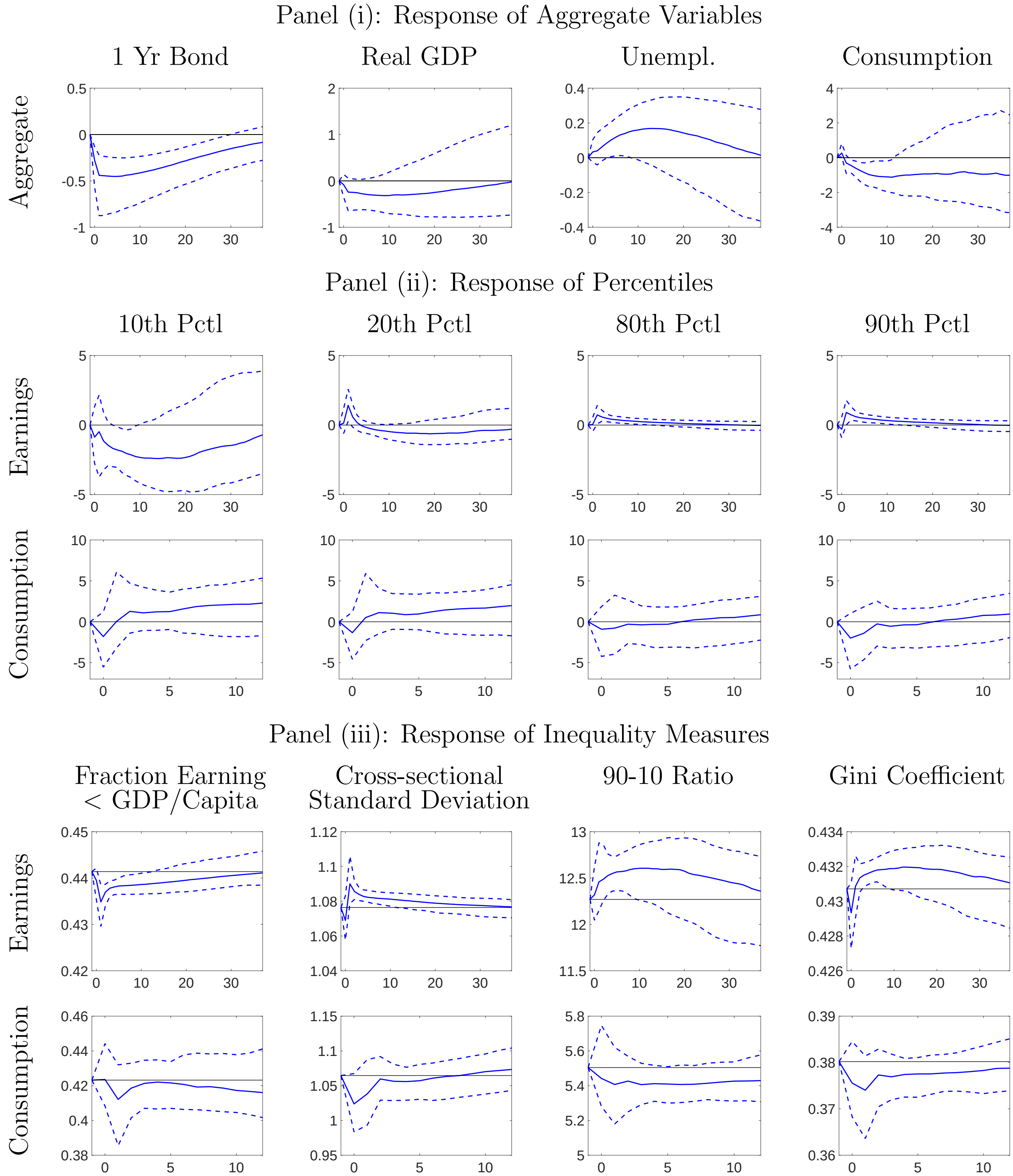
Thus far, we have focused on responses to a conventional monetary policy shock. The JK instruments also allow us to compute IRFs for an information shock, meaning the unanticipated change in the interest rate conveys private information of the central bank about the state of the economy. For instance, a reduction in interest rates may signal that aggregate output and prices will be lower than expected by the public. The shock is set identified through the assumption that surprises in interest rates have the same sign as surprises to the stock market index. Summary statistics for the earnings and consumption distribution responses are depicted in Figure 10.

Aggregate Responses. Panel (i) summarizes the responses of the aggregate variables: a reduction of bond yields indicates that real activity and prices will be below expectation. At the posterior median real GDP drops 20 bp, consumption falls, and the unemployment rate increases up to 0.15 percentages after one year. The posterior mean responses revert back to zero after three years, except for consumption which exhibits a more persistent response.

Earnings Distribution Response. The top row of Panel (ii) contains the responses of the earnings density percentiles. Most notably, earnings at the 10th percentile experience a persistent drop of about 2% at the posterior median, whereas there is a small and short-lived increase for the other percentiles. Panel (iii) contains responses of inequality measures. The fraction of individuals earning less than GDP per capita drops in the short run. The cross-sectional standard deviation increases and eventually reverts back to its initial level. According to the response of the 90-10 ratio and the Gini coefficient inequality rises in the long run. Thus, overall the simultaneous unanticipated drop in interest rates and stock prices increases earnings inequality because the contraction is associated with an increase in unemployment.

Consumption Distribution Response. The bottow rows of Panels (ii) and (iii) of Figure 10 depict the response of percentiles and inequality statistics to an informational shock derived from the consumption distribution. According to the posterior medians, consumption rises at the 10th and 20th percentiles, stays essentially flat at the 80th percentile, and slightly drops and subsequently reverts back to the steady state level at the 90th percentile. As a consequence, inequality as measured by the 90-10 ratio and the Gini coefficient decreases. As in Section 6, it is important to note that we measure unit-level consumption relative to aggregate per capita consumption. The latter, shown in the right most graph in Panel (i), drops in response to the information shock. Thus, the rise of relative consumption at the 10th and 20th percentiles is dampened by the simultaneous drop in aggregate consumption.

