Measuring the Effects of Aggregate Shocks on Unit-Level Outcomes and Their Distribution

Stephanie Ettmeier*
University of Bonn

Chi Hyun Kim
University of Bonn

Frank Schorfheide
University of Pennsylvania, CEPR, PIER, NBER

This Version: October 25, 2024
Very preliminary and incomplete. Please do not circulate.

Abstract

This paper studies the effect of aggregate shocks on micro-level outcomes. We develop and estimate a cross-sectional units vector autoregression (csuVAR) that combines aggregate variables with unit-level outcomes, earnings in our application. The csuVAR also allows us to reconstruct the cross-sectional distribution from the unit-level outcomes. We contrast the csuVAR with a functional VAR model (fVAR) that is designed to directly track the evolution of macroeconomic aggregates and a cross-sectional distribution, but not individual units. In an empirical application we examine the effect of productivity shocks on the unit-level labor earnings dynamics in Germany, using a panel data set constructed from the Sample of Integrated Labour Market Biographies (SIAB) published by Institute for Employment Research (IAB) of the German Federal Employment Agency. (JEL C11, C32, C52, E32)

Key words: Aggregate Shock, Earnings Distribution, Functional Model, IAB Data, Panel Data Analysis, Structural Vector Autoregression

* Correspondence: S. Ettmeier, C.H. Kim: Institute of Finance and Statistics, Department of Economics, University of Bonn, Adenauerallee 24-26, 53113 Bonn, Germany. Email: ettmeier@uni-bonn.de and ckim@uni-bonn.de. F. Schorfheide: Department of Economics, University of Pennsylvania, Philadelphia, PA 19104-6297, USA. Email: schorf@ssc.upenn.edu. We thank participants at various seminars and conferences for helpful comments and suggestions.

1 Introduction

Traditionally, business cycle fluctuations and the effects of macroeconomic policy interventions have been predominantly 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 representative agent dynamic stochastic equilibrium (DSGE) 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 macroeconomic shocks.

The existing literature has considered two related, but distinct questions. First, what is the effect of an aggregate shock on the cross-sectional distribution of a micro-level outcome x_{it} for unit i in period t? Second, how does a particular cross-sectional unit (household, firm, etc.) or group of units respond to an aggregate shock? In the case of a monetary or fiscal policy shock, the answer to the first question provides guidance to policy makers who are concerned about the effect of their actions on the income, wealth, or consumption distribution in the economy. Answers to the second question shed light on the shock propagation mechanism.

From a methodological perspective, the first question can be answered with a functional VAR (fVAR) model, that combines a vector of macroeconomic variables with (the log density of) a cross-sectional distribution; for instance, the one developed in Chang, Chen, and Schorfheide (2024), henceforth CCS. A repeated cross-section of unit-level observations is sufficient to estimate an fVAR. The data and modeling requirements to answer the second question is more stringent: panel data are needed, and one has to develop time series models that can track unit-level histories and capture the cross-sectional heterogeneity.

Our paper makes two main contributions: first, we develop a novel model that combines a VAR for macroeconomic aggregates with a panel data component that tracks unit-level outcomes. The two components are connected as follows: the cross-sectional units respond to aggregate outcomes. Morever, the aggregate variables are a function of the lagged cross-sectional distribution of micro-level outcomes. We refer to this model as cross-sectional units VAR, henceforth csuVAR, discuss its properties, and provide an estimation strategy. Second, we estimate the csuVAR and the fVAR based on a random sample from a German administrative data set that contains information about unit-level labor earnings. From the estimated fVAR, we can obtain an impulse response function (IRF) for the cross-sectional distribution of earnings to a labor productivity shock. From the estimated csuVAR, we construct unit-level earnings responses, which can then be converted into the response of the cross-sectional distribution, and compared to the fVAR responses.

The csuVAR is tailored to the data set that is used in the empirical analysis. We use detailed and high-frequency micro-level information entailed in the Sample of Integrated Labour Market Biographies (SIAB) published by Institute for Employment Research (IAB) of the German Federal Employment Agency. From the raw data we are able to construct a panel data set at quarterly frequencies. Individuals i can be in one of three states: they can be employed (E), unemployed (U), or they can be out of the sample (O) because either they were randomly replaced or they decided to leave the sample. For each unit i we observe the labor earnings x_{it} conditional on being employed and the employment state (E, U, or O). Thus, the csuVAR not only needs to determine the earnings dynamics x_{it} but also the employment state. There are two sets of latent variables in the csuVAR: the time-dependent state-transition probabilities, and the cross-sectional densities of x_{it} that feed back into the law of motion of the macroeconomic aggregates. We provide an asymptotic argument (number of cross-sectional units tends to infinity) that allows us to replace the transition probabilities by empirical frequencies, and to replace the unobserved cross-sectional densities, by coefficient estimates, for finite-dimensional sieve approximations of the log densities, similar to what is done in the fVAR.

An important aspect of the csuVAR modeling is to capture the heterogeneity at the micro-level. We pursue a Bayesian approach which specifies a parametric correlated random effects (CRE) distribution for the heterogeneous coefficients that determine the unit-level earnings dynamics. The cross-sectional information in the panel data set identifies the hyperparameters associated with the CRE distribution. The CRE distribution, in turn, implicitly serves as a prior distribution in unit-level time series regressions that determine the heterogeneous parameters. Our setup is very similar to the one used in recent work on forecasting with Bayesian panel data models, such as, Liu (2023) and Liu, Moon, and Schorfheide (2023).

Our paper is related to several strands of literature. The fVAR framework is taken from CCS, who use it to estimate the effect of productivity shocks on the earnings distribution in the U.S. using data from the Current Population Survey. The framework also has been recently used by Chang and Schorfheide (2022) to study the effects of monetary policy shocks, and by Ettmeier (2023) to study the distributional effects of aggregate fiscal policy shocks.

As mentioned previously, while the fVAR can be estimated based on the repeated cross sections, the csuVAR requires panel data. For many countries, including the U.S., highquality 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. The existing work utilizing administrative data has mostly used panel local projections (LP) and coefficient heterogeneity takes the form of observed group heterogeneity, meaning the researcher assumes that the response of say, income or consumption to an aggregate shock, is identical for units that belong to a particular group. For instance, Holm, Paul, and Tischbirek (2021) use administrative panel data from Norway and define groups in terms of liquid asset distribution).[1] 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. While the panel LP approach provides group specific responses to aggregate shocks, it is challenging to aggregate the estimates to responses of the cross-sectional distribution, because typically a large fraction of heterogeneity is ignored and group membership may be correlated with unit-level outcomes.

Compared to the panel LP analysis we raise the bar and build a model that combines aggregate time series with a panel model and captures the interactions between the macro and micro level. We examine to what extent the fVAR and the csuVAR modeling approaches deliver similar answers to questions about the effect of aggregate shocks on cross-sectional distributions. The key challenges for any panel approach is 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.

There exists a large literature at the intersection of labor economics and macroeconomics on the estimation of idiosyncratic earnings processes. Recent contributions include Guvenen (2007, 2009), Browning, Ejrnjes, and Alvarez (2010), Hryshko (2012), Browning and Ejrnjes (2013), and Hoffmann (2019). Such an earnings process is an important part of the micro-level component of our csuVAR. The income dynamics literature emphasizes the decomposition of income in a permanent and transitory compoment and favors unobserved components models, in part to rationalize consumption decisions. In this regard our earnings process is fairly simple as it evolves according to an AR(1) process driven by a single idiosyncratic shock alongside some aggregate variables.

There also exists a literature on how aggregate shocks affect the estimation of panel data models, e.g., Hahn, Kuersteiner, and Mazzocco (2020). These authors focus on the

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

fact that aggregate shocks affect units' decisions and inference on cross-sectional data may fail to correctly account for decision making of rational agents facing aggregate uncertainty. We exploit the time series variation in the aggregate variables, to estimate how agents at the micro level respond to aggregate shocks. Moreover, we use timing assumptions common in the structural VAR literature to identify the structural shock. In particular, we assume that aggregate outcomes respond to the cross-sectional distribution with a one-period lag, whereas unit-level outcomes can respond to aggregate shocks contemporaneously.

The remainder of this paper is organized as follows: Section 2 provides a stylized example of the interaction of macro- and micro-level dynamics. In Section 3 we review the functional VAR (fVAR) framework of CCS. The csuVAR is introduced in Section 4. It is tailored to some specific features of the data set used in the empirical analysis. The model consists of two interlinked parts: one that describes the evolution of the macroeconomic aggregates and one that determines the micro-level dynamics. The estimation of the csuVAR and the computation of impulse response functions (IRFs) are discussed in Section 5. We consider various generalizations of the baseline model in Section 6. The empirical analysis is presented in Section 7 and Section 8 concludes. Formal derivations, further information about the fVAR and csuVAR estimation, and details about the data sets used in the empirical analysis are relegated to an Online Appendix.

2 Interaction of Macro- and Micro-level Dynamics

To fix ideas, we now present a stylized model that captures the interaction of macro-level and micro-level dynamics. Let Υ_t be an $n_y \times 1$ vector of macroeconomic aggregates that evolves according to a first-order vector autoregression (VAR):

$$\Upsilon_t = B_{vv}\Upsilon_{t-1} + \int B_{vx}(\tilde{x})[\ln p^x_{t-1}(\tilde{x})]d\tilde{x} + u_{v,t}, \quad u_{v,t} \overset{iid}{\sim} p^u(u_t). \tag{1}$$

The non-standard feature of this specification is that aggregate outcomes also depend on the lagged cross-sectional density $p^x_{t-1}(x)$ of a cross-sectional variable x_{it} which will be individual-level income in the application in this paper. The kernel $B_{vx}(\tilde{x})$ is used to define an integral operator that maps the log cross-sectional density into an $n \times 1$ vector.

Suppose that at the micro-level the unit-level variable x_{it} follows the law of motion

$$x_{it} = \lambda_{i1}\Upsilon_t + \lambda_{i2}\Upsilon_{t-1} + \phi_{xx}x_{it-1} + \eta_{it}, \quad \eta_{it} \overset{iid}{\sim} p^\eta(\eta). \tag{2}$$

Here x_{it} responds to contemporaneous (Υ_t) and lagged (Υ_{t-1}) aggregate conditions, and it depends on its own lag (x_{it-1}). Moreover, it is assumed that the responses to the aggregate conditions are heterogeneous across units. For now we assume a random effects setting with

$$(\lambda_{i1}, \lambda_{i2}) \overset{iid}{\sim} p_\lambda(\lambda_1, \lambda_2). \tag{3}$$

If one assumes that the cross-sectional distribution of x_{it-1} was $p^x_{t-1}(x)$, then we can combine (2) and (3) to deduce that (dropping the i subscripts)

$$p^x_t(x) = \int \left[\int p_\eta(x - \lambda_1 \Upsilon_t - \lambda_2 \Upsilon_{t-1} - \phi_{xx}\tilde{x}) p_\lambda(\lambda_1, \lambda_2) d(\lambda_1, \lambda_2) \right] p_{t-1}(\tilde{x}) d\tilde{x}, \tag{4}$$

which, according to (1), affects aggregate conditions in period $t+1$ and can be interpreted as a general equilibrium effect. The system has a lower-triangular structure in that Υ_t affects x_{it} and $p^x_t(x)$ contemporaneously, but $p^x_t(x)$ affects Υ_t only with a one-period lag; see (1). We will maintain the triangular structure throughout the paper.

