Firm Heterogeneity and Aggregate Fluctuations: a Functional VAR model for multidimensional distributions*

Massimiliano Marcellino † Andrea Renzetti ‡ Tommaso Tornese §

This Draft: August 25, 2025

Abstract

We develop a Functional Augmented Vector Autoregression (FunVAR) model to explicitly incorporate firm-level heterogeneity observed in more than one dimension and study its interaction with aggregate macroeconomic fluctuations. Our methodology employs dimensionality reduction techniques for tensor data objects to approximate the joint distribution of firm-level characteristics. More broadly, our framework can be used for assessing predictions from structural models that account for micro-level heterogeneity observed on multiple dimensions. Leveraging firm-level data from the Compustat database, we use the FunVAR model to analyze the propagation of total factor productivity (TFP) shocks and monetary policy shocks on the US macroeconomy, examining their impact on both macroeconomic aggregates and the cross-sectional distribution of capital and labor across firms. Then, we use the proposed framework to identify cross-sectional uncertainty shocks and evaluate their effects on aggregate macroeconomic fluctuations.

J.E.L Classification Code: C32; E32

Keywords: Functional VARs; Tensor data; Firm heterogeneity; Cross-sectional uncertainty; Risk shocks

*Andrea Renzetti and Massimiliano Marcellino thank MUR-Prin 2022 - Prot. 20227YZ9JK, financed by the European Union - Next Generation EU, for partial financial support. Seminar participants at Bocconi University and Cattolica University of Milan provided useful comments on a previous draft. We thank Joshua C. Chan, Stephanie Ettmeier for helpful comments.

†Bocconi University and BAFFI Centre; email: massimiliano.marcellino@unibocconi.it

‡Bocconi University and BAFFI Centre; email: andrea.renzetti@unibocconi.it.

§Universitá Cattolica del Sacro Cuore, Milano; email: tommaso.tornese@unicatt.it

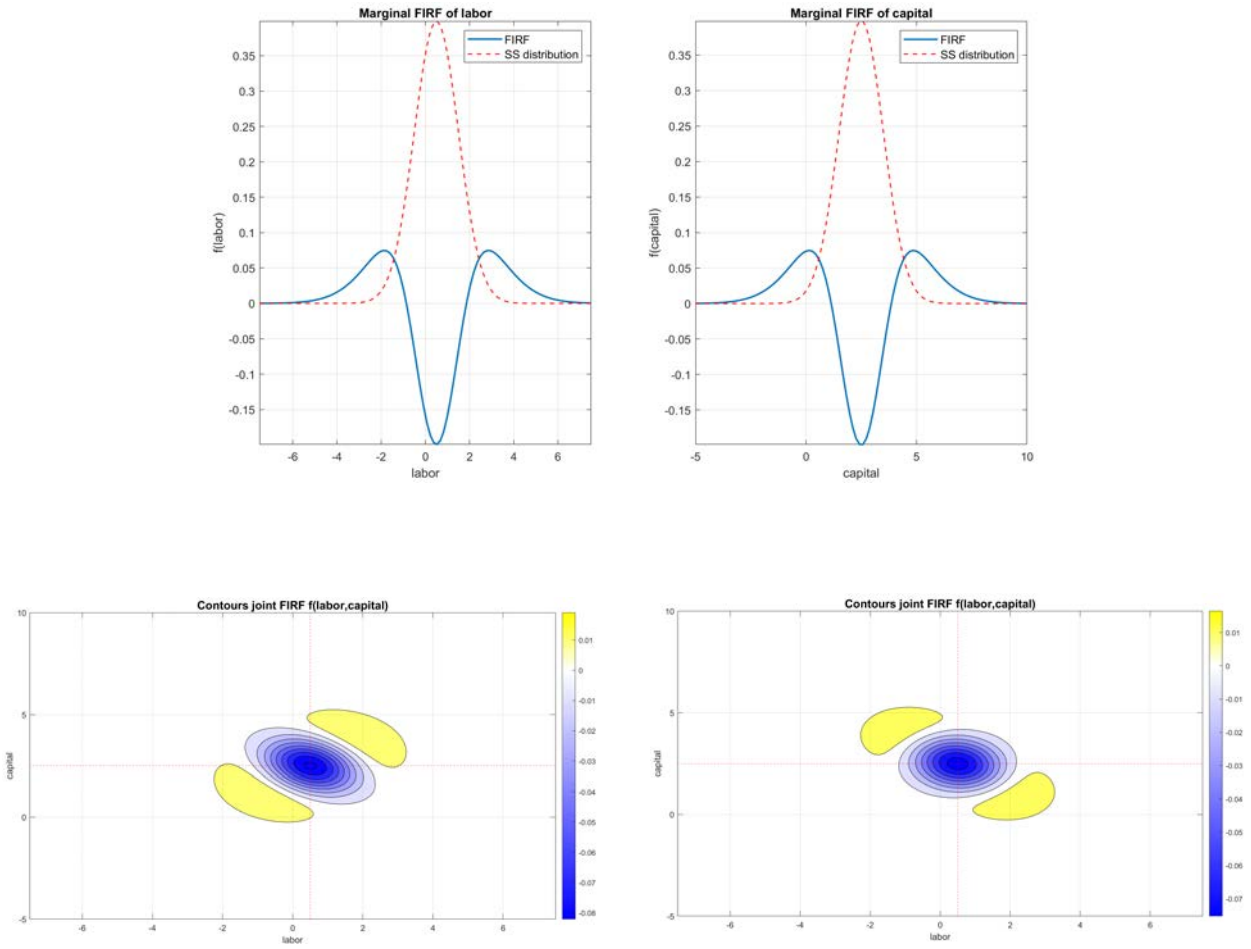
1 Introduction

The study of household and firm heterogeneity plays a crucial role in understanding macroeconomic fluctuations. Extensive research has recently focused on household heterogeneity and its implications for aggregate fluctuations within both fully structural and semi-structural frameworks (Aiyagari 1994; Kaplan et al. 2018; Bayer et al. 2019; Bilbiie et al. 2023; Chang et al. 2022). Analogously, heterogeneous firm models often suggest that firm heterogeneity offers additional nuanced insights for understanding macroeconomic fluctuations (Winberry 2018; Koby et al. 2020; Winberry 2021; Ottonello et al. 2020a). In this paper, we propose a semi-structural model based on a Functional Augmented Vector Autoregression (FunVAR) framework to model the dynamic interaction between the multidimensional distributions of firm-level characteristics and the macroeconomic aggregates. Our approach builds on recent advancements in the estimation of functional VARs, where macroeconomic aggregates evolve jointly with the distribution functions of micro-level variables (Chang et al. 2021; Huber et al. 2024; Bjørnland et al. 2023; Chang et al. 2025). Distinct from prior contributions, our methodology explicitly tackles the complexity of modeling a multidimensional distribution of the micro-level observables. We do so by leveraging dimensionality reduction techniques for tensor data objects to non-parametrically approximate the joint distribution of firm-level characteristics. These techniques make it easier to handle complex, high-dimensional dependencies compared to splines, Bernstein polynomials, and other basis function approximations. They also offer better scalability in multidimensional environments, where such basis functions can become computationally intensive and impractical.

We emphasize the necessity of modeling joint distributions of micro-level observables, not only for validating macroeconomic theories and testing predictions derived from heterogeneous agent models, but also for more broadly identifying and evaluating macroeconomic shocks and their distributional effects. First, modeling the joint distribution of micro-level characteristics, rather than marginal distributions, is crucial for qualitative assessing the distributional effects of macroeconomic shocks on the micro-level distributions. Imagine, for example, having two structural heterogeneous firm models that differ fundamentally in terms of how firm's reallocate labor and capital after the realization of a shock. In one structural model, after the aggregate shock, some firms with intermediate labor and capital levels expand both labor and capital while other similar firms reduce both inputs. In the other structural model, some firms with intermediate labor and capital increase capital while reducing labor, while others do the opposite. Importantly, the same changes in the

marginal distributions are compatible with both structural models. Figure 1, upper panel, shows (in blue) the changes in mass of the marginal distribution functions after the realization of the shock, relative to the steady-state value, alongside the steady-state distribution itself (in red).