Figure 10: Aggregate and Distribution Responses to Information Shock



Notes: Shock occurs at $h = 0$. The plots depict 10th (dashed), 50th (solid), and 90th (dashed) percentiles of the posterior distribution. Panel (i): real GDP and consumption resonses are percentage deviations from the steady state, bond yield and unemployment rate responses are absolute percentages. Panels (ii,iii): earnings responses account for the pointmass at zero. x-axis is months for earnings and quarters for consumption.

9 Conclusion

We estimated a functional VAR that stacks macroeconomic aggregates and the cross-sectional distribution of earnings, consumption, and financial income, respectively, to present semi-structural evidence about the distributional effects of monetary policy shocks. The empirical results in our paper could be used to assess the quantitative implications of HANK models more formally. We only provided an informal discussion and leave a more rigorous quantitative assessment for future research. Finally, there is a broader normative question whether central banks should track the effect of their policy interventions on inequality statistics. In regards to earnings inequality, our results suggest that distributional effects can be understood from the response of the unemployment rate. The small and uncertain responses of consumption and financial income inequality measures provide, in our view, a case for central banks to focus on the effects of their policies on macroeconomic aggregates as they have been doing traditionally.

References

Almuzara, M., and V. Sancibrian (2023): "Micro Responses to Macro Shocks," *Working Paper, FRB New York*.

Amberg, N., T. Jansson, M. Klein, and A. Rogantini Picco (2022): "Five Facts about the Distributional Income Effects of Monetary Policy Shocks," *AER: Insights*, 4(3), 289–304.

Andersen, A. L., N. Johannesen, M. Jorgensen, and J.-L. Peydro (2021): "Monetary Policy and Inequality," *Working Paper, Universitat Pompeu Fabra*, 1761.

Anderson, E., A. Inoue, and B. Rossi (2016): "Heterogeneous Consumers and Fiscal Policy Shocks," *Journal of Money, Credit, and Banking*, 48(8), 1877–1888.

Arias, J. E., J. Rubio-Ramirez, and D. Waggoner (2022): "Inference in Bayesian Proxy-SVARs," *Journal of Econometrics*, forthcoming.

Bjornland, H. C., Y. Chang, and J. Cross (2023): "Oil and the Stock Market Revisited: A Mixed Functional VAR Approach," *Manuscript, BI Norwegian Business School and Indiana University*.

Bosq, D. (2000): *Linear Processes in Function Spaces*. Springer Verlag, New York.

Brady, P. J., and S. Brass (2021): "Comparing the Current Population Survey to Income Tax Data," *SSRN Working Paper*, 4025470.

Caldara, D., and E. Herbst (2019): "Monetary Policy, Real Activity, and Credt Spreads: Evidence from Bayesian Proxy SVARs," *American Economic Journal: Macroeconomics*, 11(1), 157–192.

Chan, J. C. C. (2022): "Asymmetric Conjugate Priors for Large Bayesian VARs," *Quantitative Economics*, 13(3), 1145–1169.

Chang, M., X. Chen, and F. Schorfheide (2024): "Heterogeneity and Aggregate Fluctuations," *Journal of Political Economy*, forthcoming.

Chang, Y., C. S. Kim, and J. Park (2016): "Nonstationarity in Time Series of State Densities," *Journal of Econometrics*, 192, 152–167.

Chang, Y., and S.-B. Kim (2006): "From Individual to Aggregate Labor Supply: A Quantitative Analysis Based on a Heterogeneous Agent Macroeconomy," *International Economic Review*, 47, 1–27.

Cloyne, J., C. Ferreira, and P. Surico (2020): "Monetary Policy When Households Have Debt: New Evidence on the Transmission Mechanism," *Review of Economic Studies*, 87, 102–129.

Coibion, O., Y. Gorodnichenko, L. Kueng, and J. Silvia (2017): "Innocent Bystanders? Monetary Policy and Inequality," *Journal of Monetary Economics*, 88, 70–89.

Del Canto, F., J. Grigsby, E. Qian, and C. Walsh (2023): "Are Inflationary Shocks Regressive? A Feasible Set Approach," *Manuscript, Columbia University.*

Diebold, F. X., and C. Li (2006): "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics*, 130, 337–364.

Doepke, M., and M. Schneider (2006): "Inflation and the Redistribution of Nominal Wealth," *Journal of Political Economy*, 114(6), 1069–1097.

Fisher, I. (1933): "The Debt-Deflation Theory of Great Depressions," *Econometrica*, 1(4), 337–357.

Furceri, D., P. Loungani, and A. Zdzienicka (2018): "The Effects of Monetary Policy Shocks on Inequality," *Journal of International Money and Finance*, 85, 168–186.

Guerello, C. (2018): "Conventional and Unconventional Monetary Policy vs. Household Income Distribution: An Empirical Analysis for the Euro Area," *Journal of International Money and Finance*, 85, 187–214.

Holm, M. B., P. Paul, and A. Tischbirek (2021): "The Transmission of Monetary Policy under the Microscope," *Journal of Political Economy*, 129(10), 2861–2904.

Horvath, L., and P. Kokoszka (2012): *Inference for Functional Data with Applications.* Springer Verlag, New York.

Hu, B., and J. Y. Park (2017): "Econometric Analysis of Functional Dynamics in the Presence of Persistence," *Manuscript, Department of Economics, Indiana University.*

Inoue, A., and B. Rossi (2020): "The Effects of Conventional and Unconventional Monetary Policy: A New Approach," *Manuscript, Vanderbilt University and Pompeu Fabra.*

Jarocinski, M., and P. Karadi (2020): "Deconstructing Monetary Policy Surprises - The Role of Information Shocks," *Amercian Economic Journal: Macroeconomics*, 12, 1–43.