We will subsequently utilize two empirical strategies to estimate the system. First, we consider the fVAR framework developed in Chang, Chen, and Schorfheide (2024) to estimate (1) in combination with a log linearized version of (4). The fVAR approach tracks the evolution of the cross-sectional density $p^x_t(x)$, but not the histories x_{it} of the units i. The second approach, which we label cross-sectional unit VAR, or csuVAR in short, uses (2) to track the units x_{it} and combines it with the aggregate law of motion (1). Note that from the cross-section of x_{it} we can construct (an approximation of) $p^x_t(x)$, which is needed for the forward iteration.

3 A Functional VAR for Cross-Sectional Data

To make the exposition self-contained, we provide a summary of the functional framework developed in CCS, which has also been used in Chang and Schorfheide (2022) and Ettmeier (2023) to study the effects of monetary and fiscal policy shocks, respectively, using U.S. data. Following the notation in 2, we let Υ_t be a vector of aggregate variables. In our application Υ_t will be partitioned into

$$\Upsilon_t = \big[Y_t, UR_t\big], \tag{5}$$

where UR_t is the unemployment rate, and the vector Y_t stacks three macroeconomic time series: log labor productivity, log GDP, and the logarithm of the average cross-sectional labor earnings $\bar{x}_t$, which will be defined in (18) below.[2]

[2]The reason for separating the unemployment rate will become apparent in Section 4.

Moreover, we continue to use $p_t^x(x)$ to denote the cross-sectional density of x_{it}. Rather than working with the densities directly, we take the logarithmic transformation $\ell_t(x) = \ln p_t^x(x)$. 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 propagated using a linear law of motion and then *ex post* normalized to integrate to one in each period.

Starting point of the functional model is a nonlinear state-space representation. The measurement equation described in Section 3.1 connects the micro-level observations x_{it} to the unobserved log density $\ell_t(x)$. The state transition equation discussed in Section 3.2 provides a joint vector autoregressive law of motion for the macro variables and $\ell_t(x)$. In Section 3.3 we discuss some simplifications of the setup that lead to a finite-dimensional VAR representation. The computation of IRFs is described in Section 3.4.

3.1 Sampling and Measurement

We assume that in every period $t = 1, \dots, T$ an econometrician observes the macroeconomic aggregates Υ_t as well as a sample of N_t draws x_{it}, $i = 1, \dots, N_t$ 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. The measurement equation for the cross-sectional observations takes the form

$$x_{it} \overset{iid}{\sim} p_t^x(x) = \frac{\exp\{\ell_t(x)\}}{\int \exp\{\ell_t(x)\}dx}, \quad i = 1, \dots, N, \quad t = 1, \dots, T. \tag{6}$$

Subsequently, we denote the set of time t cross-sectional observations by $x_{1:N,t}$.

3.2 State Transition

The log density ℓ_t in (6) can be viewed as an infinite-dimensional state variable. We assume that Υ_t and ℓ_t evolve according to a joint autoregressive law of motion that we express in terms of deviations from a deterministic component $\big(\Upsilon_*, \ell_*(x)\big)$. 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 (Υ_*, ℓ_*) depend on t. Let

$$\Upsilon_t = \Upsilon_* + \tilde{\Upsilon}_t, \quad \ell_t = \ell_* + \tilde{\ell}_t. \tag{7}$$

The deviations from the deterministic component $\big(\tilde{\Upsilon}_t, \tilde{\ell}_t(x)\big)$ evolve jointly according to the following linear functional vector autoregressive (fVAR) law of motion:

$$\begin{aligned}\tilde{\Upsilon}_t &= B_{vv}\tilde{\Upsilon}_{t-1} + \int B_{vl}(\tilde{x})\tilde{\ell}_{t-1}(\tilde{x})d\tilde{x} + u_{v,t} \\ \tilde{\ell}_t(x) &= B_{lv0}(x)\tilde{\Upsilon}_t + B_{lv1}(x)\tilde{\Upsilon}_{t-1} + \int B_{ll}(x,\tilde{x})\tilde{\ell}_{t-1}(\tilde{x})d\tilde{x} + u_{l,t}(x).\end{aligned} \tag{8}$$

The second equation in (8) can be interpreted as a log-linearized (in function space) version of (4). It includes $\tilde{Y}_t$ on the right-hand side to capture the assumed triangular structure of the system. We assume that $u_{v,t}$ is mean-zero random vector with covariance Ω_{vv}. Moreover, $u_{l,t}(x)$ is a random element in a Hilbert space with covariance function $\Omega_{ll}(x,\tilde{x})$, something that is not present in (4) but needed to fit the data. We denote the covariance function for $u_{v,t}$ and $u_{l,t}(x)$ by $\Omega_{vl}(x)$. For the empirical analysis below we add more lags to the system. (8) can be viewed as the state-transition equation in a functional state-space model.

3.3 Three Simplifications

Equations (6), (7), and (8) define an infinite-dimensional nonlinear state-space model for the observables $\{\Upsilon_t, x_{1:N,t}\}_{t=1}^T$. Unfortunately, the estimation of this model is not practical, and we will simplify it in several 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

$$\ell_t^{(K)}(x) = \sum_{k=1}^K \alpha_{k,t}\zeta_k(x) = \big[\zeta_1(x), \ldots, \zeta_K(x)\big] \cdot \begin{bmatrix} \alpha_{1,t} \\ \vdots \\ \alpha_{K,t} \end{bmatrix} = \zeta'(x)\alpha_t \tag{9}$$

and $\ell_*^{(K)}(x) = \zeta'(x)\alpha_*$. Here $\zeta_1(x), \zeta_2(x), \ldots$ 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 $\tilde{\ell}^{(K)}(x) = \ell_t^{(K)}(x) - \ell_*^{(K)}(x)$.

To construct the measurement equation of the cross-sectional observations in (6), we define the K-dimensional vector of sufficient statistics

$$\bar{\zeta}(x_{1:N,t}) = \frac{1}{N}\sum_{i=1}^N \zeta(x_{it}).$$

This allows us to write a K'th order representation of the density of $x_{1:N,t}$:

$$\begin{aligned} p^{(K)}\left(x_{1:N,t}|\alpha_t\right) &= \exp\left\{N\mathcal{L}^{(K)}(\alpha_t|x_{1:N,t})\right\}, \\ \mathcal{L}^{(K)}(\alpha_t|x_{1:N,t}) &= \bar{\zeta}'(x_{1:N,t})\alpha_t - \ln\int \exp\left\{\zeta'(x)\alpha_t\right\}dx. \end{aligned} \tag{10}$$

We represent the kernels $B_{ll}(x,\tilde{x})$ and $B_{vl}(\tilde{x})$, the function $B_{lv}(x)$, and the functional innovation $u_{l,t}(x)$ that appear in the state-transition equation (8) as follows:

$$\begin{aligned} B_{ll}^{(K)}(x,\tilde{x}) &= \zeta'(x)B_{ll}\xi(\tilde{x}), \quad B_{vl}^{(K)}(x) = B_{vl}\xi(\tilde{x}) \\ B_{lv}^{(K)}(x) &= \zeta'(x)B_{lv}, \quad u_{l,t}^{(K)}(x) = \zeta'(x)u_{\alpha,t}, \end{aligned} \tag{11}$$

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_{vl} is of dimension $n_y\times K$, and B_{lv} is of dimension $K\times n_y$. Combining (7), (8), and (11) yields the following vector autoregressive system for the macroeconomic aggregates and the sieve coefficients (omitting K superscripts):

$$\begin{bmatrix} \Upsilon_t - \Upsilon_* \\ \alpha_t - \alpha_* \end{bmatrix} = \begin{bmatrix} B_{vv} & B_{vl}C_\alpha \\ B_{lv} & B_{ll}C_\alpha \end{bmatrix} \begin{bmatrix} \Upsilon_{t-1} - \Upsilon_* \\ \alpha_{t-1} - \alpha_* \end{bmatrix} + \begin{bmatrix} u_{v,t} \\ u_{\alpha,t} \end{bmatrix}, \tag{12}$$

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

$$u_t \sim \mathcal{N}(0,\Sigma). \tag{13}$$

The finite-dimensional state-space representation is given by the measurement equation (10) and the state-transition equation (12). To obtain a more compact notation, we define $W_t = [\Upsilon_t', \alpha_t']'$, absorb the matrix C_α into a general regression coefficient matrix Φ_1, and introduce an intercept Φ_0, which leads to

$$W_t = \Phi_0 + \Phi_1 W_{t-1} + u_t, \quad u_t \sim \mathcal{N}(0,\Sigma). \tag{14}$$

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 (10) around the maximum likelihood estimator (MLE) $\hat{\alpha}_t$. This approximation can be written as a linear Gaussian measurement equation:

$$\hat{\alpha}_t = \alpha_t + N^{-1/2}\eta_t, \quad \eta_t \sim \mathcal{N}\big(0, \hat{V}_t^{-1}\big), \tag{15}$$

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 (12) 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 (10). This leads to the sequence $\hat{\alpha}_t$.

2. Estimate a version of the VAR in (14), replacing the "true" sieve coefficients α_t in the definition of 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.

3.4 Impulse Response Function

Conditional on a set of fVAR parameters, the IRF can be computed as follows. First, starting from some initial conditions Υ_0^* and α_0^* we iterate (14) forward to create a baseline trajectory W_h^0 (all shocks are zero) and a shocked trajectory W_h^* (the macro shock of interest hits in period $h = 1$, all other shocks are zero). Taking the paths of α_h^0 and α_h^{sh} we obtain paths for the unnormalized log densities which can be converted into approximations of $p_h^0(x)$ and $p_h^{sh}(x)$ using (6).

4 A Cross-Sectional Units VAR

The csuVAR introduced in this section is a generalized version of the model described by (1) and (2) in Section 2. For the model to be suitable for the data set considered in the empirical analysis, we introduce a discrete state s_{it} that captures the employment status of unit i and describe the state transition in Section 4.1. In Section 4.2 we present the aggregate component of the csuVAR which in addition to the variables Y_t also provides a law of motion for the employment state transition probabilities. The cross-sectional unit dynamics are specified in Section 4.3. Finally, we derive an (approximate) likelihood function for the csuVAR in Section 4.4.

4.1 Micro-data Features and Some Definitions

Our application uses a random subsample of an administrative data set that encompasses all individuals who have ever been registered with the German social insurance system. A detailed description of the data set is provided in the Online Appendix. For each unit i we consider three states $s_{it} \in \{1, 2, 3\}$: employed (E), unemployed (U), and out-of-sample (O). An individual reaches the O state by leaving the labor force or by randomly being dropped from the subsample. In the former case, the individual may re-appear in the sample in a future period, whereas in the latter case the individual will not re-enter the sample. The introduction of the O state allows us to treat the panel as balanced, with the convention that the unit-level outcome x_{it}, which is labor earnings in our application, is unobserved for units in the O state. More precisely, in every period $t = 1, \ldots, T$ our raw data comprise units that are either employed or unemployed. We define the set of units $i = 1, \ldots, N$ in the panel used for the econometric analysis as the union of individuals that appeared as employed or unemployed for at least one of the periods $t = 1, \ldots, T$ in the raw data. In periods in which an individual is neither employed nor unemployed, and hence is not part of the raw data set, we assign the O state.[3]

Let $\mathbb{I}\{x = a\}$ be the indicator function that is equal to one if $x = a$ and equal to zero otherwise. Based on state counts we can define

$$E_t = \sum_{i=1}^{N} \mathbb{I}\{s_{it} = 1\}, \quad U_t = \sum_{i=1}^{N} \mathbb{I}\{s_{it} = 2\}, \quad O_t = \sum_{i=1}^{N} \mathbb{I}\{s_{it} = 3\}, \quad UR_t = \frac{U_t}{E_t + U_t}, \tag{16}$$

where UR_t is the unemployment rate in the sample. The (unobserved) state transition probabilities are defined as

$$\Pi_{jk,t} = \mathbb{P}_t\{s_{it} = k | s_{it-1} = j\}. \tag{17}$$

[3]Consider a unit i that enters the mail in period $t = 3$ as employed, then stays employed in $t = 4$, is unemployed in periods $t = 5$, and subsequently leaves the data set. Then $s_{i1} = s_{i2} = 3$ (O), $s_{i3} = s_{i4} = 1$ (E), $s_{i5} = 2$ (U), and $s_{it} = 3$ for $t > 5$ (O).