Figure 1: Marginal and joint functional IRF of labor and capital



Notes: The figure shows the functional IRFs of the marginal labor and capital distributions to a shock (upper panel). The lower panel reports the contours from two different bivariate functional IRFs, both equally compatible with the changes in the marginal distributions in the upper panel. The red dashed lines in the lower panel represent the steady state mean values.

After the shock, the mass of firms both with low and high capital endowment increases, as does the mass of firms with low and high laborendowment. However, the marginals alone do not indicate how capital and labor are adjusting relative to each other. In the lower panel, the figure shows the two contour plots of the change in the mass of the bivariate density of labor and capital with respect to the bivariate steady-state distribution implied by the two structural models. In the scenario on the left, some firms with intermediate levels of capital and labor expand both inputs while others

reduce both. In the scenario on the right, some firms specialize by increasing capital while reducing labor, while others specialize by increasing labor and reducing capital. This simple example shows that, while challenging, modeling the joint distribution rather than just the marginal distributions of the observed micro variables is essential for qualitatively evaluating the effects of macroeconomic shocks on the micro-level distributions and understanding whether the data favor a certain scenario, supported for example by a specific structural model, over another. In the empirical analysis, we use the FunVAR to evaluate the effects and propagation of aggregate TFP shocks and monetary policy shocks both on macroeconomic aggregates and on the joint distribution of firms-level labor and capital .

Moreover, modeling the multidimensional joint distribution of micro-level observables can also be instrumental for identifying macroeconomic shocks—especially when these shocks are naturally defined as exogenous shifts in specific features of this joint distribution. A prominent example are cross-sectional uncertainty shocks, which can be interpreted as an exogenous shift in the cross-sectional dispersion of firm-level productivity (Bloom 2009; Christiano et al. 2014; Arellano et al. 2019; Dew-Becker et al. 2023). In observable terms, these shocks manifest as exogenous changes in the cross-sectional dispersion of firms' output, conditional on their capital and labor inputs. Such shocks, sometimes referred to as risk shocks, have been shown to play a significant role in driving business cycle fluctuations. To identify and quantify the macroeconomic effects of these cross-sectional uncertainty shocks, we therefore adopt a Functional VAR framework that links the joint distribution of firms' output, capital, and labor with the macroeconomic aggregates. Consistently with the theoretical literature, we exploit this framework to identify cross-sectional uncertainty shocks as distributional shocks which maximize the variation in the dispersion of firms' output conditional on firms' capital and labor.

Our econometric approach leverages dimensionality reduction techniques for tensor data objects to non-parametrically approximate the multidimensional distribution of firm-level characteristics. Specifically, in the paper we discuss three alternative approaches. First, we consider applying principal component analysis (PCA) to the unfolded tensor—a process known as flattening or unfolding in the context of tensor analysis (Kolda et al. 2009). Next, we examine multilinear principal component analysis (Tucker 1966) and CP - canonical decomposition/parallel factor decomposition - (Carroll et al. 1970; Harshman 1970). Unlike PCA on unfolded data, both multilinear PCA and CP decomposition preserve the tensor's multi-dimensional structure, yielding more parsimonious representations, though being more sensitive to misspecification. All these methods have been

widely applied in feature extraction and pattern recognition tasks, particularly in contexts such as 2-D/3-D images and video sequences, where data are naturally represented as tensors. We show that they can also be successfully used to non-parametrically approximate an unconstrained transformation of the multidimensional distribution function—specifically, the joint distribution of firm-level characteristics in our application.

In our empirical analysis, we focus on firm heterogeneity, an aspect that is typically addressed within fully structural models in the literature. These models are often calibrated or estimated using only macroeconomic data.[1] In contrast, semi-structural models—commonly employed to assess the impact of macroeconomic shocks on aggregate fluctuations—tend to neglect firm-level heterogeneity (see Ramey (2016) for a general review).[2] We employ the Compustat database, exploiting both the Quarterly and the Annual datasets to recover the cross sectional distribution of firm-level production, capital and labor for the US companies from 1984-Q4 to 2019-Q4.

In the first part of our empirical analysis, we use an internal instrument identification strategy to evaluate the effects of aggregate TFP shocks and monetary policy shocks on both macroeconomic aggregates and the joint distribution of firm-level labor and capital, through the lens of the FunVAR model. We find that TFP shock have a persistent effect on the joint labor-capital distribution, with the effects peaking between one and two years after the shock. TFP shocks lead to a simultaneous increase of both capital and labor above their steady state level and an increase in the dispersion of the two inputs across the pool of firms. The effects of monetary policy shocks are found to be less persistent and vanish within one year. A contractionary monetary-policy shock reduces, on average, both labor and capital inputs across firms, but weakens their co-movement: the cross-sectional correlation between labor and capital falls in the short run and gradually recovers as the two factors realign.

In the second part of our empirical contribution, we use the funVAR model for a "micro-to-macro" application. We identify shocks to the dispersion of cross-sectional firm's productivity and evaluate their effects on aggregate macroeconomic fluctuations. We find that cross sectional uncertainty shocks are associated with declines in investment, output, employment and consumption, consistent with theoretical predictions.

The paper is structured as follows. Section 2 introduces the FunVAR model. Section 3 discusses

1. An exception is Liu et al. (2023), who develop a method for estimating heterogeneous agent models using both micro and macro data, though they apply it solely to simulated firm-level data.

2. An exception is Lenza et al. (2024), that recently study the role of heterogeneity in the revenues of individual firms for the transmission of a business cycle shock in the euro area using the framework of Chang et al. (2021).

estimation and inference. Section 4 tests the FunVAR using simulated data from an heterogeneous agents model. Sections 5 and 6 study the distributional effects of, respectively, TFP and monetary shocks in the US. Section 7 studies the effects of cross-sectional uncertainty shocks. Section 8 summarizes and concludes.

2 Model

We assume that we observe a vector of macroeconomic variables $\boldsymbol{y}_t$ and repeated cross sections of firms level characteristics $\boldsymbol{x}_{it}$ for $t = 1, \dots, T$ periods and $i = 1, \dots, N_t^{cross}$ firms. We also assume that the cross-sectional observations are drawn from a time-varying multivariate distribution with density $f_t(\boldsymbol{x})$. To model the dynamic interaction between the distribution function of firms-level characteristics and the aggregate macroeconomic time series, we consider the following function augmented VAR model

$$\boldsymbol{y_t} = \boldsymbol{c_y} + \sum_{s=1}^{p} \boldsymbol{B}_{l,yy}\boldsymbol{y_{t-s}} + \sum_{s=1}^{p} \int \boldsymbol{B_{s,yl}}(\boldsymbol{x}) l_{t-s}(\boldsymbol{x}) d\boldsymbol{x} + \boldsymbol{u_{y,t}} \,, \tag{1}$$

$$l_t(\boldsymbol{x}) = c_l(\boldsymbol{x}) + \sum_{s=1}^{p} \boldsymbol{B}_{s,ly}(\boldsymbol{x})\boldsymbol{y}_{t-s} + \sum_{s=1}^{P} \int B_{ll}(\boldsymbol{x}, \boldsymbol{x}') l_{t-s}(\boldsymbol{x}') d\boldsymbol{x}' + u_{l,t}(\boldsymbol{x}) \,. \tag{2}$$

The macroeconomic aggregates are stored in the $n_y \times 1$ column vector $\boldsymbol{y_t}$ while $l_t(\boldsymbol{x})$ is defined to be the Centered-Log Ratio (CLR) transform of the multivariate density function of the vector of firms-level characteristics. For example, assuming that in each period we observe firms' specific labor and capital endowments, we have $\boldsymbol{x} = [x_1, x_2]'$ where x_1 stands for the labor input while x_2 for the capital input. The function $l_t(\boldsymbol{x})^{obs}$ contains the information about the distribution of the firm-level capital and labor at each time t. In particular, the CLR transformation of the distribution function is given by

$$l_t(\boldsymbol{x}) := \text{CLR}(f_t(x_1, x_2)) = \log(f_t(x_1, x_2)) - \frac{1}{|\Omega|} \int_{\Omega} \log(f_t(x_1, x_2))\, dx_1 dx_2 \,. \tag{3}$$

This transformation greatly simplifies the econometric analysis of the time variation of the multivariate distribution function as it maps a density function, that needs both to integrate to one and to obey non-negativity constraints, to an unconstrained real-valued space.[3] We assume that the

3. The CLR transformation, traditionally used in the context of compositional data, has been used for modelling distribution functions by Hron et al. (2016). For an extensive discussion on the use of the CLR transformation in the context of distribution functions we refer to Petersen et al. (2022).