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

Lee, D. (2021): "The Effects of Monetary Policy on Consumption and Inequality," *Manuscript, Federal Reserve Bank of New York.*

Lenza, M., and J. Slacalek (2023): "How Does Monetary Policy Affect Income and Wealth Inequality? Evidence from Quantitative Easing in the Euro Area," *Manuscript, European Central Bank.*

Ma, E. (2021): "Monetary Policy and Inequality: How Does One Affect the Other," *Working Paper, Lousiana State University.*

McKay, A., and C. K. Wolf (2023): "Monetary Policy and Inequality," *Journal of Economic Perspectives*, 37(1), 121–144.

Meeks, R., and F. Monti (2019): "Heterogeneous Beliefs and the Phillips Curve," *Bank of England Staff Working Paper*, 807.

Mitman, K., T. Broer, and J. Kramer (2022): "The Curious Incidence of Monetary Policy Across the Income Distribution," *CEPR Discussion Paper*, 17589.

Nakamura, E., and J. Steinsson (2018): "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect," *Quarterly Journal of Economics*, 133, 1283–1330.

Plagborg-Møller, M., and C. K. Wolf (2021): "Local Projections and VARs Estimate the Same Impulse Responses," *Econometrica*, 89(2), 995–980.

Ramsey, J. O., and B. W. Silverman (2005): *Functional Data Analysis.* Springer Verlag, New York, 2nd edn.

Stock, J. H., and M. W. Watson (2010): "Monthly GDP and GNI – Research Memorandum," *Manuscript, Princeton University.*

Online Appendix: On the Effects of Monetary Policy Shocks on Earnings and Consumption Heterogeneity

Minsu Chang, and Frank Schorfheide

This Appendix consists of the following sections:

A Density Estimation

A.1 Top Coding

Likelihood Function with Censoring. We define the censoring point c_t as

$$c_t = \max_{i=1,\ldots,N} x_{it}$$

Moreover, we let

$$N_{t,max} = \sum_{i=1}^{N} \mathbb{I}\{x_{it} = c_t\}.$$

If $N_{t,max} = 1$, we assume that the observed sample is not constrained by the top-coding and use the standard likelihood function described in the main text. If $N_{t,max} > 1$ we use a likelihood function that assumes that any earnings value exceeding c_t is coded as c_t.

Recall that in the main text we ignored the dependence of the cross-sectional sample size N on t in the notation and defined $p^{(K)}(X_t|\alpha_t) = \exp\left\{N\mathcal{L}^{(K)}(\alpha_t|X_t)\right\}$, where

$$\mathcal{L}^{(K)}(\alpha_t|X_t) = \bar{\zeta}'(X_t)\alpha_t - \ln\int_0^\infty \exp\left\{\zeta'(x)\alpha_t\right\}dx, \quad \bar{\zeta}(X_t) = \frac{1}{N}\sum_{i=1}^{N_t}\zeta(x_{it}).$$

We introduce the unknown parameter $\pi_t = \mathbb{P}\{x_{it} \geq c_t\}$. We drop the top-coded observations from the definition of $\bar{\zeta}(X_t)$ und make the time dependence explicit in the notation. Let

$$\bar{\zeta}_t(X_t) = \frac{1}{N_t}\sum_{i=1}^{N_t}\zeta(x_{it})\mathbb{I}\{x_{it} < c_t\}. \tag{A.1}$$

The log likelihood function is obtained as follows: the sample contains $N_{t,max}$ top-coded observations where the probability of sampling a top-coded observation is π_t. The probability of samping an observation that is not top-coded is $(1-\pi_t)$. Conditional on not being top-coded, the observation $x_{it} < c_t$ is sampled from a continuous density with a domain that is truncated at c_t. Thus, dividing the log-likelihood by the sample size N_t, we obtain

$$\begin{aligned}\mathcal{L}^{(K)}(\alpha_t, \pi_t|X_t) \;\; = \;\; & \frac{N_{t,max}}{N_t}\ln\pi + \frac{N_t - N_{t,max}}{N_t}\ln(1-\pi_t) \\ & + \bar{\zeta}_t'(X_t)\alpha_t - \frac{N_t - N_{t,max}}{N_t}\ln\int_0^{c_t}\exp\left\{\zeta'(x)\alpha_t\right\}dx.\end{aligned} \tag{A.2}$$

Notice that regardless of the value of α_t, the MLE of π_t is

$$\hat{\pi}_t = \text{argmax}_{\pi \in [0,1]} \mathcal{L}^{(K)}(\alpha_t, \pi_t | X_t) = N_{t,max}/N_t. \tag{A.3}$$

Moreover, regardless of the value of π_t, the MLE of α_t is given by

$$\begin{aligned} \hat{\alpha}_t &= \text{argmax}_{\alpha_t} \mathcal{L}^{(K)}(\alpha_t, \pi_t | X_t) \\ &= \text{argmax}_{\alpha_t} \bar{\zeta}_t'(X_t)\alpha_t - \frac{N_t - N_{t,max}}{N_t} \ln \int_0^{c_t} \exp\left\{\zeta'(x)\alpha_t\right\} dx. \end{aligned} \tag{A.4}$$

The objective function for α_t is almost identical to what we had without top coding, except for a definition of $\bar{\zeta}_t(X_t)$ that drops the top-coded observations in the summation and the factor of $(N_t - N_{t,max})/N_t$ in front of the normalization constant of the density.