Raw labor earnings $\tilde{x}_{it}$ are only observed for individuals who are working, meaning that we observe $\tilde{x}_{it}\mathbb{I}\{s_{it=1}\}$. To remove a common trend from the unit-level income data and thereby reduce the spatial correlation of incomes induced by a business cycle, we define the average of the cross-sectional earnings as

$$\bar{x}_t = \frac{\sum_{i=1}^{N} \tilde{x}_{it}\mathbb{I}\{s_{it} = 1\}}{\sum_{i=1} \mathbb{I}\{s_{it} = 1\}} \tag{18}$$

and transform the unit-level earnings data as follows:

$$x_{it} = f\big(\tilde{x}_{it}/\bar{x}_t\big). \tag{19}$$

In the application $f(\cdot)$ is the inverse hyperbolic sine function, which is approximately linear for values close to zero and logarithmic for large positive values.

4.2 Aggregate Model Component

Our goal is to keep the aggregate model component of the csuVAR similar to that of the fVAR in Section 3. We assume that the aggregate variables follow a VAR law of motion. In our application, we use the same set of variables Y_t as for the fVAR: log labor productivity, log GDP, and the logarithm of the average cross-sectional labor earnings $\bar{x}_t$, defined in (18). However, we have to change the definition of Υ_t: (5) is replaced by

$$\Upsilon_t = \big[Y_t, \{\Pi_{j1,t}, \Pi_{j2,t}\}_{j=1}^{3}\big]. \tag{20}$$

Rather than including the unemployment rate UR_t we have to use the unobserved transition probabilities $\Pi_{jk,t}$ defined in (17) to be able to characterize the evolution of s_{it}. Given initial levels (E_0, U_0, O_0) the evolution of the unemployment rate is determined by the $\Pi_{jk,t}$s and can be excluded from Υ_t.[4] We assume that Υ_t evolves according to (the extension to a pth order process is straightforward)

$$\Upsilon_t = B_{vv}\Upsilon_{t-1} + \int B_{vx}(\tilde{x})\ell_{t-1}(\tilde{x})d\tilde{x} + u_{v,t}, \quad u_{v,t} \sim \mathcal{N}(0, \Sigma_{vv}), \tag{21}$$

where $\ell_t(x)$ is the (unnormalized) log density of x_{it}, see (6), and captures the feedback from the micro level to the aggregate level. To simplify the notation going forward, we combine B_{vv}, $B_{vx}(\cdot)$, and the non-redundant elements of Σ_{vv} in the (at this point infinite-dimensional) parameter vector θ_{VAR}.

[4]The use of probabilities instead of unnormalized probabilities in the vector Υ_t allowed for a more accurate tracking of the unemployment rate UR_t, defined in (16).

4.3 Micro-level Model Component

We make the simplifying assumption that the state-transitions at the micro-level only depend on the aggregate probabilities $\Pi_{jk,t}$ but not on unit-level characteristics. Thus, the only part of the model that remains to be specified is the labor income of the employed ($s_{it} = 1$). Define the selection matrix M_v that selects elements of the aggregate variables Υ_t and their lags Υ_{t-1}, i.e.,

$$Z_t = M_v \begin{bmatrix} 1 \\ \Upsilon_t \\ \Upsilon_{t-1} \end{bmatrix}.$$

We assume that the conditional density of earnings takes the following form

$$p(x_{it}|\cdot) = \begin{cases} p_{\mathcal{N}}\big(x_{it}|\gamma_i' Z_t + \rho x_{it-1}, \sigma_i^2\big) & \text{if } s_{it-1} = 1, s_{it} = 1 \\ p_{\mathcal{N}}\big(x_{it}|\gamma_{UE}' Z_t, \sigma_{UE}^2\big) & \text{if } s_{it-1} = 2, s_{it} = 1 \\ p_{\mathcal{N}}\big(x_{it}|\gamma_{OE}' Z_t, \sigma_{OE}^2\big) & \text{if } s_{it-1} = 3, s_{it} = 1 \end{cases}, \tag{22}$$

where $p_{\mathcal{N}}(x|\mu, \Sigma)$ is the probability density function for $X \sim \mathcal{N}(\mu, \Sigma)$. The assumption of conditional Normal distributions is preliminary and has been made for computational convenience. For the heterogeneous parameters of the E-to-E transitions we assume the following correlated random effects distribution:

$$\begin{aligned} p(\gamma_i|x_{i0}, \sigma_i^2, \xi) &= p_{\mathcal{N}}\big(\gamma_i|\underline{\gamma}_0 + \underline{\gamma}_s x_{i\tau_{i0}}, \sigma_i^2 \underline{V}_\gamma\big) \\ p(\sigma_i^2|\xi) &= p_{IG}\big(\sigma_i^2|\underline{\nu}, \underline{s}^2\big), \end{aligned} \tag{23}$$

where the period τ_{i0} is the first period in which unit i is employed. Moreover, $p_{IG}(\cdot)$ is the probability density function of an Inverse Gamma distribution, parameterized as scaled inverse χ^2 distribution in terms of degrees of freedom $\underline{\nu}$ and sum of squared residuals $\underline{s}^2$. Thus, we can define the vector of hyperparameters as

$$\xi = \big[\underline{\gamma}_0, \underline{\gamma}_s, \text{vech}(\underline{V}_\gamma), \underline{\nu}, \underline{s}^2\big],$$

where vech$(\cdot)$ is the vector half that collects the non-redundant elements of a symmetric matrix. We stack the homogeneous parameters in the vector

$$\theta_{csu} = \big[\rho, \gamma_{UE}, \sigma_{UE}^2, \gamma_{OE}, \sigma_{OE}^2\big].$$

If we combine (22) and (23) we can obtain a generalization of the law of motion of the cross-sectional density in (4). Because of the discrete states and the correlated random effects structure, the model implies that one has to keep track of the conditional distribution $p_t^x(x|x_{\tau_0})$ and the marginal distribution $p_t^{x_{\tau_0}}(x_{\tau_0})$, from which the marginal density $p_t(x)$ and its logarithm $\ell_t(x)$ can be obtained. In general, such a calculation is not practical because it involves high-dimensional integration and it is very sensitive to assumptions about micro-level dynamics and heterogeneity. Instead, we reconstruct $\ell_{t-1}(x)$ from the micro data, using the same technique as in Section 3. Details follow in the next subsection.

4.4 Likelihood Function

Data Generating Process (DGP). Let Π_t collect the transition probabilities $\Pi_{jk,t}$ and $\mathcal{D}_{it}$ collect the observables s_{it} and $x_{it}\mathbb{I}\{s_{it}=1\}$. We use $\Pi_{1:\tau}$ to denote the sequence $\{\Pi_1,\dots,\Pi_\tau\}$ and let

$$\mathcal{D}_{1:N,1:\tau} = \{\mathcal{D}_{11},\dots,\mathcal{D}_{N1},\dots,D_{1\tau},\dots,\mathcal{D}_{N\tau}\}.$$

Recall that the only observables are $Y_{1:T}$ and $\mathcal{D}_{1:N,1:T}$. All other objects are unobserved. Starting point is the joint density of $(Y_{1:T},\Pi_{1:T},\ell_{1:T},\mathcal{D}_{1:N,1:T},\gamma_{1:N},\sigma^2_{1:N})$ conditional on the homogeneous parameters $(\theta_{VAR},\theta_{csu},\xi)$ and the initial conditions (Y_0,ℓ_0). In a first-order vector autoregressive setting $(p=1)$ the joint density can be factorized as follows:

$$\begin{aligned}
&p\big(Y_{1:T},\Pi_{1:T},\ell_{1:T},\mathcal{D}_{1:N,1:T},\gamma_{1:N},\sigma^2_{1:N}|Y_0,\Pi_0,\ell_0,\theta_{VAR},\theta_{csu},\xi\big) \qquad (24)\\
&= \prod_{t=1}^{T}\Bigg\{\underbrace{p(Y_t,\Pi_t|Y_{t-1},\Pi_{t-1},\ell_{t-1},\theta_{VAR})}_{\text{VAR Part}}\times\underbrace{p(s_{1:N,t}|\Pi_t,\mathcal{D}_{1:N,t-1})}_{\text{Panel Part I}}\\
&\quad\times\underbrace{\left(\prod_{i=1}^{N}\big[p(\gamma_i,\sigma_i^2|x_{it},\xi)\big]^{\mathbb{I}\{t=\tau_{i0}\}}\big[p(x_{it}|Y_t,Y_{t-1},\mathcal{D}_{it-1},\gamma_i,\sigma_i^2,\theta_{csu})\big]^{\mathbb{I}\{s_{it}=1\}}\right)}_{\text{Panel Part II}}\\
&\quad\times\underbrace{\left(\prod_{t=1}^{T}\mathbb{I}\{\ell_t=f_t(\cdot)\}\right)}_{\text{Panel Part III}}\Bigg\}.
\end{aligned}$$

The *VAR Part* in (24) describes the evolution of the aggregate variables Υ_t, which is given by the VAR in (21). We assume a lower-triangular structure of the system, in which cross-sectional outcomes only enter with a lag, through the log density $\ell_{t-1}(x)$, but not contemporaneously. *Panel Part I* summarizes the evolution of the unit-level states s_{it}, which depends on the aggregate transition probabilities Π_t, determined in the *VAR Part.* For now, we assume that the state transition does not depend on unit-level features, meaning that entry and exit into employment is random. *Panel Part II* describes the conditional distribution of $x_{it}\mathbb{I}\{s_{it}=1\}$ which is given by (23). We assume that earnings conditional on the aggregate variables (Y_t, Y_{t-1}) are independent across i. Recall that τ_{i0} is the first period t in which unit i enters the employment state. At this point its unit-level parameters are determined which leads to the term $\big[p(\gamma_i,\sigma_i^2|x_{it},\xi)\big]^{\mathbb{I}\{t=\tau_{i0}\}}$. From the unit-level laws of motion and assumptions about coefficient heterogeneity, the cross-sectional densities ℓ_t are determined; see the example in (4). We denote this relationship by the function $f_t(\cdot)$. We summarize the key assumptions which will be relaxed in Section 6:

Assumption 1 *The data generating process satisfies the following conditions:*

- **(LT)** *The system is lower triangular in that the cross-sectional distribution of x does not affect Υ_t contemporaneously.*
- **(No-Sel)** *The evolution of s_{it} does not depend on unit-level characteristics.*
- **(P-CRE)** *Parametric correlated random effects: the distribution of (γ_i,σ_i^2) is parametric and conditional on the initial value $x_{i\tau_{i0}}$; see (23).*