$l_t(\boldsymbol{x})$ that we can observe, or estimate on a grid, is a noisy realization of the CLR transformation of the true multivariate density function, namely

$$l_t(\boldsymbol{x})^{\text{obs}} = l_t(\boldsymbol{x}) + \varepsilon_t, \tag{4}$$

where ε_t is a noise with $\mathbb{E}[\varepsilon_t] = 0$ and $\mathbb{E}[\varepsilon_t^2] = \sigma^2$ and $l_t(\boldsymbol{x})$ is the true CLR transformation of the distribution function evaluated at the grid points. For example, in the bivariate labor and capital example we can denote $\mathcal{X}_1 = \{x_{1,1}, x_{1,2}, \ldots, x_{1,N_1}\}$ and $\mathcal{X}_2 = \{x_{2,1}, x_{2,2}, \ldots, x_{2,N_2}\}$ the sets of grid points for labor and capital, respectively, and define $\boldsymbol{x} \in \mathcal{X}_1 \times \mathcal{X}_2$, the Cartesian product of all pairs (x_1, x_2). We can then estimate $f_t(x_1, x_2)$ from available data obtaining noisy observations of the true CLR transformation of the distribution function on $N^{grid} = N_1 N_2$ grid points, where the noise is due to the density estimation error. We assume that the true centered log-ratio transformation of the multivariate density function admits the following finite basis expansion

$$l_t(x_1, x_2) = \sum_{i=1}^{K} \beta_{i,t} h_i(x_1, x_2) \ , \tag{5}$$

which in a more general multidimensional setting can be written as

$$l_t(\boldsymbol{x}) = \sum_{i=1}^{K} \beta_{i,t} h_i(\boldsymbol{x}) \ . \tag{6}$$

This expansion let us to rewrite the functional VAR model as a factor augmented VAR model for the aggregate macroeconomic variables $\boldsymbol{y}_t$ and the factors $\boldsymbol{\beta_t}$, that is[4]

$$\begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{\beta}_t \end{bmatrix} = \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \begin{bmatrix} \boldsymbol{y}_{t-1} \\ \boldsymbol{\beta}_{t-1} \end{bmatrix} + \ldots + \boldsymbol{\Phi}_p \begin{bmatrix} \boldsymbol{y}_{t-p} \\ \boldsymbol{\beta}_{t-p} \end{bmatrix} + \begin{bmatrix} \boldsymbol{u}_{y,t} \\ \tilde{\boldsymbol{u}}_{l,t} \end{bmatrix} \ . \tag{7}$$

In this factor augmented VAR model, the dynamic behavior of the factors $\boldsymbol{\beta}_t$ is governing the time variation of firm-level characteristics over time. Macroeconomic shocks, such as aggregate TFP shocks, are driving the joint dynamics of the cross-sectional distribution of firm level characteristics and the macroeconomic aggregates through the vector $[\boldsymbol{u}'_{y,t} \tilde{\boldsymbol{u}}'_{l,t}]$, that similarly to the standard Structural VAR framework, can be interpreted as linear combinations of the structural shocks.

4. In the appendix A.1 we report the steps, which directly follow Chang et al. (2021), to derive the factor augmented representation of the functional VAR.

3 Estimation and inference

The CLR transformation maps the multivariate density function $f_t(\boldsymbol{x})$ to the space L^2 of square-integrable real measurable functions, where statistical methods for unconstrained data can be applied. For any t we compute the CLR transformation as follows:

$$l(\boldsymbol{x})^{\text{obs}} = CLR(\hat{f}(\boldsymbol{x})) \,, \tag{8}$$

where $\hat{f}(.)$ is a kernel density estimate of the distribution function on the set of grid points N^{grid}. For example, in the bivariate labor and capital example, it is obtained as

$$\hat{f}(x_1, x_2) = \frac{1}{N_1 N_2 h_1 h_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} K\left(\frac{x_1 - x_{1,i}}{h_1}, \frac{x_2 - x_{2,j}}{h_2}\right) \,. \tag{9}$$

Considering all the points on the grid, our observation equation becomes

$$\boldsymbol{l}_t^{obs} = \boldsymbol{H}\boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t \,, \tag{10}$$

where $\boldsymbol{l}_t^{obs}$ is the $N^{grid} \times 1$ vector of observed values for the CLR transformation of the multidimensional distribution function, $\boldsymbol{H}$ is the $N^{grid} \times K$ matrix of loadings, while $\boldsymbol{\beta}_t$ is the $K \times 1$ vector of scores and $\boldsymbol{\varepsilon}_t$ is the vector of noises.

To estimate the model, two alternative two-step approaches can be considered. In the first, one estimates the loadings $\boldsymbol{H}$, and then directly substitutes them into the Factor-Augmented VAR model. In the second approach, following Doz et al. (2011), the loadings $\boldsymbol{H}$ are first estimated for the basis expansion approximation, and Bayesian posterior inference is then performed on the parameters of the VAR model (7)—namely, $(\boldsymbol{\Phi}, \boldsymbol{\Sigma})$—and the latent factors $\boldsymbol{\beta}_{1:T}$. The former approach neglects estimation errors in the scores, which can be justified when the cross-sectional dimension is large. The latter takes it into account and can result in more efficient estimates, as it considers dynamic properties in the factor extraction step.

In the next sections, we describe three approaches for estimating the loadings in $\boldsymbol{H}$ and the scores $\boldsymbol{\beta}_{1:T}$, under a different set of assumptions concerning the approximation of the true CLR transformed multidimensional distribution function in (6). We then detail the estimation of the factor augmented VAR model approximation of the FunVAR model.

3.1 Unfolding and approximation by principal component analysis

One approach for estimating the loadings $\boldsymbol{H}$ needed for the approximation of the multidimensional distribution function consists in vectorizing the matrix of observable $\boldsymbol{L}_t = CLR(\hat{\boldsymbol{F}}_{\boldsymbol{t}})$, where $\hat{\boldsymbol{F}}_{\boldsymbol{t}}$ is the matrix containing the value of $\hat{f}(x_1, x_2)$ when evaluated on the chosen grid, and then performing principal component analysis on the T vectors $\boldsymbol{l}_t^{obs}$, each of dimension $N^{grid} \times 1$. This practice is known as flattening or unfolding in the context of tensor analysis (Kolda et al. 2009). In particular, going back to the labor and capital example, we can define the tensor object $\mathcal{L} \in \mathcal{R}^{N_1 \times N_2 \times T}$, which is storing the values of the CLR transformation of the multivariate density function on the finite grid for labor and capital, for all the time periods in the sample. First, we unfold the tensor obtaining the $N_1 N_2 \times T$ matrix $\tilde{\boldsymbol{L}} = [\boldsymbol{l}_1^{obs}, ..., \boldsymbol{l}_T^{obs}]$. Then, we perform principal component analysis on $\tilde{\boldsymbol{L}}$ since this matrix is stacking on its columns the $N_1 N_2$ dimensional column vectors $\boldsymbol{l}_t^{obs}$ for $t = 1, \ldots, T$. Doing so we aim at estimating the $N_1 N_2 \times K$ loading matrix $\boldsymbol{H}$ minimizing the reconstructing error, that is:

$$\min_{\boldsymbol{H}, \boldsymbol{\beta}_t} \frac{1}{T} \sum_{t=1}^{T} \|vec(\boldsymbol{L}_t) - \boldsymbol{H}\boldsymbol{\beta}_t\|^2 \ . \tag{11}$$