Recovering the Density for Uncensored Observations. To reconstruct the full density we can use

$$p(x|\alpha_t) = \frac{\exp\left\{\sum_{k=1}^K \alpha_{k,t}\zeta_k(x)\right\}}{\int_0^\infty \exp\left\{\sum_{k=1}^K \alpha_{k,t}\zeta_k(x)\right\} dx}. \tag{A.5}$$

Note that here we dropped the censoring indicator function and the integration is now from 0 to ∞. Once the α_t's have been estimated based on the censored observations, we work with the full density in the functional state-space model and its K-dimensional approximation.

Modification of Hessian Matrix. We now re-compute the score and the Hessian. Dropping the (K) superscript we obtain the following first derivatives with respect to α_k for $k = 1, \ldots, K$:

$$\mathcal{L}_k^{(1)}(\alpha_t|\pi_t, X_t) = \bar{\zeta}_{t,k}(X_t) - \left(\frac{N_t - N_{t,max}}{N_t}\right) \int_0^{c_t} \zeta_k(x)\bar{p}(x|\alpha_t)dx,$$

where

$$\bar{p}(x|\alpha_t) = \frac{\exp\left\{\sum_{k=1}^K \alpha_{k,t}\zeta_k(x)\right\}}{\int_0^{c_t} \exp\left\{\sum_{k=1}^K \alpha_{k,t}\zeta_k(x)\right\} dx} \mathbb{I}\{x < c_t\}.$$

We can now deduce from our previous calculations that

$$\begin{aligned} &\mathcal{L}_{kl}^{(2)}(\alpha_t|\pi_t, X_t) \\ &= -\left(\frac{N_t - N_{t,max}}{N_t}\right) \int_0^{c_t} \left(\zeta_k(x) - \int_0^{c_t} \zeta_k(x)\bar{p}(x|\alpha_t)dx\right) \\ &\qquad \times \left(\zeta_l(x) - \int_0^{c_t} \zeta_l(x)\bar{p}(x|\alpha_t)dx\right) \bar{p}(x|\alpha_t)dx. \end{aligned} \tag{A.6}$$

Thus, compared to the standard case, the limits of integration change and there is an additional factor $(N_t - N_{t,max})/N_t$.

A.2 Transformations of the $\hat{\alpha}_t$s

Compression/Standardization. The vector $\hat{\tilde{\alpha}}_t = \hat{\alpha}_t - \alpha_*$ may exhibit collinearity. Even though K basis functions may be necessary to approximate the cross-sectional densities, the time variation might be concentrated in a lower-dimensional space, because, for instance, only the means of the cross-sectional distributions are varying over time. This feature can be captured by assuming that the time-variation is captured by a $\tilde{K} < K$ dimensional factor a_t:

$$(\alpha_t - \alpha_*)' = a_t'\Lambda, \tag{A.7}$$

where Λ is a $\tilde{K} \times K$ matrix. As is well known from the factor model literature, Λ and a_t are only identified up to a $\tilde{K} \times \tilde{K}$ dimensional invertible matrix. In principle, the matrix Λ and the sequence of vectors a_t, $t = 1, \dots, T$ have to be estimated simultaneously under this factor structure,

To avoid the simultaneous estimation of the cross-sectional densities, we take the following short cut. First, we compute the $\hat{\alpha}_t$s period-by-period without imposing any restrictions. Second, conditional on α_* we compute the demeaned (and potentially seasonally adjusted) MLEs $\hat{\tilde{\alpha}}_t = \hat{\alpha}_t - \alpha_*$ and arrange them in a $T \times K$ matrix $\hat{\tilde{\alpha}}$ with rows $\hat{\tilde{\alpha}}_t'$. Third, we conduct a principal components analysis which is based on the eigenvalue decomposition of the sample covariance matrix $\hat{\tilde{\alpha}}'\hat{\tilde{\alpha}}/T$. Let $\hat{M}$ be $K \times \tilde{K}$ matrix of eigenvectors associated with the $\tilde{K}$ non-zero eigenvalues (in practice greater than 10^{-10}).[13] Then, let

$$\hat{a} = \hat{\tilde{\alpha}}\hat{M}, \quad \hat{\Lambda} = (\hat{a}'\hat{a})^{-1}\hat{a}'\hat{\tilde{\alpha}}, \tag{A.8}$$

where $\hat{a}$ is the $T \times \tilde{K}$ matrix with rows $\hat{a}_t'$. Even if $\tilde{K} = K$ this operation standardizes the basis function coefficients α_t. To evaluate the MDD formula (25), we replace K by $\tilde{K}$, $\mathcal{L}^{(K)}(\hat{\alpha}_t|X_t)$ by $\mathcal{L}^{(\tilde{K})}(\alpha_* + \hat{\Lambda}'\hat{a}_t|X_t)$, and we change the term $\sum_{t=1}^T \ln|\hat{V}_t|^{1/2}$ to $\sum_{t=1}^T \ln|(\hat{\Lambda}\hat{V}_t^{-1}\hat{\Lambda}')^{-1}|^{1/2}$.

[13]Because our goal is to eliminate perfect collinearities, we choose an eigenvalue cut-off that yields $\alpha_* + \hat{\Lambda}'\hat{a}_t = \hat{\alpha}_t$.

Seasonal Adjustments. Deterministic seasonal adjustments of the cross-sectional densities can be incorporated in the model, as needed, by replacing the vector of constants $\alpha_* = \alpha_t - \tilde{\alpha}_t$ by a time-varying process. In our application the time period t is either a quarter or a month. For quarterly data we let $\alpha_{*,t} = \sum_{q=1}^{4} \alpha_{q,t} s_q(t)$, where $s_q(t) = 1$ if period t is associated with quarter q and $s_q(t) = 0$ otherwise. For monthly data we we use $\alpha_{*,t} = \sum_{m=1}^{12} \alpha_{m,t} s_m(t)$, where $s_m(t) = 1$ if period t is associated with month m and $s_m(t) = 0$ otherwise.