Re-arranging Terms. To facilitate the estimation of the csuVAR, it is convenient to integrate out (γ_i,σ_i^2) and re-arrange the terms in (24) as follows:

$$
\begin{aligned}
& p\big(Y_{1:T},\Pi_{1:T},\ell_{1:T},\mathcal{D}_{1:N,1:T}|Y_0,\Pi_0,\ell_0,\theta_{VAR},\theta_{csu},\xi\big) \qquad (25)\\
&= \underbrace{\left(\prod_{t=1}^{T} p(Y_t,\Pi_t|Y_{t-1},\Pi_{t-1},\ell_{t-1},\theta_{VAR})\right)}_{\text{VAR Part}} \times \underbrace{\left(\prod_{t=1}^{T} p(s_{1:N,t}|\Pi_t,\mathcal{D}_{1:N,t-1})\right)}_{\text{Panel Part I}} \\
&\quad \times \underbrace{\left(\prod_{i=1}^{N}\int\prod_{t=1}^{T}\big[p(\gamma_i,\sigma_i^2|x_{it},\xi)\big]^{\mathbb{I}\{t=\tau_{i0}\}}\big[p(x_{it}|Y_t,Y_{t-1},\mathcal{D}_{it-1},\gamma_i,\sigma_i^2,\theta_{csu})\big]^{\mathbb{I}\{s_{it}=1\}}d(\gamma_i,\sigma_i^2)\right)}_{\text{Panel Part II}} \\
&\quad \times \underbrace{\left(\prod_{t=1}^{T}\mathbb{I}\{\ell_t=f_t(\cdot)\}\right)}_{\text{Panel Part III}}.
\end{aligned}
$$

If Π_t were observed and ℓ_t be approximated by a sieve as in (9) with known coefficient α_t, then the estimation of θ_{VAR} in (*VAR Part*) would be equivalent to estimating the first n_y equations of the fVAR in Section (4), except that Y_t is replaced by the larger vector Υ_t. *Panel Part I* describes the employment state transition for each unit i. The parameters θ_{csu} are concentrated in *Panel Part II* and can be estimated using dynamic panel data techniques. In the remainder of this subsection, we will discuss how to integrate out the unobserved transition probabilities Π_t and to determine ℓ_t from an approximation of the function $f_t(\cdot)$ based on the cross-sectional observations $x_{1:N,t}$.

Replacing Π_t by $\hat{\Pi}_t$. Suppose that the sequence ℓ_t is observed. Formally, Π_t is a latent variable and needs to be integrated out from (25). Because of Assumption 1(No-Sel) we can rewrite the *Panel Part I* expression as

$$p(s_{1:N,t}|\Pi_t, \mathcal{D}_{1:N,t-1}) = p(s_{1:N,t}|\Pi_t, s_{1:N,t-1}).$$

This density can be viewed as a measurement equation of a state-space model, and the *VAR Part* $p(Y_t, \Pi_t|Y_{t-1}, \Pi_{t-1}, \ell_{t-1}, \theta_{VAR})$ is the state-transition equation. For each period t, starting from a distribution $p(\Pi_{t-1}|Y_{1:t-1}, s_{1:N,1:t-1}, \ell_{1:t-2}, \theta_{VAR})$, the filter computes

$$\begin{aligned} & p(Y_t, \Pi_t|Y_{1:t-1}, s_{1:N,1:t-1}, \ell_{1:t-1}, \theta_{VAR}) \quad (26) \\ & \quad = \int p(Y_t, \Pi_t|Y_{t-1}, \Pi_{t-1}, \ell_{t-1}, \theta_{VAR}) p(\Pi_{t-1}|Y_{1:t-1}, s_{1:N,1:t-1}, \ell_{1:t-2}, \theta_{VAR}) d\Pi_{t-1} \\ & p(Y_t, s_{1:N,t}|Y_{1:t-1}, s_{1:N,1:t-1}, \ell_{1:t-1}, \theta_{VAR}) \quad (27) \\ & \quad = \int p(s_{1:N,t}|\Pi_t, s_{1:N,t-1}) p(Y_t, \Pi_t|Y_{1:t-1}, s_{1:N,1:t-1}, \ell_{1:t-1}, \theta_{VAR}) d\Pi_t \\ & p(\Pi_t|Y_{1:t}, s_{1:N,1:t}, \ell_{1:t-1}, \theta_{VAR}) \quad (28) \\ & \quad \propto \; p(s_{1:N,t}|\Pi_t, s_{1:N,t-1}) p(Y_t, \Pi_t|Y_{1:t-1}, s_{1:N,1:t-1}, \ell_{1:t-1}, \theta_{VAR}). \end{aligned}$$

(26) is the predictive density for (Y_t, Π_t), (27) is the predictive density for $s_{1:N,t}$, and (28) is the updating equation. The output of the filter can be used to express

$$\begin{aligned} & \int \left(\prod_{t=1}^{T} p(Y_t, \Pi_t|Y_{t-1}, \Pi_{t-1}, \ell_{t-1}, \theta_{VAR}) p(s_{1:N,t}|\Pi_t, \mathcal{D}_{1:N,t-1}) \right) d\Pi_{1:T} \quad (29) \\ & \qquad = \prod_{t=1}^{T} p(Y_t, s_{1:N,t}|Y_{1:t-1}, s_{1:N,1:t-1}, \ell_{1:t-1}, \theta_{VAR}). \end{aligned}$$

The two terms on the right-hand side of (28) can be interpreted as likelihood function and prior for Π_t and the left-hand side is the posterior. The log likelihood function can be approximated using a second-order Taylor approximation around the maximum likelihood estimator (MLE)

$$\hat{\Pi}_{jk,t} = \frac{N_{jk,t}}{\sum_{j=1}^{3} \hat{\Pi}_{jk,t}}, \quad (30)$$

where $N_{jk,t} = \sum_{i=1}^{N} \mathbb{I}\{s_{it-1} = j,\, s_{it} = k\}$ is the number of transitions from state j to k. The negative inverse Hessian evaluated at the MLE is proportional to $1/N$ which means that information accumulates at rate N. At the same time, the precision of the prior distribution is bounded by a function of Σ_{vv} in (21). Using standard Bayesian large sample arguments, the posterior distribution in (28) concentrates around $\hat{\Pi}_t$. In turn, in period $t+1$:

$$\begin{aligned} p(Y_{t+1}, \Pi_{t+1}|Y_{1:t}, s_{1:N,1:t}, \ell_{1:t}, \theta_{VAR}) &\approx p(Y_{t+1}, \Pi_{t+1}|Y_t, \hat{\Pi}_{t-1}, \ell_t, \theta_{VAR}) \\ p(Y_{t+1}, s_{1:N,t+1}|Y_{1:t}, s_{1:N,1:t}, \ell_{1:t}, \theta_{VAR}) &\approx p(Y_{t+1}, \hat{\Pi}_{t+1}|Y_t, \hat{\Pi}_t, \ell_t, \theta_{VAR}). \end{aligned}$$

Formally, one can show the following result:[5]

Theorem 1 *Suppose that Assumption 1 is satisfied. Then,*

$$\left| \sum_{t=1}^{T} \ln p(Y_t, s_{1:N,t}|Y_{1:t-1}, s_{1:N,1:t-1}, \ell_{1:t-1}, \theta_{VAR}) - \ln p(Y_{t+1}, \hat{\Pi}_{t+1}|Y_t, \hat{\Pi}_t, \ell_t, \theta_{VAR}) \right| \lesssim \frac{T}{N}.$$

The *Panel Part III* Term. The cross-sectional density ℓ_t is determined by the law of motion of x_{it} in Section 4.3. As discussed previously, the direct calculation of ℓ_t generally involves a complicated integration. Instead, we will construct ℓ_t from the cross sectional distribution of x_{it}. To fix ideas, consider the following example. Suppose that

$$x_{it} = \gamma_i(1-\rho) + \rho x_{it-1} + \eta_{it}, \quad \gamma_i \sim \mathcal{N}(0, \underline{V}_\gamma), \quad \eta_{it} \sim \mathcal{N}(0,1).$$

Let $\mu_{x,t} = 0$ and $v_{x,t} = \underline{V}_\gamma + 1/(1-\rho^2)$. The direct calculation leads to

$$\ell_t(x) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln v_{x,t} - \frac{1}{2v_{x,t}}(x - \mu_{x,t})^2. \tag{31}$$

Alternatively, we can use $x_{1:N,t}$ to compute sample mean and variance $\hat{\mu}_{x,t}$ and $\hat{v}_{x,t}$ and use it to construct the approximation $\hat{\ell}_t(x)$ by replacing population moments in (31) by sample moments. This generates an approximation error that vanishes as $N \longrightarrow \infty$.

In practice we use the same approach as in Section 3 to approximate ℓ_t. We use the K-dimensional sieve approximation $\ell_t^{(K)}(x) = \zeta'(x)\alpha_t$ in (9). Let $\hat{\alpha}_t$ be the MLE of α_t based on the sample $x_{1:N,t}$. We note that $\|\alpha_t - \hat{\alpha}_t\| \lesssim \sqrt{K/N}$; see Stone (1990). Provided that K

[5] We write $|f(N,K;\omega)| \lesssim \alpha_{K,N}$ to mean that there is a constant C such that for every small $\epsilon > 0$ we can find (K_0, N_0) such that $\mathbb{P}^\omega \{|f(N,K;\omega)| \leq C\alpha_{K,N}\} \geq 1-\epsilon$ for $K > K_0$ and $N > N_0$. Moreover, we write $f(N,K;\omega) \asymp 1 \pm \alpha_{K,N}$ to denote that $|f(N,K;\omega) - 1| \lesssim \alpha_{K,N}$.

grows sufficiently fast with N such that the approximation bias through the K-dimensional sieve is of strictly smaller order than the standard deviation of $\hat{\alpha}_t$. Let

$$\hat{\ell}_t^{(K)}(x) = \zeta'(x)\hat{\alpha}_t. \tag{32}$$

Then, uniformly for $x \in [\underline{x}, \bar{x}]$

$$\|\ell_t(x) - \hat{\ell}_t(x)\| = O_p(\sqrt{K/N}).$$

Theorem 2 *Suppose that Assumption 1 is satisfied. Then,*

$$\left| \sum_{t=1}^{T} \ln p(Y_{t+1}, \hat{\Pi}_{t+1}|Y_t, \hat{\Pi}_t, \ell_t, \theta_{VAR}) - \ln p(Y_{t+1}, \hat{\Pi}_{t+1}|Y_t, \hat{\Pi}_t, \hat{\ell}_t, \theta_{VAR}) \right| \lesssim T\sqrt{\frac{K}{N}}.$$

Approximate Likelihood Function. Based on Theorems 1 and 2 we estimate the csuVAR using the approximate likelihood function

$$\begin{aligned} &\hat{p}\big(Y_{1:T}, \hat{\Pi}_{1:T}, \hat{\ell}_{1:T}, \mathcal{D}_{1:N,1:T}|Y_0, \Pi_0, \hat{\ell}_0, \theta_{VAR}, \theta_{csu}, \xi\big) \qquad (33) \\ &= \underbrace{\left(\prod_{t=1}^{T} p(Y_t, \hat{\Pi}_t|Y_{t-1}, \Pi_{t-1}, \hat{\ell}_{t-1}, \theta_{VAR})\right)}_{\text{VAR Part}} \\ &\times \underbrace{\left(\prod_{i=1}^{N}\int \prod_{t=1}^{T} \big[p(\gamma_i, \sigma_i^2|x_{it}, \xi)\big]^{\mathbb{I}\{t=\tau_{i0}\}} \big[p(x_{it}|Y_t, Y_{t-1}, \mathcal{D}_{it-1}, \gamma_i, \sigma_i^2, \theta_{csu})\big]^{\mathbb{I}\{s_{it}=1\}} d(\gamma_i, \sigma_i^2)\right)}_{\text{Panel Part II}}, \end{aligned}$$

where $\hat{\ell}_t$ is given in (32).