To estimate $\boldsymbol{H}$ we apply the singular value decomposition (SVD) to the unfolded matrix $\tilde{\boldsymbol{L}}$, that is:

$$\tilde{\boldsymbol{L}} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\boldsymbol{T}},$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal matrices, and $\boldsymbol{S}$ is a diagonal matrix containing the singular values. More specifically, we select K columns from $\boldsymbol{V}$, which represent the eigenbasis associated with the K largest eigenvalues. Note that, in this approach, PCA on the flattened data identifies the top K modes of variation without distinguishing between the individual dimensions. Consequently, it treats all dimensions as a single combined dimension, ignoring the multi-dimensional structure and the interactions specific to each dimension. The approach can be naturally extended to the case in which we have more than two dimensions, for example when d is the number of micro-level variables, we unfold the $d+1$ dimensional tensor $\mathcal{L} \in \mathcal{R}^{N_1 \times N_2 \times \ldots \times N_d \times T}$ into a $(\prod_{i=1}^{d} N_i) \times T$ matrix and then perform principal component analysis for identifying the top K modes of variation. Note also that, in general, this approach requires the estimation of $N^{grid} \times K$ loadings in the $\boldsymbol{H}$ matrix, which in small samples can be challenging. In the next sections we consider two approaches for reducing the number of parameters to be estimated, assuming specific basis expansions for the CLR transformation of the multivariate density function.

3.2 Approximation by multilinear principal component analysis

One approach that considerably reduces the number of loadings to be estimated and explicitly leverages the multi-dimensional structure of the data, applying dimensionality reduction to each mode separately, is multilinear principal component analysis. This methodology aligns with the Tucker decomposition framework (Tucker 1966), which allows for the decomposition of a tensor into mode-specific factors and a core tensor. Specifically, going back to the example on the approximation of the bivariate labor and capital distribution, we assume that the true centered log ratio transformation of the density function can be expanded as

$$l_t(x_1, x_2) = \sum_{i=1}^{K_1} \sum_{j=1}^{K_2} \beta_{ij,t} h_i(x_1) h_j(x_2) \tag{12}$$

Note that this expansion is a particular case of the more general basis expansion in (6). The representation is bilinear, meaning it is expressed as a product of components that vary across two separate dimensions, that is labor x_1 and capital x_2. This structure assumes that the function lies in a lower-dimensional subspace (rank $K_1 \times K_2$) of the full space spanned by the basis functions. In practice this implies a dimensionality reduction, as it approximates the full function by focusing on the most important modes of variation in the two separate dimensions (controlled by K_1 and K_2). The expansion can be written using the Kronecker product as

$$l_t(x_1, x_2) = (\boldsymbol{h}(x_2) \otimes \boldsymbol{h}(x_1))' \boldsymbol{\beta}_t \, . \tag{13}$$

where $\boldsymbol{h}(x_1) \otimes \boldsymbol{h}(x_2)$ is of dimension $K \times 1$ where now $K = K_1 K_2$, while $\boldsymbol{\beta}_t$ is of dimension $K \times 1$:

$$\boldsymbol{h}(x_1) = \begin{bmatrix} h_1(x_1) \\ h_2(x_1) \\ \vdots \\ h_{K_1}(x_1) \end{bmatrix}, \quad \boldsymbol{h}(x_2) = \begin{bmatrix} h_1(x_2) \\ h_2(x_2) \\ \vdots \\ h_{K_2}(x_2) \end{bmatrix}, \quad \boldsymbol{\beta}_t = \begin{bmatrix} \beta_{11,t} \\ \beta_{21,t} \\ \vdots \\ \beta_{K_1 1,t} \\ \beta_{12,t} \\ \vdots \\ \beta_{K_1 K_2,t} \end{bmatrix} .$$

Therefore, in terms of the observable CLR, (4) and (13), translate in the following bilinear form

$$\boldsymbol{L}_t = \boldsymbol{H}_1 \boldsymbol{B}_t \boldsymbol{H}_2' + \boldsymbol{E}_t \ , \tag{14}$$

where $\boldsymbol{L}_t$ is $N_1 \times N_2$, $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are the basis functions evaluated on the grid of values for labor x_1 and capital x_2, respectively of dimension $N_1 \times K_1$ and $N_2 \times K_2$. $\boldsymbol{B}_t$ is the $K_1 \times K_2$ matrix of factors at time t while $\boldsymbol{E}_t$ is the $N_1 \times N_2$ matrix of noises. Vectorizing equation (14) and exploiting the properties of the Kronecker product we get

$$\boldsymbol{l}_t^{obs} = (\boldsymbol{H}_2 \otimes \boldsymbol{H}_1)\boldsymbol{\beta_t} + \boldsymbol{\varepsilon_t} \ , \tag{15}$$

where $\boldsymbol{l}_t^{obs}$ is the $N^{grid} \times 1$ vector of observable with $N^{grid} = N_1 N_2$, $\boldsymbol{\beta}_t = vec(\boldsymbol{B}_t)$ and $\varepsilon_t = vec(\boldsymbol{E}_t)$. Note that the number of loadings is now equal to $N_1 K_1 + N_2 K_2$, while in the previous approach based on principal component analysis on the vectorized data we had $N_1 N_2 \times K$ loadings. We apply bilinear principal component analysis to estimate the loadings in $\boldsymbol{H} = (\boldsymbol{H}_2 \otimes \boldsymbol{H}_1)$. Estimation by bilinear principal component (see Ye (2004)) seeks to minimize:

$$\min_{\boldsymbol{H}_1, \boldsymbol{H}_2, \{\boldsymbol{B}_t\}} \frac{1}{T} \sum_{t=1}^{T} \left\| \boldsymbol{L}_t - \boldsymbol{H}_1 \boldsymbol{B}_t \boldsymbol{H}_2' \right\|^2 \ . \tag{16}$$

Theorem 1 in Hung et al. (2012) let us re-frame the problem in terms of the population expected Frobenius norm $\mathbb{E}\{||\boldsymbol{L} - \boldsymbol{H}_1 \boldsymbol{B} \boldsymbol{H}_2'||^2\}$. In particular, the minimizers $\boldsymbol{H}_1 \in \mathcal{O}_{K_1, \tilde{K}_1}$ and $\boldsymbol{H}_2 \in \mathcal{O}_{K_2, \tilde{K}_2}$ will be equal to the maximizers of the following problem:

$$\max_{\boldsymbol{H}_1, \boldsymbol{H}_2} \mathbb{E}\{||\boldsymbol{H}_1' \boldsymbol{L} \boldsymbol{H}_2||^2\} \ . \tag{17}$$

We use an iterative approach to estimate $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$. Starting with initial random matrices $\boldsymbol{H}_1^{(0)}$, $\boldsymbol{H}_2^{(0)}$ we iterate the following steps:

1. For fixed $\boldsymbol{H}_2^{(k)}$ update ($\boldsymbol{H}_1^{(k+1)}$) by maximizing

$$\boldsymbol{H}_1^{(k+1)} = \text{argmax}_{\boldsymbol{H}_1} T^{-1} \sum_{t=1}^{T} \left\| \boldsymbol{H}_1 \boldsymbol{L}_t \boldsymbol{H}_2^{\prime(k)} \right\|^2 \ . \tag{18}$$