A.3 Recovering Cross-Sectional Densities

Based on the estimated state-transition equation we can generate forecasts and impulse response functions for the compressed coefficients a_t. However, the dynamics of these coefficients in itself are not particularly interesting. Thus, we have to convert them back into densities using the following steps (which can be executed for each prior/posterior draw of a_t from the relevant posterior distribution). First, use (A.7) with $\Lambda = \hat{\Lambda}$ to transform a_t into α_t. If the estimation is based on a seasonal adjustment, α_* can be replaced by $\alpha_{*,t}$, or, if the goal is to compute impulse responses, one could use the average of the seasonal dummies as intercept. Second, compute

$$p^{(K)}(x|\alpha_t) = \frac{\exp\left\{\zeta'(x)\alpha_t\right\}}{\int \exp\left\{\zeta'(\tilde{x})\alpha_t\right\} d\tilde{x}}.$$

B Prior and Posterior Computations

B.1 More Details on the Prior

Recall that the ith equation of the W_t VAR is given by (19). We assume that the parameters (β_i, D_i) are *a priori* independent across equations, i.e.,

$$p(\beta, D) = \prod_{i=1}^{n} p(\beta_i|D_i)p(D_i). \tag{A.9}$$

For each pair (β_i, D_i) we use a Normal-Inverse Gamma (NIG) distribution of the form

$$\beta_i|D_i \sim \mathcal{N}\big(\underline{\beta}_i, D_i \underline{V}_i^{\beta}\big), \quad D_i \sim IG\big(\underline{\nu}_i, \underline{S}_i\big). \tag{A.10}$$

The prior density takes the form

$$
\begin{aligned}
p(\beta_i, D_i) &= (2\pi)^{-k_i/2} |\underline{V}_i^\beta|^{-1/2} \frac{\underline{S}_i^{\underline{\nu}_i}}{\Gamma(\underline{\nu}_i)} D_i^{-(\underline{\nu}_i + 1 + k_i/2)} \\
&\quad \times \exp\left\{ -\frac{1}{D_i} \left[\underline{S}_i + \frac{1}{2} (\beta_i - \underline{\beta}_i)' (\underline{V}_i^\beta)^{-1} (\beta_i - \underline{\beta}_i) \right] \right\}.
\end{aligned}
$$

In the remainder of this subsection we discuss the construction of $\underline{\beta}_i$, $\underline{V}_i^\beta$, $\underline{\nu}_i$, and $\underline{S}_i$. The prior is obtained by transforming a prior for the reduced-form parameters (Φ, Σ) into a prior for the quasi-structural parameters $(\beta_1, \ldots, \beta_{n_w}, D)$.

Prior for D_i and the α_i component of β_i. We start from a prior for $\Sigma = A^{-1'} D A^{-1}$:

$$
\Sigma \sim IW\big(\underline{\nu}, \underline{S}\big), \quad \underline{S} = \mathrm{diag}(\underline{s}_1^2, \ldots, \underline{s}_n^2). \tag{A.11}
$$

Chan (2021) shows that this prior implies

$$
D_i \;\sim\; IG\left(\frac{\underline{\nu} + i - n}{2}, \frac{\underline{s}_i^2}{2} \right), \quad i = 1, \ldots, n. \tag{A.12}
$$

Thus, a comparison with (A.10) indicates that we are setting

$$
\underline{\nu}_i = \frac{\underline{\nu} + i - n}{2}, \quad \underline{S}_i = \frac{\underline{s}_i^2}{2}. \tag{A.13}
$$

Moreover, (A.11) implies that

$$
A_{ij} | D_i \;\sim\; \mathcal{N}\left(0, \frac{D_i}{\underline{s}_j^2} \right), \quad 1 \le j < i, \quad i = 2, \ldots, n, \tag{A.14}
$$

which determines the prior for the α_i component of β_i.

Prior for the $B_{\cdot i}$ component of β_i. Recall that $B_{\cdot i}$ consists of np coefficients on lagged elements of y_t and an intercept. The overall dimension of the vector is $k \times 1$. The prior will take the form

$$
B_{\cdot i} \sim \mathcal{N}\big(\underline{B}_{\cdot i}, \underline{V}_i^B\big), \tag{A.15}
$$

where the $k \times k$ matrix $\underline{V}_i^B$ is assumed to be diagonal.

To specify a prior for $B_{\cdot i}$, we loosely map *a priori* beliefs about $(\alpha_i, \Phi_{\cdot i})$ into beliefs about $B_{\cdot i}$. To simplify the notation a bit, let $\phi_i = \Phi_{\cdot i}$ and suppose that the researcher starts with the belief that

$$
\phi_i \sim \mathcal{N}\big(\underline{\phi}_i, \, D_i \underline{V}_i^\phi\big) \tag{A.16}
$$

with $\underline{\phi}_i = 0$. Because the macroeconomic variables are in log-level we let for $i = 1, \dots, n_y$:

$$[\underline{\phi}_i]_j = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases} \qquad j = 1, \dots, k. \tag{A.17}$$

The $k \times k$ prior covariance matrix $\underline{V}_i^{\phi}$ is assumed to be diagonal with elements $l = 1, \dots, k$:

$$[\underline{V}_i^{\phi}]_{ll} = \begin{cases} \frac{1}{\lambda_1}\frac{1}{\underline{s}_i^2 h^{\lambda_4}} & \text{for coeff. on the } h\text{-th lag if vars } (i,j) \text{ belong to same block} \\ \frac{1}{\lambda_1}\frac{1}{\lambda_2 \underline{s}_i^2 h^{\lambda_4}} & \text{for coeff. on the } h\text{-th lag if var } i \text{ belongs to } Y \text{ and } j \text{ belongs to } a \\ \frac{1}{\lambda_1}\frac{1}{\lambda_3 \underline{s}_i^2 h^{\lambda_4}} & \text{for coeff. on the } h\text{-th lag if var } i \text{ belongs to } a \text{ and } j \text{ belongs to } Y \\ \frac{1}{\lambda_5} & \text{for the intercept} \end{cases}$$