5 csuVAR Estimation and IRF Computation

We now turn to the estimation of the csuVAR. We consider a Bayesian approach based on the approximate likelihood function in (33). The csuVAR is set up so that the aggregate model component and the micro-level component can be estimated separately. Details follow in Sections 5.1 and 5.2, respectively. Section 5.3 summarizes the main steps of the model estimation. The computation of impulse response functions (IRFs) is discussed in Section 5.4.

5.1 Estimation of Aggregate Model Component

To make the estimation of the aggregate component of the csuVAR in (21) operational, we proceed as in the case of the fVAR and replace the log density $\ell_{t-1}(x)$ and the kernel $B_{vx}(x)$ by K-dimensional approximations. Allowing for $p \geq 1$ lags we approximate (21) by

$$\Upsilon_t = \Phi_0 + \sum_{j=1}^{p} \left(\Phi_j^{v} \Upsilon_{t-j} + \Phi_j^{\alpha} \hat{\alpha}_{t-j} \right) + u_{v,t}, \quad u_{v,t} \sim \mathcal{N}(0, \Sigma_{vv}). \tag{34}$$

To make the aggregate part of the csuVAR structural, we assume that

$$u_{v,t} = \Sigma_{vv}^{tr} \Omega \epsilon_t, \tag{35}$$

where Σ_{vv}^{tr} is the lower triangular Cholesky factor of Σ_{vv} and Ω is an orthonormal matrix. In our application we will set $\Omega = I$ and focus on the response to the first shock. Note that (34) has the same structure as the finite-dimensional approximation of the fVAR (14) with $W_t = [Y_t', \hat{\alpha}_t']$. The only difference is that in the csuVAR the vector $\hat{\alpha}_t$ is not determined as part of the vector autoregressive law of motion. Hence, we use the same estimation method for (14) and (34). We follow the approach taken in Chang, Chen, and Schorfheide (2024) and Chang and Schorfheide (2022), which builds on Chan (2022). We sketch the basic idea in the context of the csuVAR and provide further details in the Online Appendix.

Likelihood Function. The structural VAR by (34) and (35) can be rewritten as follows:

$$A\Upsilon_t = \sum_{j=1}^{p} \left(B_j^{v} \Upsilon_{t-j} + B_j^{\alpha} \hat{\alpha}_{t-j} \right) + B_0 + \eta_t, \quad \eta_t = D^{1/2} \Omega \epsilon_t, \tag{36}$$

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 $A^{-1} D^{1/2} = \Sigma_{vv}^{tr}$, $A^{-1} B_j^{\alpha} = \Phi_j^{\alpha}$, and $A^{-1} B_j^{v} = \Phi_j^{v}$, $j = 0. \dots, p$. 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[6]

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

Moreover, let $k_i = p(n_v + n_\alpha) + i - 1$ and define the $k_i \times 1$ vectors

$$Z_{it} = \left[\tilde{Y}_{<i,t}', \Upsilon_{t-1}', \hat{\alpha}_{t-1}', \dots, \Upsilon_{t-p}', \hat{\alpha}_{t-p}', 1 \right], \quad \beta_i = \left[\mathcal{A}_i', B_{i\cdot,1}^{v\prime}, B_{i\cdot,1}^{\alpha\prime}, \dots, B_{i\cdot,p}^{v\prime}, B_{i\cdot,p}^{\alpha\prime}, B_{i\cdot,0}' \right]',$$

[6]In slight abuse of notation, we are changing the definition of "Z" (see also (22)) here to generically denote the regressors in the VAR equation.

where $B_{i\cdot,j}$ is the ith row of the matrix B_j. Finally, define Υ_i to be the $T \times 1$ vector with elements Υ_{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

$$\Upsilon_i = Z_i \beta_i + \eta_i. \tag{37}$$

Because A is lower-triangular with ones on the diagonal, the Jacobian associated with the change-of-variables from η_t to Υ_t in (36) is equal to one. In turn, the likelihood function for the system is the product of the likelihood functions for each variable i. Let $\beta = (\beta_1, \dots, \beta_n)$. Then:

$$p(\Upsilon|\beta, D) \propto \prod_{i=1}^{n} |D_i|^{-1/2} \exp\left\{ -\frac{1}{2D_i} (\Upsilon_i - Z_i\beta_i)'(\Upsilon_i - Z_i\beta_i) \right\}. \tag{38}$$

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

$$p(\beta, D|\lambda) = \prod_{i=1}^{n} p(\beta_i|D_i, \lambda) p(D_i|\lambda), \tag{39}$$

where λ is a vector of hyperparameters. For each pair (β_i, D_i) we use a Normal-Inverse Gamma ($\mathcal{N}IG$) distribution of the form

$$\beta_i|(D_i, \lambda) \sim \mathcal{N}\big(\underline{\beta}_i, D_i \underline{V}_i^{\beta}\big), \quad D_i|\lambda \sim IG\big(\underline{\nu}_i, \underline{S}_i\big). \tag{40}$$

This setup is very flexible, because it is straightforward to impose equation-specific restrictions. Further details are provided in Section 7 and the Online Appendix.

Posterior Sampling. The conjugate form of the prior implies that the posterior distribution of (β, D) also belongs to the $\mathcal{N}IG$ 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), $i = 1, \dots, n$ we can sample the parameters for each equation separately.

5.2 Micro-level Model Component

For the micro-level component we need to estimate the vector of homogeneous coefficients θ_{csu}, the heterogeneous coefficients $(\gamma_{1:N}, \sigma^2_{1:N})$, and the hyperparameters ξ of the correlated random effects distribution. The estimation is based on the *Panel Part II* in (33). For illustrative purposes assume that $T = 6$ and suppose that unit $i = \iota$ is unemployed in periods $t = 1, 2$, she is employed in periods $t = 3, 4, 5$, and subsequently leaves the panel. Then,

$$s_{\iota 1} = s_{\iota 2} = 2, \quad s_{\iota 3} = s_{\iota 4} = s_{\iota 5} = 1, \quad s_{\iota 6} = 3,$$

and $\tau_{\iota 0} = 3$. Hence, using (22) the contribution of unit ι to *Panel Part II* in (33) can be written as

$$\begin{aligned}
&\int \prod_{t=1}^{T} \big[p(\gamma_\iota, \sigma_\iota^2 | x_{\iota t}, \xi)\big]^{\mathbb{I}\{t=\tau_{\iota 0}\}} \big[p(x_{\iota t}|Y_t, Y_{t-1}, \mathcal{D}_{\iota t-1}, \gamma_\iota, \sigma_\iota^2, \theta_{csu})\big]^{\mathbb{I}\{s_{\iota t}=1\}} d(\gamma_\iota, \sigma_\iota^2) \qquad (41)\\
&= \; p_{\mathcal{N}}\big(x_{\iota 3}|\gamma_{UE}' Z_3, \sigma_{UE}^2\big) \int \prod_{t=4}^{5} p_{\mathcal{N}}\big(x_{\iota t}|\gamma_\iota' Z_t + \rho x_{\iota t-1}, \sigma_\iota^2\big) p(\gamma_\iota, \sigma_\iota^2 | x_{\iota 3}, \xi) d(\gamma_\iota, \sigma_\iota^2).
\end{aligned}$$

We deduce that for the estimation of $(\gamma_{UE}, \sigma_{UE}^2)$ and $(\gamma_{OE}, \sigma_{OE}^2)$ we can pool observations across units and time periods based on $(s_{it-1} = 2, s_{it} = 1)$ and $(s_{it-1} = 3, s_{it} = 1)$, respectively. Moreover, the unit-specific parameters (γ_i, σ_i^2) can be integrated out from the likelihood function, conditional on (ρ, ξ) for each unit separately. In the remainder of this subsection we provide further details on the Bayesian estimation.

Estimating $(\gamma_{UE}, \sigma_{UE}^2)$ and $(\gamma_{OE}, \sigma_{OE}^2)$. For the initial distribution of earnings conditional on entering the employment state $s_{it} = 1$ from unemployment $s_{it-1} = 2$ or from outside of the earnings panel $s_{it-1} = 3$ is given by the linear Gaussian regression models

$$x_{it} = \gamma_{\cdot E}' Z_t + u_{it}^{\cdot E}, \quad u_{it}^{\cdot E} \sim \mathcal{N}(0, \sigma_{\cdot E}^2), \qquad (42)$$

where the dot in the $E\cdot$ super- and subscripts is a placeholder for U or O; see (22). We use a conjugate $\mathcal{N}IG$ prior of the form

$$p(\gamma_{E\cdot}|\sigma_{E\cdot}^2) = p_{\mathcal{N}}\big(\gamma_{E\cdot}|\underline{\gamma}_{E\cdot}, \sigma_{E\cdot}^2 \underline{V}_\gamma^{E\cdot}\big), \quad p(\sigma_{E\cdot}^2) = p_{IG}\big(\sigma_{E\cdot}^2|\underline{\nu}_{E\cdot}, \underline{s}_{E\cdot}^2\big). \qquad (43)$$

The posterior distribution also belongs to the $\mathcal{N}IG$ family:

$$\begin{aligned}
p(\gamma_{E\cdot}|\sigma_{E\cdot}^2, \Upsilon_{1:T}, \mathcal{D}_{1:N,1:T}) &= p_{\mathcal{N}}\big(\gamma_{E\cdot}|\bar{\gamma}_{E\cdot}, \sigma_{E\cdot}^2 \bar{V}_\gamma^{E\cdot}\big) \qquad (44)\\
p(\sigma_{E\cdot}^2, \Upsilon_{1:T}, \mathcal{D}_{1:N,1:T}) &= p_{IG}\big(\sigma_{E\cdot}^2|\bar{\nu}_{E\cdot}, \bar{s}_{E\cdot}^2\big),
\end{aligned}$$

where the parameters of the posterior distribution are given by the standard formulas:

$$\begin{aligned}
\bar{V}_\gamma^{E\cdot} &= \left((\underline{V}_\gamma^{E\cdot})^{-1} + \sum_{i=1}^N \sum_{t=1}^T \mathbb{I}\{s_{it-1} = \cdot, s_{it} = 1\} Z_t Z_t' \right)^{-1} \quad (45) \\
\bar{\gamma}_{E\cdot} &= \bar{V}_\gamma^{E\cdot} \left((\underline{V}_\gamma^{E\cdot})^{-1} \underline{\gamma}_{E\cdot} + \sum_{i=1}^N \sum_{t=1}^T \mathbb{I}\{s_{it-1} = \cdot, s_{it} = 1\} Z_t x_{it} \right) \\
\bar{s}_{E\cdot}^2 &= \underline{s}_{E\cdot}^2 + \sum_{i=1}^N \sum_{t=1}^T \mathbb{I}\{s_{it-1} = \cdot, s_{it} = 1\} Z_t Z_t' + \underline{\gamma}_{\cdot E}' (\underline{V}_\gamma^{E\cdot})^{-1} \underline{\gamma}_{\cdot E} - \bar{\gamma}_{\cdot E}' (\bar{V}_\gamma^{E\cdot})^{-1} \bar{\gamma}_{\cdot E} \\
\bar{\nu}_{E\cdot} &= \underline{\nu}_{E\cdot} + \sum_{i=1}^N \sum_{t=1}^T \mathbb{I}\{s_{it-1} = \cdot, s_{it} = 1\}.
\end{aligned}$$

Estimation of (γ_i, σ_i^2) Conditional on (ρ, ξ). Given (ρ, ξ) the posterior distribution of γ_i, σ_i^2 can be calculated separately for each unit i. Write

$$\tilde{x}_{it} = x_{it} - \rho x_{it-1} = \gamma_i' Z_t + u_{it}, \quad u_{it} \sim \mathcal{N}(0, \sigma_i^2). \quad (46)$$

This regression is combined with the $\mathcal{N}IG$ prior distribution in (23). The resulting posterior distribution has the same form as the posterior in (44), after making straightforward adjustments to the formulas in (45): there is no summation over i and the indicator function has to be replaced by $\mathbb{I}\{s_{it} = 1, s_{it-1} = 1\}$.