2. For fixed $\boldsymbol{H}_1^{(k+1)}$ update ($\boldsymbol{H}_2^{(k+1)}$) by maximizing

$$\boldsymbol{H}_2^{(k+1)} = \text{argmax}_{\boldsymbol{H}_2} T^{-1} \sum_{t=1}^{T} \left\| \boldsymbol{H}_1^{(k+1)} \boldsymbol{L}_t \boldsymbol{H}_2' \right\|^2 . \tag{19}$$

Both maximization problems are standard eigenvalue problems, which can be formulated in terms of SVDs. As the algorithm may find only a local maximum, multiple random initial values are considered to ensure that the global maximum is found. Iterations stop when the change in the objective function is found to be lower than a predefined small ϵ, indicating convergence. Hung et al. (2012) develop the asymptotic theory for this type of order-two multilinear principal component analysis, addressing both asymptotic efficiency and the distributions of the principal components and associated projections. Their work provides comprehensive details on convergence rates and efficiency, which we refer to for further details.[5] Because principal component analysis on unfolded data requires the estimation of many parameters, multilinear principal component analysis is expected to outperform conventional principal component analysis, especially when the sample size is small to moderate. However, principal component analysis in unfolded data may provide better approximations when these interactions are weak or noisy, or the basis expansion assumed in the Tucker decomposition is too restrictive.

In the multilinear principal component analysis, it is also natural to handle more than two dimensions. For example, in the d dimensional case the basis expansion of the true CLR transformation of the distribution function just becomes

$$l_t(x_1, x_2, \ldots, x_d) = (\boldsymbol{h}(x_d) \otimes \cdots \otimes \boldsymbol{h}(x_2) \otimes \boldsymbol{h}(x_1))' \boldsymbol{\beta}_t , \tag{20}$$

where $(\boldsymbol{h}(x_d) \otimes \cdots \otimes \boldsymbol{h}(x_2) \otimes \boldsymbol{h}(x_1))$ is of dimension $\prod_{i=1}^{d} K_i \times 1$. In terms of the observables we have

$$\mathcal{L}_t = \mathcal{B}_t \times_1 \boldsymbol{H}_1 \times_2 \boldsymbol{H}_2 \times_3 \boldsymbol{H}_3 \cdots \times_d \boldsymbol{H}_d + \mathcal{E}_t , \tag{21}$$

where the notation $\times_n$ refers to the mode-n product of a tensor and a matrix and $\mathcal{B}_t$ is the core tensor of the Tucker decomposition. The iterative procedure described above is then just extended to estimate all the matrices with the loadings $\boldsymbol{H}_1, \ldots, \boldsymbol{H}_d$, each of dimension $N_i \times K_i$ for $i = 1, \ldots, d$.

5. Additionally, they propose a method for selecting the dimensionality parameters K_1 and K_2 based on a test concerning the explained proportion of total variance. This test determines whether the variance explained by the selected dimensions exceeds a predefined threshold, to which we also refer for guidance on dimensionality selection.

3.2.1 CP decomposition

The multilinear decomposition is particularly well-suited for datasets with interactions across multiple modes, since it permits to perform dimensionality reduction in a flexible manner. An alternative to the multilinear decomposition is the canonical decomposition (CANDECOMP) (Carroll et al. 1970) or parallel factors (PARAFAC)(Harshman 1970), henceforth CP decomposition. In this decomposition, the rank is shared across all modes, which means that the same number of components is used for each dimension of the tensor. In particular, going back to the bivariate labor and capital example, it is assumed that:

$$l(x_1, x_2) = \sum_{k=1}^{K} \beta_{t,k} h_k^{(1)}(x_1) h_k^{(2)}(x_2) \, . \tag{22}$$

Note that in this case the number of loadings to be estimated becomes $K(N_1 + N_2)$. The loadings can be estimated by Alternating Least Squares (ALS) (Carroll et al. 1970; Harshman 1970). Babii et al. (2024) have recently shown how to estimate the loadings of the CP decomposition by iterative principal components on the unfolded tensor along each of its dimensions. They label this estimator Tensor-PCA and derive its asymptotic properties. The CP decomposition framework also naturally extends to more than two dimensional settings, developing the expansion:

$$l(\boldsymbol{x}) = \sum_{k=1}^{K} \beta_{t,k} \prod_{i=1}^{d} h_k^{(i)}(x_i) \, . \tag{23}$$

In terms of the observables we have:

$$\mathcal{L}_t = \sum_{k=1}^{K} \beta_{t,k} \bigotimes_{i=1}^{d} \boldsymbol{h}_k^{(i)} + \mathcal{E}_t, \tag{24}$$

where $\bigotimes_{i=1}^{d} \boldsymbol{h}_k^{(i)}$ denotes the outer product of vectors $\boldsymbol{h}_k^{(i)}$ across each dimension $i = 1, 2, \ldots, d$. In general, the CP decomposition is less flexible when different modes of the data exhibit varying complexity or correlation structures. In contrast, the multilinear decomposition allows for different ranks in each mode, enabling more precise control over dimensionality reduction in each dimension.

In the empirical application, we rely on cross-validation in order to compare the basis function approximation of the multivariate density function with the three different approaches.

3.3 Factor augmented VAR

Once having obtained estimates for the loadings $\boldsymbol{H}$ and the scores $\boldsymbol{\beta}_{1:T}$, one straightforward approach is to plug these estimates directly into the Factor-augmented VAR model. An alternative strategy, following Doz et al. (2011), is to condition on the estimated loadings and jointly estimate the scores along with the parameters of the Factor-augmented VAR model. Besides accounting for the uncertainty associated with the estimation of the scores, this approach also accommodates missing densities and mixed-frequency data, as discussed in the Appendix, Section A.2. Here we provide further details of this second approach. Conditional on the estimate of the loadings $\boldsymbol{H}$ in the first step, obtained through one of the three alternative approaches proposed above, we perform Bayesian inference on the parameters of the factor augmented VAR model:

$$\boldsymbol{w}_t = \boldsymbol{\Phi}\boldsymbol{x}_t + \boldsymbol{u}_t \;, \qquad \boldsymbol{u}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}) \;, \tag{25}$$

$$\boldsymbol{l}_t^{obs} = \boldsymbol{H}\boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t \;, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \boldsymbol{I}_{N^{grid}}\sigma^2) \;, \tag{26}$$

where $\boldsymbol{w}_t = [\boldsymbol{y}_t, \boldsymbol{\beta}_t']'$ and $\boldsymbol{x}_t = [1, \boldsymbol{w}_{t-1}', \ldots, \boldsymbol{w}_{t-p}']'$. We cast the model in state space form and treat $\boldsymbol{\beta_{1:T}}$ as latent stochastic states. We need to specify a prior distribution for $(\boldsymbol{\Phi}, \boldsymbol{\Sigma})$ and for the noise variance σ^2. A common approach is to use the standard Normal-Inverse Wishart conjugate prior for $(\boldsymbol{\Phi}, \boldsymbol{\Sigma})$, which is computationally attractive in large-dimensional settings (Carriero et al. 2009; Bańbura et al. 2010). However, this imposes symmetric shrinkage across all equations in the VAR, which can be restrictive. To address this issue, we utilize the asymmetric conjugate prior proposed by Chan (2022), which permits to estimate the VAR parameters in high-dimensional settings while maintaining computational feasibility and enabling asymmetric shrinkage across equations. For σ^2, instead, we specify a standard independent Inverse Gamma prior. We combine these prior distributions with the likelihood implied by (26) and (25), obtaining the following posterior distribution for $(\boldsymbol{\beta}_{1:T}, \sigma^2, \boldsymbol{\Phi}, \boldsymbol{\Sigma})$,

$$p(\boldsymbol{\beta}_{1:T}, \sigma^2, \boldsymbol{\Phi}, \boldsymbol{\Sigma} | \boldsymbol{y}_{1:T}, \boldsymbol{l}_{1:T}) \propto p(\boldsymbol{y}_{1:T}, \boldsymbol{\beta}_{1:T} | \boldsymbol{\Phi}, \boldsymbol{\Sigma}) p(\boldsymbol{l}_{1:T} | \boldsymbol{\beta}_{1:T}, \sigma^2) p(\boldsymbol{\Phi}, \boldsymbol{\Sigma}) p(\sigma^2) \;. \tag{27}$$

In order to perform inference based on (27), we device a *Gibbs sampler* that iteratively draws from the following conditional posterior distributions

1. Draw from $p(\sigma^2 | \boldsymbol{\Sigma}, \boldsymbol{\Phi}, \boldsymbol{\beta}_{1:T}, \boldsymbol{y}_{1:T}, \boldsymbol{l}_{1:T})$.