We now turn the prior for ϕ_i into a prior for $B_{\cdot i}$, utilizing the relationship between the quasi-structural-form coefficients, B, and the reduced form coefficients, Φ. For the coefficients on lagged elements of y_t we obtain:

$$[B_h]_{ij} = [\Phi_h]_{ij} + \sum_{l=1}^{i-1} A_{il}[\Phi_h]_{lj}. \tag{A.18}$$

Likewise, the $n \times 1$ vector of intercepts, B_0, is related to the reduced form intercept via

$$[B_0]_i = [\Phi_0]_i + \sum_{l=1}^{i-1} A_{il}[\Phi_0]_l. \tag{A.19}$$

Taking expectations of (A.18) and (A.19), and using $\mathbb{E}[A_{ij}] = 0$, we deduce that

$$\mathbb{E}\big[[B_h]_{ij}\big] = \mathbb{E}\big[[\Phi_h]_{ij}\big], \quad \mathbb{E}\big[[B_0]_i\big] = \mathbb{E}\big[[\Phi_0]_i\big] \tag{A.20}$$

We use a prior covariance matrix $\underline{V}_i^B$ that is diagonal. The entries on the diagonal are specified as follows: we first express the variance of a generic element $[B_h]_{ij}$ in terms of the variances of A_{ij} and $[\Phi_h]_{ij}$:

$$\begin{aligned} \mathbb{V}\big[[B_h]_{ij}\big] &= \mathbb{E}\big[\mathbb{V}\big([B_h]_{ij}\big)|A\big] + \mathbb{V}\big[\mathbb{E}\big([B_h]_{ij}\big)|A\big] \\ &= \mathbb{E}\left[\mathbb{V}\big([\Phi_h]_{ij}\big) + \sum_{l=1}^{i-1} A_{il}^2 \mathbb{V}\big([\Phi_h]_{lj}\big)\right] + \mathbb{V}\left[\mathbb{E}([\Phi_h]_{ij}) + \sum_{l=1}^{i-1} A_{il}\mathbb{E}([\Phi_h]_{lj})\right] \\ &= \mathbb{V}\big([\Phi_h]_{ij}\big) + \sum_{l=1}^{i-1} \mathbb{V}(A_{il})\mathbb{V}\big([\Phi_h]_{lj}\big) + \sum_{l=1}^{i-1} \mathbb{V}(A_{il})\big(\mathbb{E}([\Phi_h]_{lj})\big)^2. \end{aligned}$$

The same calculation for the variance of the intercept leads to:

$$\begin{aligned}
\mathbb{V}\big[[B_0]_i\big] &= \mathbb{E}\big[\mathbb{V}\big([B_0]_i\big)|A\big] + \mathbb{V}\big[\mathbb{E}\big([B_0]_i\big)|A\big] \\
&= \mathbb{E}\left[\mathbb{V}\big([\Phi_0]_i\big) + \sum_{l=1}^{i-1} A_{il}^2 \mathbb{V}\big([\Phi_0]_l\big)\right] + \mathbb{V}\left[\mathbb{E}([\Phi_0]_i) + \sum_{l=1}^{i-1} A_{il}\mathbb{E}([\Phi_0]_l)\right] \\
&= \mathbb{V}\big([\Phi_0]_i\big) + \sum_{l=1}^{i-1} \mathbb{V}(A_{il})\mathbb{V}\big([\Phi_0]_l\big) + \sum_{l=1}^{i-1} \mathbb{V}(A_{il})\big(\mathbb{E}([\Phi_0]_l)\big)^2.
\end{aligned}$$

To arrange the first np $\mathbb{V}\big[[B_h]_{ij}\big]$ terms on the diagonal of the $k \times k$ matrix $\underline{V}_i^B$ we use the index function

$$f(j,h) = (h-1)n + j. \tag{A.21}$$

Here h corresponds to the lag and j is the index for the element of the y_{t-h} vector. Using the definition of the index function, the expressions for $\mathbb{V}[A_{il}]$ and $\mathbb{E}([\Phi_h]_{lj})$ from the Normal distribution in (A.14), and the expression for $\mathbb{V}\big([\Phi_h]_{lj}\big)$ from the Normal distribution in (A.16), we can write

$$\mathbb{V}\big[[B_h]_{ij}\big] = D_i[\underline{V}_i^{\phi}]_{f(j,h)} + \sum_{l=1}^{i-1} \frac{D_i}{\underline{s}_l^2}\left[D_l[\underline{V}_l^{\phi}]_{f(j,h)} + [\underline{\phi}_l]^2_{f(j,h)}\right].$$

For the intercept in equation i we obtain

$$\mathbb{V}\big[[B_0]_i\big] = D_i\frac{D_i}{\lambda_5} + \sum_{l=1}^{i-1}\frac{D_iD_l}{\underline{s}_l^2\lambda_5}.$$

We now replace the variance parameter D_l by the hyperparameter $\underline{s}_l^2$. This ensures that $\underline{V}_i^B$ is not a function of the (unknown) variance parameter D_i and simplifies posterior calculations. Using

$$\mathbb{V}\big[[B_h]_{ij}\big] = D_i[\underline{V}_i^B]_{f(j,h)} \quad \text{and} \quad \mathbb{V}\big[[B_0]_i\big] = D_i[\underline{V}_i^B]_{np+1}$$

we obtain

$$[\underline{V}_i^B]_{f(j,h)} = [\underline{V}_i^{\phi}]_{f(j,h)} + \sum_{l=1}^{i-1}\left([\underline{V}_l^{\phi}]_{f(j,h)} + \frac{1}{\underline{s}_l^2}[\underline{\phi}_l]^2_{f(j,h)}\right). \tag{A.22}$$

$$[\underline{V}_i^B]_{np+1} = \frac{1}{\lambda_5} + \sum_{l=1}^{i-1}\frac{1}{\lambda_5} = \frac{i}{\lambda_5}. \tag{A.23}$$