Estimation of (ρ, ξ). Inference for (ρ, ξ) is based on the the marginal likelihood function

$$\begin{aligned}
&\prod_{i=1}^N \int \prod_{t=1}^T \left[p(\gamma_i, \sigma_i^2 | x_{it}, \xi) \right]^{\mathbb{I}\{t = \tau_{i0}\}} \left[p_{\mathcal{N}}(x_{it} | \gamma_i' Z_t + \rho x_{it-1}, \sigma_\iota^2) \right]^{\mathbb{I}\{s_{\iota t} = 1, s_{it-1} = 1\}} d(\gamma_i, \sigma_i^2) \\
&= \prod_{i=1}^N (2\pi)^{-(\bar{\nu}_i - \underline{\nu}_i)/2} \frac{2^{\bar{\nu}_i/2} \Gamma(\bar{\nu}_i/2) |\underline{s}_i^2|^{\underline{\nu}_i/2} |\underline{V}_\gamma^i|^{-1/2}}{2^{\underline{\nu}_i/2} \Gamma(\underline{\nu}_i/2) |\bar{s}_i^2|^{\bar{\nu}_i/2} |\bar{V}_\gamma^i|^{-1/2}}, \quad (47)
\end{aligned}$$

where $(\bar{V}_\gamma^i, \bar{s}_i^2, \bar{\nu}_i)$ are defined similarly as $(\bar{V}_\gamma^{E\cdot}, \bar{s}_{E\cdot}^2, \bar{\nu}_{E\cdot})$ in (45). The marginal likelihood is combined with a prior distribution $p(\rho, \xi)$ and draws from the posterior are generated using a random walk Metropolis Hastings algorithm; see, for instance, Herbst and Schorfheide (2015).

5.3 Summary of csuVAR Estimation

Posterior inference can be implemented sequentially, using the following steps:

1. Use transition counts to estimate the transition probabilities $\hat{\Pi}_{jk,t}$; see (30).

2. Follow the steps outline in Section 3 to estimate the sieve coefficients $\hat{\alpha}_t$ for the approximation $\hat{\ell}_t^{(K)}(x) = \zeta'(x)\hat{\alpha}_t$ in (32).

3. Generate draws from the posterior distribution of the VAR block coefficients in (36); see Section 5.1. Note that $\{\Pi_{j1,t}, \Pi_{j2,t}\}_{j=1}^3$ in the definition of Υ_t, see (20), is replaced by $\{\hat{\Pi}_{j1,t}, \hat{\Pi}_{j2,t}\}_{j=1}^3$.

4. Generate draws from the posterior distributions of $(\gamma_{UE}, \sigma^2_{UE})$ and $(\gamma_{OE}, \sigma^2_{OE})$; see Section 5.2.

5. Generate draws from the posterior distribution of (ρ, ξ); see Section 5.2.

6. For each $(\rho, \xi)^j$ draw, generate a draw from the conditional posterior distribution of (γ_i, σ_i^2), $i = 1, \dots, N$, given $(\rho, \xi)^j$.

Currently we are taking the following short-cuts: Step 4: we run an OLS regression and then later on resample the residuals in our simulations. Step 5: we use a fixed effects regression of x_{it} on x_{it-1}, $x_{i\tau_0}$ and Z_t to construct estimates $\hat{\rho}$, and $\hat{\xi}$. Step 6: conditioning on $(\hat{\rho}, \hat{\xi})$ we compute the posterior means of (γ_i, σ_i^2).

5.4 IRF Computation

The empirical analysis focuses on micro-level impulse responses to aggregate shocks. We will describe the IRF computation conditional on a particular set of parameters $(\theta_{VAR}, \theta_{csu}, \xi, \gamma_{1:N}, \sigma^2_{1:N})$. The computation can either be conducted once for a point estimate, or repeatedly for parameter draws from the posterior distribution. Because of the nonlinearities in the model, the IRFs are, at least in principle, state, sign, and size dependent. We start from an initial value $(\Upsilon^*_{-p+1:0}, \mathcal{D}^*_{1:N,0})$ which in the application we match to a particular time period. We generate a baseline trajectory, denoted by a 0 superscript, and a shocked trajectory, denoted by a *sh* superscript. This simulation can be repeated multiple times. For horizons $h = 1$ to H

1. Aggregate Variables: iterate (34) forward by one period to generate Υ_h^0 and Υ_h^{sh}, respectively. Use the same innovation ϵ_t drawn from a $\mathcal{N}(0, I)$ for both trajectories. The only exception is period $h = 1$, where the innovation for the shocked trajectory is set to the desired value.

2. Unit-level Employment States: based on the transition probabilities (Π_h^0, Π_h^{sh}) generate the states $(s_{1:N,h}^0, s_{1:N,h}^{sh})$. To reduce the simulation noise, we generate a single draw $U_{it} \sim \mathcal{U}[0,1]$ for each unit i and use probability integral transformations based on $(\Pi_{j\cdot,h}^0, \Pi_{j\cdot h}^{sh})$ to determine s_{ih}^0 and s_{ih}^{sh}, respectively.

3. Unit-level Earnings for the Employed: define Z_t^0 and Z_t^{sh} based on $(\Upsilon_t^0, \Upsilon_{t-1}^0)$ and $(\Upsilon_t^{sh}, \Upsilon_{t-1}^{sh})$. Then use (22) to draw x_{it}^0 and x_{it}^{sh}.

4. Cross-Sectional Density: estimate K-dimensional approximations of ℓ_t^0 and ℓ_t^{sh} based on $x_{1:N,t}^0$ and $x_{1:Nt}^{sh}$ following the steps in Section 3. The log densities can be converted into densities as needed.

6 Generalizations

(In this section we will discuss the relaxation of parts of Assumption 1 and how to make the parametric features of the micro-level component of the csuVAR more flexible.)

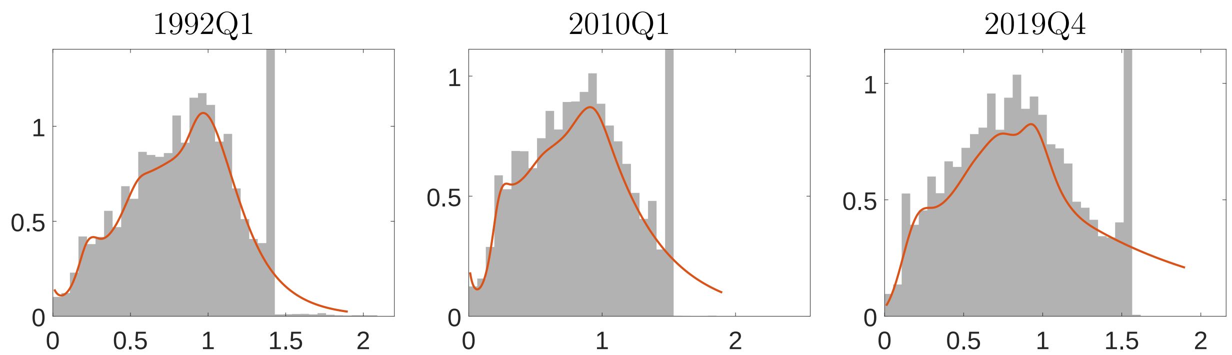
7 Empirical Analysis

In the empirical application we estimate the fVAR and the csuVAR based on a subsample of an administrative data set that contains earning information for individuals registered in the social security system. One of the questions that we will investigate is whether the estimation of the fVAR delivers a similar estimate of the effect of a productivity shock on the cross-sectional distribution of earnings as the fVAR analysis. A brief summary of the aggregate and micro-level data and the shock identification is provided in Section 7.1. The results from the fVAR and csuVAR analysis are presented in Sections 7.2 and 7.3, respectively. Finally, we comment on the relationship between our empirical strategies and the widely-used local projections in Section 7.4.

7.1 Data and Shock Identification

The micro-level observations used in our analysis are obtained from the Institute für Arbeitsmarkt- und Berufsforschung (IAB), Germany. The dataset covers approximately 80% of the total German labor force, where self-employed and civil servants are not included

Figure 1: Estimated Densities for Three Time Periods



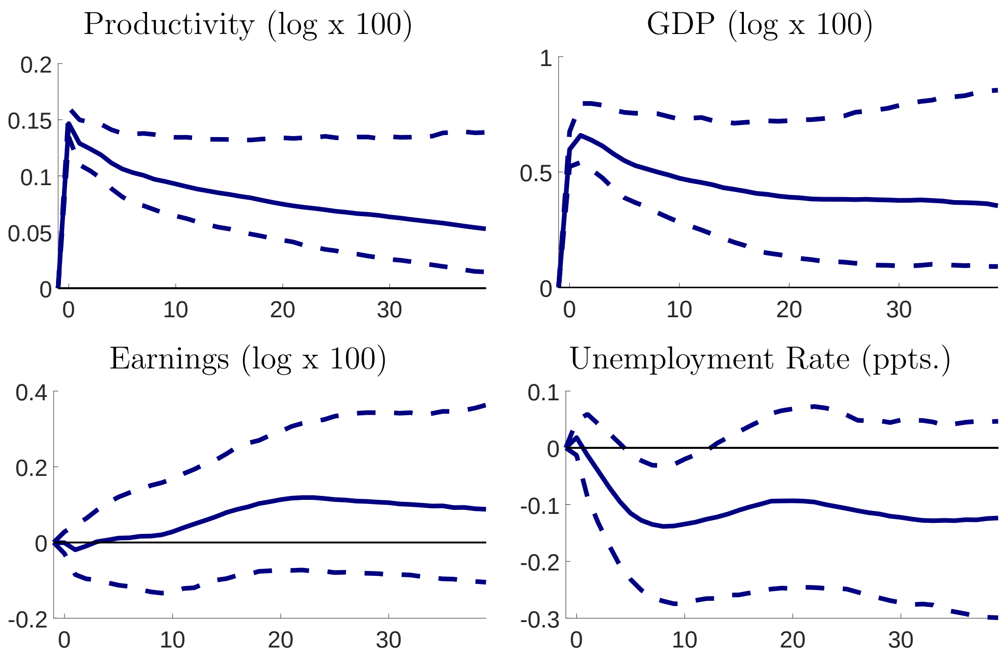
Notes: The solid lines represent the ML estimates $\hat{k}_t^x(x)$ based on the K-dimensional sieve approximation of the log density. We overlay histograms of the raw data.

in the sample. The data is in the form of labor market spells, which report the average daily wage during the spell together with the number of working days. We use a random 2% sample of the population (all individuals who have ever been registered with the social insurance system). From this data set we construct the observations (x_{it}, s_{it}) as discussed in Section 4.1.