2. Draw from $p(\boldsymbol{\Sigma}|\sigma^2, \boldsymbol{\Phi}, \boldsymbol{\beta}_{1:T}, \boldsymbol{y}_{1:T}, \boldsymbol{l}_{1:T})$.

3. Draw from $p(\boldsymbol{\Phi_0}, \boldsymbol{\Phi_1}, \ldots, \boldsymbol{\Phi_p}|\boldsymbol{\Sigma}, \sigma^2, \boldsymbol{\beta}_{1:T}, \boldsymbol{y}_{1:T}, \boldsymbol{l}_{1:T})$.

4. Draw from $p(\boldsymbol{\beta}_{1:T}|\boldsymbol{\Phi_0}, \boldsymbol{\Phi_1}, \ldots, \boldsymbol{\Phi_p}|\boldsymbol{\Sigma}, \sigma^2, \boldsymbol{\Phi}, \boldsymbol{y}_{1:T}, \boldsymbol{l}_{1:T})$.

Step 1 to 3 of the Gibbs Sampler are standard. In order to draw from the conditional distribution of the latent states in step 4 of the Gibbs Sampler, we exploit the linearity of the system, and sample the vector of the latent states jointly in a single step from $t = 1, \ldots, T$. This approach operationally borrows from Beyeler et al. (2018) and Chan et al. (2023). More precisely, we define:

$$\boldsymbol{w} = \boldsymbol{S}_o \boldsymbol{y} + \boldsymbol{S}\boldsymbol{b} \,, \tag{28}$$

where $\boldsymbol{w} = ([\boldsymbol{y}_1', \boldsymbol{\beta}_1'], \ldots ([\boldsymbol{y}_T', \boldsymbol{\beta}_T'])' = (\boldsymbol{w}_1', \ldots, \boldsymbol{w}_T')'$ is a $T(n_y + K) \times 1$ vector that stores the macro aggregates together with the latent factors from the cross sectional distribution, while $\boldsymbol{y} = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_T')'$ and $\boldsymbol{b} = (\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_T')'$ respectively store the macroeconomic aggregates and the latent states. The selection matrices $\boldsymbol{S}_o$ and $\boldsymbol{S}$ are given by $\boldsymbol{S}_o = \boldsymbol{I}_T \otimes \boldsymbol{S}_{oo}$ and $\boldsymbol{S} = \boldsymbol{I}_T \otimes \boldsymbol{S}_{ff}$ with $\boldsymbol{S}_{oo} = [\boldsymbol{I}_{n_y}; \boldsymbol{0}_{K \times n_y}]$ and $\boldsymbol{S}_{ff} = [\boldsymbol{0}_{n_y \times K}; \boldsymbol{I}_K]$. Defining

$$\boldsymbol{c_\Phi} = \begin{bmatrix} \boldsymbol{\Phi}_0 + \sum_{j=1}^{p} \boldsymbol{\Phi}_j \boldsymbol{w}_{1-j} \\ \boldsymbol{\Phi}_0 + \sum_{j=2}^{p} \boldsymbol{\Phi}_j \boldsymbol{w}_{2-j} \\ \vdots \\ \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_p \boldsymbol{w}_0 \\ \boldsymbol{\Phi}_0 \\ \vdots \\ \boldsymbol{\Phi}_0 \end{bmatrix}, \quad \boldsymbol{D_\Phi} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ -\boldsymbol{\Phi}_1 & \boldsymbol{I} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ -\boldsymbol{\Phi}_2 & -\boldsymbol{\Phi}_1 & \boldsymbol{I} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \boldsymbol{0} & \boldsymbol{0} \\ -\boldsymbol{\Phi}_p & \cdots & -\boldsymbol{\Phi}_2 & -\boldsymbol{\Phi}_1 & \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \ddots & \ddots & \ddots & -\boldsymbol{\Phi}_1 & \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & -\boldsymbol{\Phi}_p & \ldots & -\boldsymbol{\Phi}_2 & -\boldsymbol{\Phi}_1 & \boldsymbol{I} \end{bmatrix}, \tag{29}$$

and the products $\boldsymbol{G}_y = \boldsymbol{D_\Phi} \boldsymbol{S}_o$ and $\boldsymbol{G} = \boldsymbol{D_\Phi} \boldsymbol{S}$, we can rewrite the factor augmented VAR as:

$$\boldsymbol{G}_o \boldsymbol{y} + \boldsymbol{G}\boldsymbol{b} = \boldsymbol{c_\Phi} + \boldsymbol{u} \,, \qquad \boldsymbol{u} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_T \otimes \boldsymbol{\Sigma}) \,, \tag{30}$$

and the observation equation in terms of $\boldsymbol{\beta}$ as follows:

$$\boldsymbol{l} = \boldsymbol{M}\boldsymbol{b} + \boldsymbol{\epsilon} \,, \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{TN^{grid}}) \,, \tag{31}$$

where $\boldsymbol{l} = vec([\boldsymbol{l}_1, \ldots, \boldsymbol{l}_T])$ is $N^{grid}T \times 1$ vector of observable values of the CLR transformation of

the multidimensional distribution function on the grid points, while $\boldsymbol{M} = \boldsymbol{I}_T \otimes \boldsymbol{H}$ is the matrix containing the loadings for the basis functions approximation of all cross-sectional distributions for $t = 1, \ldots, T$. Equations (30) and (31) imply that we can sample the latent states jointly from the conditional posterior distribution of $\boldsymbol{b}$ given by

$$p(\boldsymbol{b}|.) \sim \mathcal{N}\left(\bar{\boldsymbol{K}}^{-1}\left(\frac{1}{\sigma^2}\boldsymbol{M}'\boldsymbol{l} + \boldsymbol{K}\boldsymbol{\mu}\right), \bar{\boldsymbol{K}}^{-1}\right) , \tag{32}$$

where

$$\bar{\boldsymbol{K}} = \frac{1}{\sigma^2}\boldsymbol{M}'\boldsymbol{M} + \boldsymbol{G}'(\boldsymbol{I}_{\boldsymbol{T}} \otimes \boldsymbol{\Sigma})^{-1}\boldsymbol{G} ,$$

with $\boldsymbol{\mu} = \boldsymbol{K}^{-1}\boldsymbol{G}'(\boldsymbol{I}_T \otimes \boldsymbol{\Sigma})(\boldsymbol{c}_{\boldsymbol{\Phi}} - \boldsymbol{G}_o\boldsymbol{y})$ and $\boldsymbol{K} = \boldsymbol{G}'(\boldsymbol{I}_T \otimes \boldsymbol{\Sigma})^{-1}\boldsymbol{G}$. Note that the computational intensity of this step does not depend on the number of grid points on which the CLR transformation of the distribution function is evaluated N^{grid}, since the product matrices $\boldsymbol{M}'\boldsymbol{M}$ and $\boldsymbol{M}'\boldsymbol{l}$ are respectively of dimension $TK \times TK$ and $TK \times 1$ and are a function of $\boldsymbol{H}$ which is computed outside of the Gibbs Sampler. Computational complexity hinges instead on K, the number of basis used for the expansion of the CLR transformation of the distribution function. When the combination TK is large, the latent scores can alternatively be drawn exploiting the standard Kalman filter and smoother step à la Carter et al. (1994).