Table A-1: Hyperparameters for VAR Prior

Parameter	Description
$\underline{\nu} = 2n$	Degrees of freedom for IG distribution
$\underline{s}_i = \text{StDev}(W_i)$	Shape para for IG; use sample standard dev.
λ_1	Overall precision of prior
λ_2	Relative precision for a to Y transmission
$\lambda_3 = 1$	Relative precision for Y to a transmission
$\lambda_4 = 2$	Decay rate for prior variance on lags
$\lambda_5 = 0.001$	Relative precision for intercept

Summary. The overall prior takes the form (A.10). The prior for D_i is given by (A.12). The prior for β_i is obtained by combining (A.14) with (A.15), where mean and variance are given in (A.20) and (A.22), respectively. The hyperparameters for the prior are summarized in Table A-1. We set $\underline{s}_i$ equal to the sample standard deviation of W_i.

B.2 Posterior Sampling and MDD

Model and prior are set up so that the coefficients can be estimated equation by equation:

$$p(W, \beta, D) = \prod_{i=1}^{N} \left((2\pi D_i)^{-1/2} \exp\left\{ -\frac{1}{2D_i}(W_i - Z_i\beta_i)'(W_i - Z_i\beta_i) \right\} p(\beta_i|D_i)p(D_i). \right) \tag{A.24}$$

Because the prior is conjugate, the posterior stays in the NIG family. It takes the form

$$\beta_i|(D_i, W_i) \sim \mathcal{N}\big(\bar{\beta}_i, D_i\bar{V}_i^{\beta}\big), \quad D_i \sim IG\big(\bar{\nu}_i, \bar{S}_i\big), \tag{A.25}$$

Instead of working with covariance matrices, it is more efficient to work with precision matrices. Define:

$$\underline{P}_i^{\beta} = \big(\underline{V}_i^{\beta}\big)^{-1}, \quad \bar{P}_i^{\beta} = \big(\bar{V}_i^{\beta}\big)^{-1}.$$

The updating equations for the posterior take the form

$$
\begin{aligned}
\bar{P}_i^{\beta} &= \underline{P}_i^{\beta} + Z_i'Z_i \\
\bar{\beta}_i &= \left(\bar{P}_i^{\beta}\right)^{-1}\left(\underline{P}_i^{\beta}\underline{\beta}_i + Z_i'W_i\right) \\
\bar{\nu}_i &= \underline{\nu}_i + T/2 \\
\bar{S}_i &= \underline{S}_i + \frac{1}{2}\left(W_i'W_i + \underline{\beta}_i'\underline{P}_i^{\beta}\underline{\beta}_i - \bar{\beta}_i'\bar{P}_i^{\beta}\bar{\beta}_i\right).
\end{aligned}
$$

When using the JK instruments, we set the coefficients on the lags and the intercept equal to zero: $B_i = 0$ for $i = 1, 2$. Thus, $\beta_1 = 0_{(np+1)\times 1}$ and $\beta_2 = [A_{2,1}, 0_{1\times(np+1)}]'$. The updating equations for the posterior change as follows. For $i = 1$:

$$
\begin{aligned}
\bar{P}_1^{\beta} &= N/A \\
\bar{\beta}_1 &= 0_{(np+1)\times 1} \\
\bar{\nu}_1 &= \underline{\nu}_i + T/2 \\
\bar{S}_1 &= \underline{S}_1 + \frac{1}{2}W_1'W_1.
\end{aligned}
$$

For $i = 2$ we are regressing W_{2t} on the single regressor W_{1t}. Partition $\beta_2 = [\beta_{1,2}, 0_{1\times(np+1)}]'$ and denote the precision associated with the first element of the β_2 vector by $[P_1^{\beta}]_{11}$. Then, we can write the updating equations as

$$
\begin{aligned}
[\bar{P}_2^{\beta}]_{11} &= [\underline{P}_2^{\beta}]_{11} + W_1'W_1 \\
\bar{\beta}_{2,1} &= \left([\bar{P}_2^{\beta}]_{11}\right)^{-1}\left([\underline{P}_2^{\beta}]_{11}\underline{\beta}_{2,1} + W_1'W_2\right) \\
\bar{\nu}_2 &= \underline{\nu}_2 + T/2 \\
\bar{S}_2 &= \underline{S}_2 + \frac{1}{2}\left(W_2'W_2 + [\underline{P}_2^{\beta}]_{11}\underline{\beta}_{2,1}^2 - [\bar{P}_2^{\beta}]_{11}\bar{\beta}_{2,1}^2\right).
\end{aligned}
$$

The MDD can be computed analytically as follows:

$$
\begin{aligned}
\ln p(W) &= -\frac{Tn}{2}\ln(2\pi) + \sum_{i=1}^{n}\Bigg[\frac{1}{2}\left(\ln|\underline{P}_i^{\beta}| - \ln|\bar{P}_i^{\beta}|\right) \\
&\quad + \underline{\nu}_i \ln|\underline{S}_i| - \bar{\nu}_i \ln|\bar{S}_i| - \ln\Gamma(\underline{\nu}_i) + \ln\Gamma(\bar{\nu}_i)\Bigg],
\end{aligned}
\tag{A.26}
$$

with the understanding that for $i = 1$ the expression $\ln|\underline{P}_i^{\beta}| - \ln|\bar{P}_i^{\beta}| = 0$ and for $i = 2$ it gets replaced by $\ln[\underline{P}_i^{\beta}]_{11} - \ln[\bar{P}_i^{\beta}|]_{11}$.