The vector Y_t includes real GDP per capita and labor productivity. Real GDP per capita is obtained from the Federal Statistical Office, Germany. Labor productivity is measured as total real GDP over total hours worked, which also is obtained from the Federal Statistical Office. The average level of (transformed) earnings, the unemployment rate, and the estimated transition probabilities $\hat{\Pi}_{jk,t}$ are constructed from the IAB micro data; see (16), (18), and (30). The fVAR and the csuVAR are estimated on data from 1992:Q1 to 2019:Q4. More detailed information on the data set is provided in the Online Appendix.

We order labor productivity first in the vector Y_t (and hence Υ_t) and use a recursive shock identification scheme. Formally, we set Ω in (35) equal to the identity matrix. We focus on the first shock which we interpret as a general productivity shock. Our methodology can be combined with any other identification scheme. In particular, one can examine the effect of monetary or fiscal shocks. We chose technology shocks because they generally explain a larger fraction of the business cycle fluctuations than policy shocks.

Figure 2: fVAR IRFs of Aggregate Variables



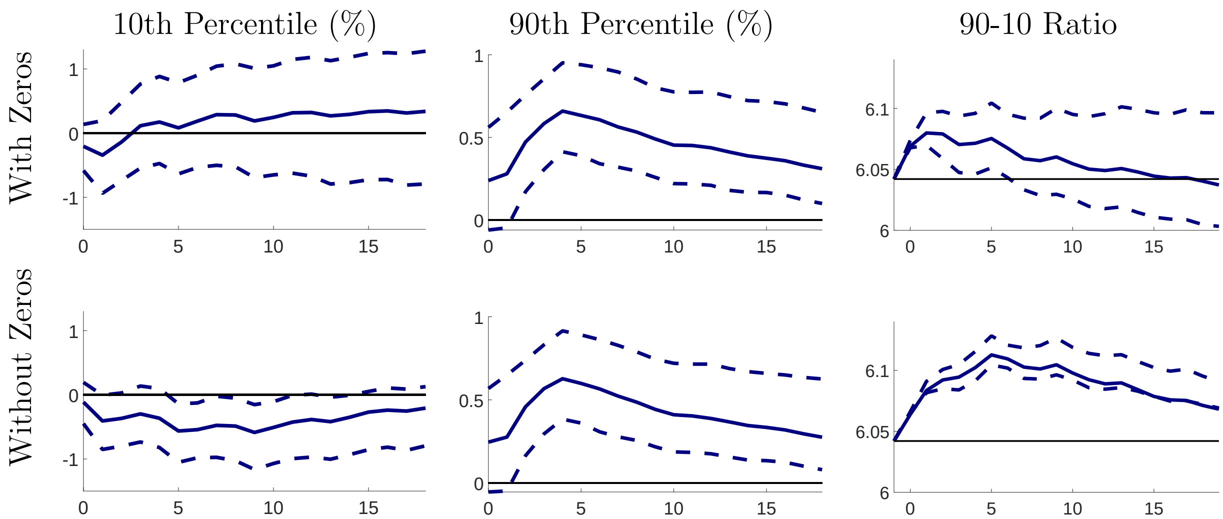
Notes: The solid lines represent posterior mean (pointwise) responses and the dashed lines delimit 90% credible intervals. x-axis corresponds to time horizon.

7.2 Results from the fVAR

Estimated Cross-sectional Densities. In Figure 1 we depict the estimated densities for 1992:Q1, 2010:Q1, and 2019:Q4, and overlay histograms. Two observations stand out: first, the density estimator smooths out some of the bumps in the histograms. An increase in K could capture some of these bumps, but would also increase the dimensionality of the VAR and in particular the number of regressors $\hat{\alpha}_{t-j}$. The model selection criterion used to determine (K, λ, p) jointly trades of fit against model complexity. Second, there is a pronounced spike in the right tail of the histograms which is caused by top coding. Our estimator uses the tails of the density $p^{(K)}(x_{1:N,t}|\alpha_t)$ to distribute the earnings at the top coding level.

IRFs of Aggregate Variables. IRFs of the aggregate variables obtained from the fVAR are depicted in Figure 2. Upon impact of the shock, labor productivity rises by 15 basis points and then slowly decays to about 7 basis points after 10 years. Meanwhile per capita GDP rises by 50 basis points in the short run and about 30 basis points in the long run. Average cross-sectional earnings in our sample rise by about 10 basis points, but with a delay of 10 quarters. Finally, the unemployment rate falls by 10 basis points. As a basis of

Figure 3: IRF of Percentiles and Inequality Statistics



Notes: The solid lines represent posterior mean (pointwise) responses and the dashed lines delimit 90% credible intervals. x-axis corresponds to time horizon.

comparison (not shown in the figures) we also estimated a VAR based on Υ_t only, without the estimated sieve coefficients, including either the unemployment rate or the transition probabilities $\hat{\Pi}_{jk,t}$ in the vector of Υ_t. The resulting IRFs are very similar to the ones depicted in Figure 2.

IRFs of Cross-sectional Distribution: Percentiles and Inequality Statistics. The computation of the cross-sectional density IRFs follows the description in Section 3.4. From the sequence of baseline densities $p_h^0(\cdot)$ and $p_h^{sh}(\cdot)$, $h = 1, \ldots, H$, we can compute, for instance, percentiles and inequality measures. These can be computed exclusively based on the continuous densities normalized to one, or by including a pointmass at zero for the unemployed, normalizing the continuous part to one minus the unemployment rate. The results are depicted in Figure 3. We applied a change of variables to undo the inverse hyperbolic sine transformation.

The panels in the top row of the figure show responses of percentiles that are computed with a pointmass at zero equal to the unemployment rate. Because the productivity shock reduces unemployment, more individuals have positive earnings. At the posterior median there is a slight increase at the 10th percentile. The earnings increase is, however, much more pronounced at the 90th percentile. Thus, in relative terms, individuals in "high" paying jobs benefit more and the 90-10 ratio rises. This result for the German data differs from the result obtained for U.S. data in CCS. In the U.S. earnings inequality in the Current Population

Survey (CPS) data is countercyclical, with individuals at the bottom end of the earnings distribution benefiting more from a technology shock.

If the pointmass at zero is excluded then we are capturing the earnings distribution conditional on being employed. The drop at the 10th percentile (relative to average earnings) could be the result of a rigid labor market for low-skilled individuals and slowly adjusting wages. Conditional on working, the 90-10 ratio response more strongly, featuring a more pronounced increase in inequality.

7.3 Results from the csuVAR

csu Parameter Estimates. The posterior sampler generates draws from the joint posterior distribution of $(\gamma_{1:N}, \sigma^2_{1:N}, \theta_{csu}, \xi)|(\Upsilon_{1:T}, \mathcal{D}_{1:N,1:T})$.[7] The presentation of the estimation results proceeds in three steps. First, we discuss the posterior distribution of the homogeneous parameters θ_{csu}. Second, we consider the posterior estimates of ξ, which characterizes the CRE distribution; see (23). Let $\hat{\xi}$ be the posterior mean estimate of ξ. Much of the discussion will focus on $p(\gamma_i, \sigma_i^2|x_{it}, \hat{\xi})$ which can be interpreted as an implicit estimate of the CRE distribution. Third, we examine the posterior estimates $\hat{\gamma}_i$ and $\hat{\sigma}_i^2$ of the heterogeneous parameters.

Priors and posteriors for the homogeneous parameters θ_{csu} are provided in Table 1. The autoregressive parameter ρ for the earnings dynamics equation is estimated to be 0.75. The remaining estimates characterize the distribution of earnings conditional on having been unemployed (UE transitions) or not in the sample (OE transitions), respectively. The distribution of earnings under the UE transitions seems to be largely unaffected by aggregate productivity and GDP growth, with posterior mean estimates of 0.01 and -0.001, respectively. The coefficient estimates for the OE transitions are larger, most notably, the coefficient on productivity growth is 0.15.

To put these numbers into perspective it is important to understand the scale of x_{it}. Recall that x_{it} is the hyperbolic inverse sine transformation of the ratio of individual earnings to average earnings. Suppose that the earnings of unit i equal the cross-sectional average ($x_{it} = 0.88$) and its γ_i coefficient for productivity growth (in %) is equal to 0.1. Then a one-percent increase in productivity growth would raise its earnings ratio from 1.0 to 1.14.

[7] The estimation results reported in this section are preliminary and based on a short-cut to the full Bayesian estimation described in Section 4.

Table 1: Homogeneous Parameters

		Prior		Posterior	
Para.	Regressor	Mean	90% Intv.	Mean	90% Intv.
ρ	x_{it-1}			0.75	
γ_{UE}	const.			0.40	
	Prod. Growth [%]			0.01	
	GDP Growth [%]			-.001	
σ^2_{UE}				0.27^2	
γ_{OE}	const.			0.35	
	Prod. Growth [%]			0.15	
	GDP Growth [%]			-0.03	
σ^2_{OE}				0.31^2	

The estimated standard deviations σ_{UE} and σ_{OE} are large: 0.27 and 0.31, respectively. Thus, there is a large amount of unexplained cross-sectional variation in the earnings of those who enter unemployment. In fact, the R^2s in pooled (across i and t) regressions for the UE and OE samples, respectively, are very low.

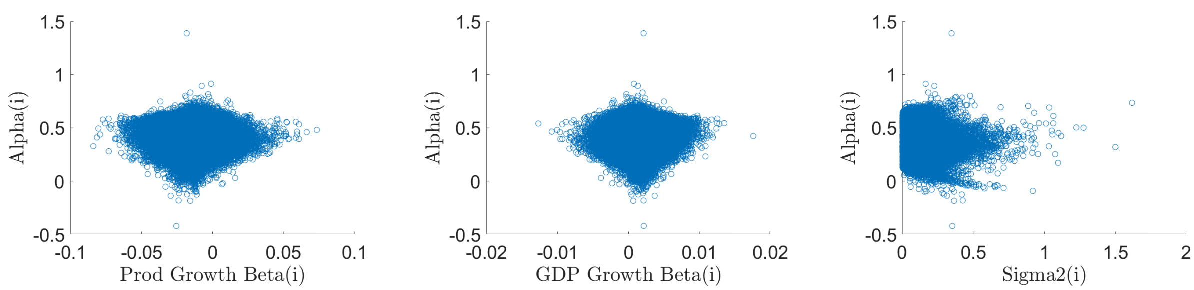
The estimation results for the hyperparameters ξ are summarized in Table 2. Columns 3 and 4 contain the prior mean and a 90% credible interval for ξ, and Columns 5 and 6 the posterior point and interval estimates. To interpret the estimates, we focus on how they shape the CRE distribution through $p(\gamma_i, \sigma^2_i | x_{it}, \hat{\xi})$. The positive estimate of $\underline{\gamma}_s$ implies that units with high initial earnings tend to also have higher earnings in the long-run. The implicit prior for the effect of productivity growth on unit-level earnings is negative (with a mean of -0.016) and it is slightly positive (with a mean of .004) for GDP growth. To capture the effect of aggregate shocks on the distribution of earnings, it is important to allow for heterogeneity in the γ_i coefficients. This heterogeneity is controlled by the variance of the CRE distribution $\underline{V}_\gamma$. Focusing on the coefficient on productivity growth, setting $\lambda = 0.0025$, and converting variances into standard deviations, the estimated CRE standard deviation of the productivity growth coefficient is 0.049.