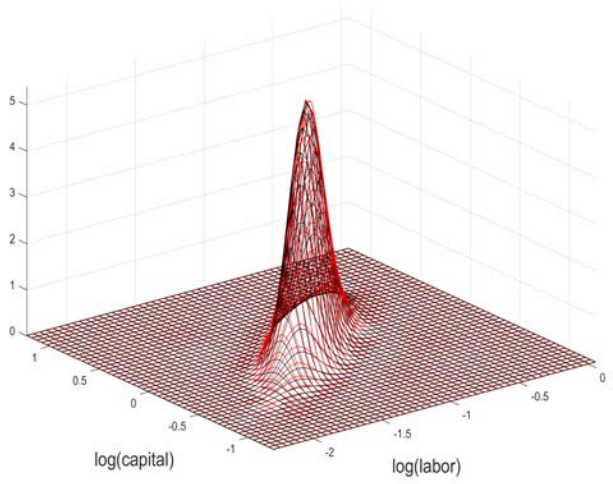
4 Simulation from an heterogeneous firm model

In this section we assess the finite sample performance of the FunVAR model to recover the impulse response functions of the macroeconomic aggregates and the cross-sectional distributions to macroeconomic shocks generated by a fully structural heterogeneous firms model. We consider the version of the standard heterogeneous firm model of Khan et al. (2008) extended in Winberry (2018). Since the model by Winberry (2018), once solved and approximated around the steady state, implies VAR dynamics in terms of the moments of the bivariate log-labor and log-capital distribution, our FunVAR model is inherently misspecified. Aware of this misspecification, we use this simulation study to assess whether and how well the FunVAR model can qualitatively replicate the effects of macroeconomic shocks on both the cross-sectional distributions and macroeconomic aggregates of the heterogeneous firm model.

In Winberry's model, firms are subject to both idiosyncratic and aggregate productivity shocks. Firms are subject to heterogeneous adjustments costs. These costs play a critical role in the decision-

making process across different firms regarding both production and input-allocation.[6] As a result of this decision-making process, the cross-sectional joint distribution of firm-level capital and labor evolves dynamically over time. Our specification of the FunVAR dynamics in terms of the joint distribution of labor and capital, rather than the marginal distributions, is designed to precisely capture the effects of structural shocks on the reallocation of both inputs. In other words, our model aims to capture the interdependence between these inputs in response to shocks, showing how firms adjust their allocation of labor and capital simultaneously. We simulate the heterogeneous firm model for $T = 250$ periods. The calibration of the parameters to simulate from the heterogeneous firm model directly follows Winberry (2018). Figure 2, shows one example of simulated bivariate density of $log(k), log(l)$ from the heterogeneous agents model and its approximation by multilinear principal component analysis with 9 basis functions obtained with $K_1 = 3$ and $K_2 = 3$.

Figure 2: One period simulated bivariate density of $log(k), log(l)$ from Winberry (2018) model and its approximation though bilinear principal components analysis



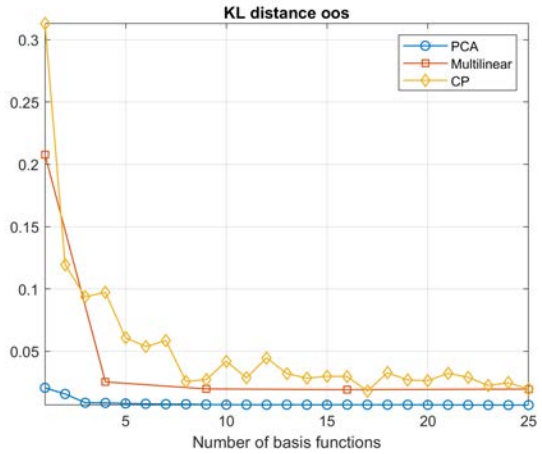
Notes: The figure shows the true (red) and the approximated (black) bivariate log-capital and log-labor distribution from for one sample period in the simulation.

To compare the approximation of the bivariate distribution across unfolded principal component analysis, bilinear principal component analysis and CP and select the number of basis functions we use cross-validation. In particular, we divide the sample in training set and test set, and compare the Kullback–Leibler divergence to the kernel density estimate of $log(k), log(l)$ in the test set. Figure 3 shows the comparison of the as a function of the number of basis functions, used for the approximation of the bivariate density. In this case, we find PCA on the unfolded data to

6. We refer to the paper of Winberry (2018) for further details on the model.

perform better than both bilinear principal component analysis and the CP decomposition. Given the simple bivariate log-normal specification, with time varying first and second moment, three basis functions provide a reliable approximation of the bivariate density function. We focus on the

Figure 3: Cross-Validated Mean KL Divergence on the Test Set: PCA vs. Multilinear PCA vs. CP



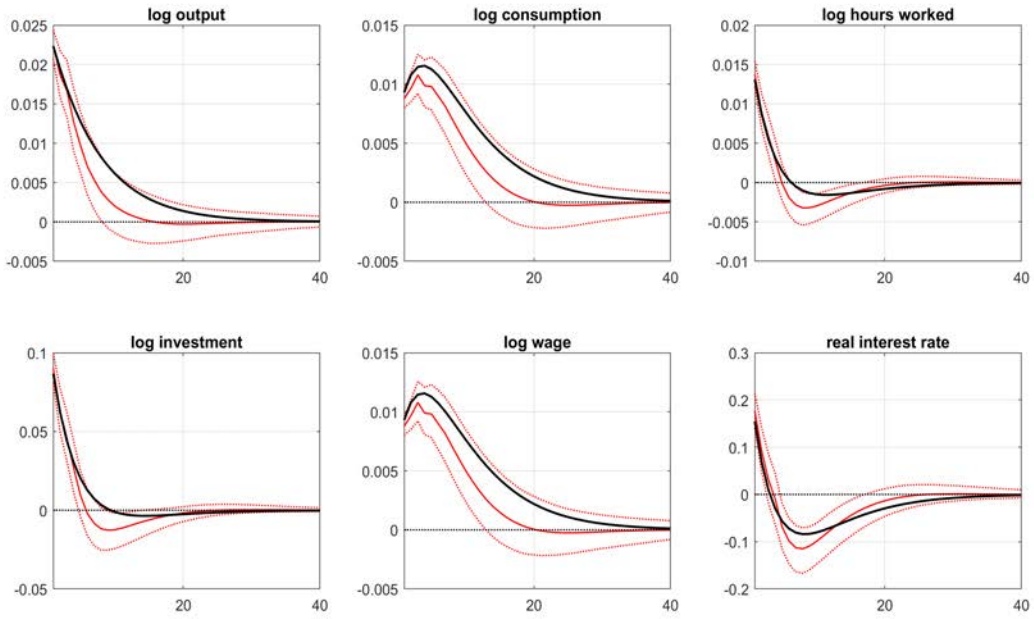
Notes: The figure shows the average KL distance to the kernel density estimate of the bivariate density function $f(log(k), log(l))$ in the test set in the cross-validation.

estimated responses of the aggregate macroeconomic time series and the cross-sectional distributions to a one standard deviation aggregate TFP shock. Figure 4, shows the impulse response functions of the main macroeconomic variables to a total factor productivity shock in the heterogeneous firm model. The figure reports in black the true responses. The red dashed lines are the 5^{th} and 95^{th} credible sets while the solid red line is the posterior median estimate obtained from our FunVAR. The TFP shocks are identified exploiting the exogeneity of the simulated TFP series. The model correctly recovers the dynamics of output, consumption, hours worked, investment and wages after the aggregate TFP shock hitting the economy.