C Data Used in the Empirical Analysis

Aggregate Data. Following Jarocinski and Karadi (2020) we use six monthly macroeconomic variables in the empirical model: (i) the monthly average of the one-year constant-maturity Treasury yield serves as the monetary policy indicator. (ii) The monthly average of the S&P 500 stock price index in log levels. (iii,iv) Real GDP and GDP deflator in log levels interpolated to monthly frequency based on Stock and Watson (2010). (v) The excess bond premium (EBP) as indicator of financial conditions. (vi) An aggregate employment rate constructed from the micro data (see below). The functional VAR with micro-level consumption data includes an additional aggregate variable: real personal consumption expenditures per capita in log levels. We use Personal Consumption Expenditures (*PCE* from FRED) and divide it by population level (*CNP16OV* from FRED) to get per capita values. Then we use GDP deflator to get the real values.

Micro-Level Earnings Data. The CPS raw data are downloaded from
http://www.nber.org/data/cps_basic.html.
The raw data files are converted into STATA using the do-files available at:
http://www.nber.org/data/cps_basic_progs.html.
We use the series PREXPLF ("Experienced Labor Force Employment"), which is the same as in the raw data, and the series PRERNWA ("Weekly Earnings"), which is constructed as PEHRUSL1 ("Hours Per Week at One's Main Job") times PRHERNAL ("Hourly Earnings") for hourly workers, and given by PRWERNAL for weekly workers. STATA dictionary files are available at:
http://www.nber.org/data/progs/cps-basic/

We pre-process the cross-sectional data as follows. We drop individuals if (i) the employment indicator is not available; and (ii) if they are coded as "employed" but the weekly earnings are missing. In addition, we re-code individuals with non-zero earnings as employed and set earnings to zero for individuals that are coded as not employed. A CPS-based unemployment rate is computed as the fraction of individuals that are coded as not employed. By construction this is one minus the fraction of individuals with non-zero weekly earnings, which is used to normalize the cross-sectional density of earnings. It turns out that the CPS-based unemployment rate tracks the aggregate unemployment rate (*UNRATE* from FRED) very closely. The levels of the two series are very similar, but the CPS unemployment rate exhibits additional high-frequency fluctuations, possibly due to seasonals that have been removed from the aggregate unemployment rate.

Micro-Level Consumption Data. We use public use microdata (PUMD) from the Consumer Expenditure Survey (CEX) conducted by the Bureau of Labor Statistics (BLS). Quarterly expediture (by respondent ID and expenditure category) are computed as the sum over three months in a quarter. We divide the expenditures by the number of individuals with age 16 and over belonging to the consumption unit (i.e. household or family) to obtain per capita expenditures. Because CEX per capita expenditures capture less than 50% of NIPA per capita expenditures we rescale them as follows. Let C_t be aggregate per capita consumption from NIPA. We calculate $\frac{1}{T}\sum_{t=1}^{T} \text{median}(c_{1t}, \dots, c_{Nt})/C_t \approx 0.46$. We then define $z_{it} = c_{it}/(0.46 \cdot C_t)$. Thus, if $z_{it} = 1$ the individual approximately consumes at the level of aggregate consumption per capita. Finally, as for the earnings data, we apply the inverse hyperbolic sine transformation to obtain x_{it}.

Transformation of Micro Data. We transform the micro data (earnings-GDP ratio, or consumption relative to aggregate per capita consumption), denoted by z below, using the inverse hyperbolic sine transformation, which is given by

$$x = g(z|\theta) = \frac{\ln(\theta z + (\theta^2 z^2 + 1)^{1/2})}{\theta} = \frac{\sinh^{-1}(\theta z)}{\theta} \tag{A.27}$$

wth $\theta = 1$. Note that $g(0|\theta) = 0$ and $g^{(1)}(0|\theta) = 1$, that is, for small values of z the transformation is approximately linear. For large values of z the transformation is logarithmic:

$$g(z|\theta) \approx \frac{1}{\theta}\ln(2\theta z) = \frac{1}{\theta}\ln(2\theta) + \frac{1}{\theta}\ln(z).$$

The inverse of the transformation takes the form

$$z = g^{-1}(x|\theta) = \frac{1}{\theta}\sinh(\theta x) = \frac{1}{2\theta}(e^{\theta x} - e^{-\theta x}).$$

Most of the calculations in the paper are based on $p_x(x)$. But in some instances, it is desirable to report for $p_z(z)$. From a change of variables (omitting the θ), we get

$$p_z(z) = p_x(g(z))|g'(z)|,$$

where

$$g'(z) = \frac{1 + \frac{\theta z}{(\theta^2 z^2 + 1)^{1/2}}}{\theta z + (\theta^2 z^2 + 1)^{1/2}} = \frac{1}{(\theta^2 z^2 + 1)^{1/2}}.$$

Whenever we do convert the estimated densities back from z to x, we recycle the density evaluations at x_j. Thus, we evaluate $p_z(z)$ for grid points $z_j = g^{-1}(x_j)$, which leads to

$$p_z(z_j) = p_x(x_j)\left|g'(g^{-1}(x_j))\right|,$$

where

$$\left|g'(g^{-1}(x_j))\right| = \frac{1}{\left(\frac{1}{4}(e^{\theta x_j} - e^{-\theta x_j})^2 + 1\right)^{1/2}} = \frac{2}{\left(e^{2\theta x_j} + e^{-2\theta x_j} + 2e^{2\theta x_j} e^{-2\theta x_j}\right)^{1/2}} = \frac{2}{e^{\theta x_j} + e^{-\theta x_j}}.$$