At last, we show scatter plots of posterior mean estimates of $\hat{\gamma}_i$ and $\hat{\sigma}^2_i$ in Figure 4. Each point in the scatter corresponds to a particular unit i. The y axis in each of the graphs is the

Table 2: Hyperparameters

		Prior		Posterior	
Para.	Regressor	Mean	90% Intv.	Mean	90% Intv.
$\underline{\gamma}_0$	const.			0.08	
	Prod. Growth [%]			-.016	
	GDP Growth [%]			0.04	
$\underline{\gamma}_s$	const.			0.11	
diag($\underline{V}_\gamma$)	const.			0.21	
	Prod. Growth [%]			$6 \cdot 10^{-6}/\lambda$	
	GDP Growth [%]			$2 \cdot 10^{-7}/\lambda$	
$\underline{\nu}$				5	
$\underline{s}^2$				$5 \cdot 0.18^2$	

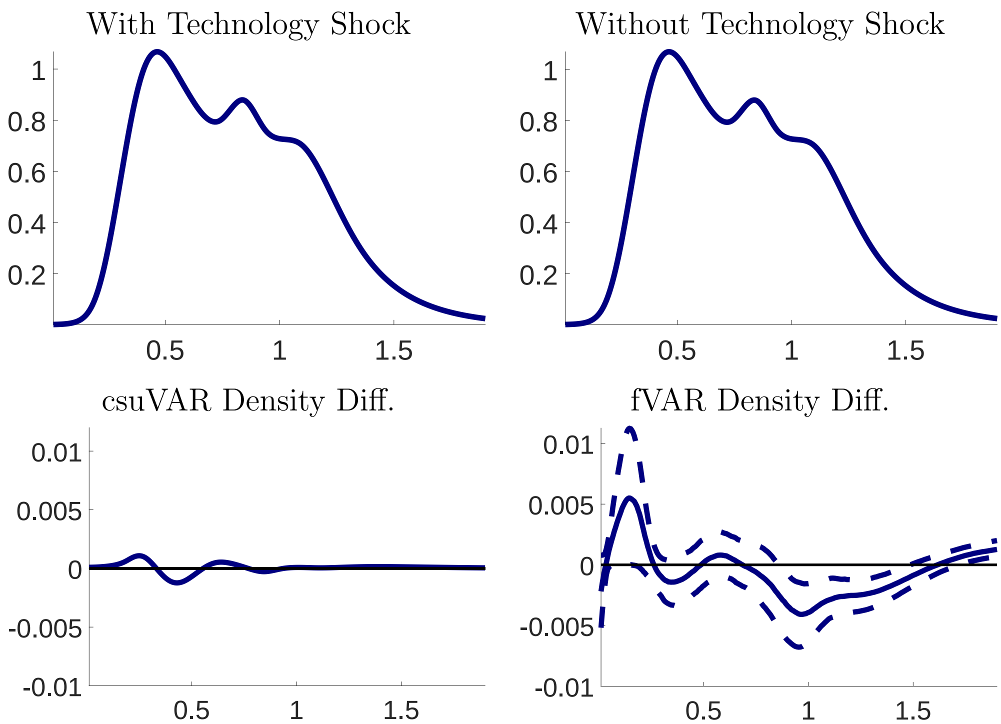
Figure 4: Scatterplots of Posterior Mean Estimates $(\hat{\gamma}_i, \hat{\sigma}_i^2)$



posterior mean of the intercept. The x-axis represents the posterior mean of the productivity growth coefficient, the GDP growth coefficient, and the variance estimate $\hat{\sigma}_i^2$. Overall, there is a substantial amount of dispersion in the unit-specific coefficient estimates. The posterior mean estimates of the intercept range from approximately -0.1 to 1.0 (abstracting from some outliers), the estimates of the unit-specific productivity growth coefficient from -0.06 to 0.05, and productivity growth from -0.01 to 0.01.

IRFs of Cross-sectional Density. We now turn to the implied response of the cross-sectional density of earnings. The computation of the IRF is described in Section 5.4 and results for the response in period $h = 4$ are plotted in Figure 5. The top left panel shows

Figure 5: Impulse Response of Cross-sectional Densities, $h = 4$



the cross-sectional density with a technology shock in period $h = 1$, whereas the top left panel shows the baseline cross-sectional distribution in period 2010Q1. The two densities look virtually indistinguishable, but the bottom left panel shows the differential obtained from the csuVAR. For comparison, we plot in the bottom right panel the density differential obtained from the fVAR analysis. The two density differentials are qualitatively similar, but quantitatively different. The distributional effect obtained from the csuVAR analysis is more muted.

7.4 Connection to Panel Local Projections

Many empirical studies, in particular those based on administrative data sets, use a form of panel local projections to study the effect of aggregate shocks on micro-level outcomes. Local projections (LPs) were initially proposed as an alternative to structural VARs to study the effects of shocks on aggregate outcomes; see Jorda (2005). The basic idea is that the vector time series has an infinite-order linear process (Wold) representation and the coefficient matrices can be directly estimated by multi-step regressions. In the past decade structural

VAR identification shifted to the use of observable shocks or shock instruments. In a local projection framework this amounts to projecting time $t+h$ observables onto a time $t+1$ shock measure and some additional control variables. Suppose that the productivity shock ϵ_{1t} is observed. Then a typical panel LP would take the form

$$x_{it+h} = \alpha_i + \beta_i \epsilon_{1t+1} + \psi x_{it} + \delta_i' Y_t + \text{(additional controls)} + \omega_{it+h} \tag{48}$$

and β_i would be interpreted as the effect of the aggregate shock ϵ_{1t} on unit-level earnings in period $t+h$. The coefficient ψ captures the effect of current unit-level earnings on future earnings. In the logic of the csuVAR, the additional control variables should also include $\ell_t(\cdot)$, because the cross-sectional log density affects future Υ_ts and future Υ_ts, in turn, affect unit-level earnings.

The effect of the aggregate shock ϵ_{1t} on the cross-sectional distribution of earnings depends on the distribution of β_i. If the coefficients were homogeneous, then the distribution of x_{it} would simply shift. However, in general we expected the β_is to be heterogeneous and their values could be correlated with x_{it}. A popular assumption in the literature is that of observed group heterogeneity, i.e.,

$$\beta_i = b_{g(i)}, \quad g(i) \in \mathcal{G}, \quad \mathcal{G} = \{1, 2, 3, \ldots, n_g\}. \tag{49}$$

This implies that the function $g(i)$ is known and the coefficients b_g can be estimated by pooling observations from units that belong to the same group. While the coefficient estimates $\hat{b}_g$ measure how members of group g respond to aggregate shocks, it requires additional work to convert the estimates into a response of the cross-sectional distribution. This requires knowledge about the dependence between $b_g(i)$ and x_{it} and the response can be obtained by evaluating (48) for each i under a baseline scenario $\epsilon_{1t} = 0$ and a shocked scenario $\epsilon_{1t} = 1$. While the panel local projection calculation appear to be easier to implement than the estimation of the csuVAR, there is a real danger that the assumed group structure misses a lot of the cross-sectional heterogeneity.

8 Conclusion

We started out from distinct but related questions: (i) what is the effect of an aggregate shock on the cross-sectional distribution of x_{it}? (ii) How does a particular cross-sectional unit (household, firm, etc.) or group of units respond to an aggregate shock? Question (i) can be answered with repeated cross-sections whereas an answer to Question (ii) typically requires panel data. The two questions are connected, because if Question (ii) is answered for all households, then Question (i) is answered as well. We use a random sample of a German administrative earnings data to estimate an fVAR that tracks cross-sectional distributions and a csuVAR that tracks cross-sectional units. Both models include a vector autoregressive law of motion for a vector of macroeconomic aggregates. While the fVAR is based on previous research, in particular CCS, the csuVAR is new. It expands existing panel data models by letting aggregate variables affect micro-level outcomes and the cross-sectional distribution feed back into the aggregate block of the model.

Both modeling approaches have advantages and disadvantages. The fVAR approach has a weaker data requirement because repeated cross sections suffice for the estimation. Moreover, unit-level behavior and heterogeneity does not need to be explicitly modeled. At the unit-level the time-series x_{it} can be highly nonlinear, e.g., significant earnings changes might be restricted to promotions, job-to-job transitions, or transitions in and out of unemployment. The cross-sectional distribution, on the other hand, is not as sensitive to these nonlinearities, because it stays unchanged if two units trade places in the x_{it} distribution. The main disadvantage of the fVAR modeling approach is that it is not designed to track specific units and hence cannot answer Question (ii) above.

The key advantage of the csuVAR is its ability to track unit-level behavior. However, the costs are substantial: one needs panel data to estimate the model and it is challenging to specify unit-level laws of motion, which may require nonlinear and non-Gaussian features. Moreover, coefficient heterogeneity needs to be carefully modeled. We contrasted both approaches in an application in which we examined the effect of productivity shocks on the cross-sectional distribution of earnings.

In practice, researchers are limited by the availability of data sets. Insights from this paper may also be useful for combining different types of data sets and conducting analyses with mixed-frequency data.

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.

BROWNING, M., AND M. EJRNJES (2013): "Heterogeneity in the Dynamics of Labor Earning," *Annual Review of Economics*, 5, 219–245.

BROWNING, M., M. EJRNJES, AND J. ALVAREZ (2010): "Modelling Income Processes with Lots of Heterogeneity," *Review of Economic Studies*, 77(4), 1353–1381.

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, M., AND F. SCHORFHEIDE (2022): "On the Effects of Monetary Policy Shocks on Earnings and Consumption Heterogeneity," *CEPR Working Paper*, 17049.

ETTMEIER, S. (2023): "No Taxation without Reallocation: The Distributional Effects of Tax Changes," *CRC Discussion Paper*, 06/2023.

GUVENEN, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, 97(3), 687–712.

——— (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1), 58–79.

HAHN, J., G. KUERSTEINER, AND M. MAZZOCCO (2020): "Estimation with Aggregate Shocks," *Review of Economic Studies*, 87, 1365–1398.

HERBST, E., AND F. SCHORFHEIDE (2015): *Bayesian Estimation of DSGE Models*. Princeton University Press, Princeton.

Hoffmann, F. (2019): "HIP, RIP, and the Robustness of Empirical Earnings Processes," *Quantitative Economics*, 10, 1279–1315.

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.

Hryshko, D. (2012): "Labor Income Profiles are Not Heterogeneous: Evidence from Income Growth Rates," *Quantitative Economics*, 3, 177–209.

Jorda, O. (2005): "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 95(1), 161–182.

Liu, L. (2023): "Density Forecasts in Panel Data Models: A Semiparametric Bayesian Perspective," *Journal of Business & Economic Statistics*, forthcoming.

Liu, L., H. R. Moon, and F. Schorfheide (2023): "Forecasting With a Panel Tobit Model," *Quantitative Economics*, 14(1), 117–159.

Stone, C. J. (1990): "Large-Sample Inference For Log-Spline Models," *Annals of Statistics*, 18(2), 717–741.

Stüber, H., W. Dauth, and J. Eppelsheimer (2023): "A guide to preparing the sample of integrated labour market biographies (SIAB, version 7519 v1) for scientific analysis," *Journal for Labour Market Research*, 57(1), 7.