Figure 5 reports the Functional Impulse Response Functions (FIRF) of the bivariate distribution of firm-level capital and labor. FIRFs are obtained by computing the difference in the mass between the bivariate log-capital and log-labor distributions after the TFP shock has occurred and the steady state distribution (this is reported in the z axis). We show the FIRF after 4 periods (one year), 8 periods (2 years) and 24 periods (6 years). In the figures in the left panels, we show the true FIRF, while in those on the right panels we report the posterior median FIRF estimated with the FunVAR on the simulated data. The posterior mean estimate from the FunVAR model correctly tracks the evolution of the bivariate density after the aggregate TFP shock hitting the economy.

As we are also concerned about whether the effects of a TFP shock are efficiently estimated,

Figure 4: Responses to a TFP shock in the heterogeneous firm model



Notes: The figure shows the Impulse Response Functions (IRFs) of the macroeconomic aggregates to a one standard deviation TFP shock. In black we report the IRF of the Winberry (2018) heterogeneous firm model. In red bold line we report the posterior mean estimate from the FunVAR while in dashed red line we report the 5^{th} and 95^{th} credible sets.

in Figure 6 we report the contours of the bivariate FIRF. The upper panels report the contours from the true FIRF in the heterogeneous firm model. The lower panels instead report the posterior median value, only in the case the 65 % credible regions of the posterior distribution do not contain zero. The figure shows that the effects of the TFP shocks are precisely estimated, and in general more accurately estimated at lower horizons.

Overall, this exercise with simulated data shows that the FunVAR is reliable also in finite samples to recover the distributional effects of TFP shocks.

5 The distributional effects of TFP shocks

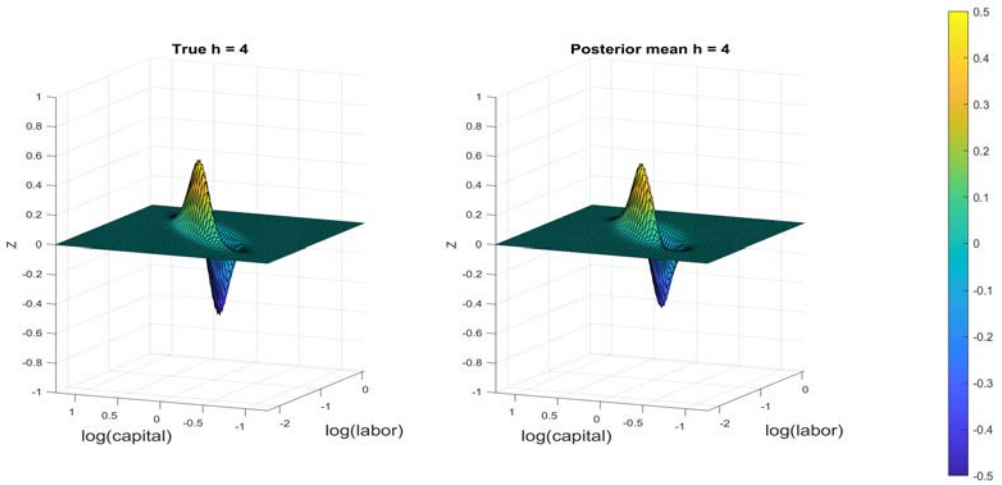
We now exploit the FunVAR model to examine the effects of TFP shocks on the US economy. An extensive literature dating back to Kydland et al. (1982) has explored the role of TFP shocks in driving business cycle fluctuations. This ongoing debate underscores the complexity of understanding how technology shocks affect firms allocation of capital and labor inputs. We use the FunVAR to investigate the effects of TFP shocks both on the macroeconomic aggregates and the firm-level joint capital and labor distribution.

5.1 Aggregate and firm-level data

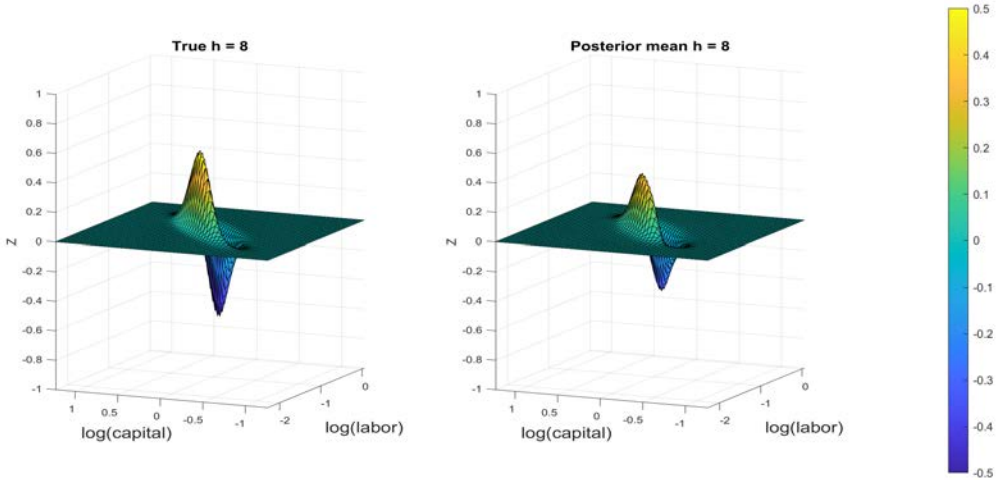
Our analysis leverages the Compustat dataset to extract the cross-sectional distribution of U.S. firm-level labor and capital. The Annual Compustat dataset provides joint data on firm-level capital and employees, while firm-level capital data is also available at a quarterly frequency through the Quarterly Compustat dataset. In what follows, we present results based a quarterly dataset obtained by micro-level interpolation to estimate missing intra-year labor observations. An alternative approach is to rely on the annual micro-level dataset for both capital and labor, and use the state-space framework outlined in Section A.2 to combine the yearly distributions with the quarterly aggregate macroeconomic time series. More specifically, we consider cross sectional data on firm's balance sheets from 1984-Q4 to 2019-Q4. The raw data are cleaned and transformed mostly following Ottonello et al. (2020b), as detailed in Appendix A.3. The final dataset is comprised of cross sections with an average size of 2.809 firms. Since balance sheet data are expressed in nominal terms (millions of dollars), we obtain the value of real capital dividing by the implicit price deflator of the nonresidential gross private domestic fixed investment. In order to clean the micro data from low frequency fluctuations and concentrate on cyclical fluctuations, we log-linearly-detrend the

Figure 5: Responses of $f(\log(k), \log(l))$ to a TFP shock in the heterogeneous firm model

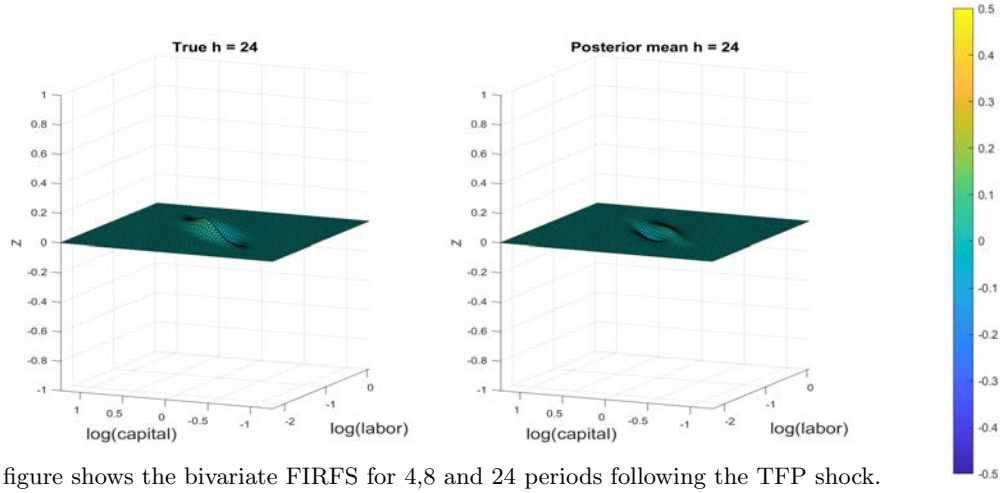
(a) $h = 4$



(b) $h = 8$

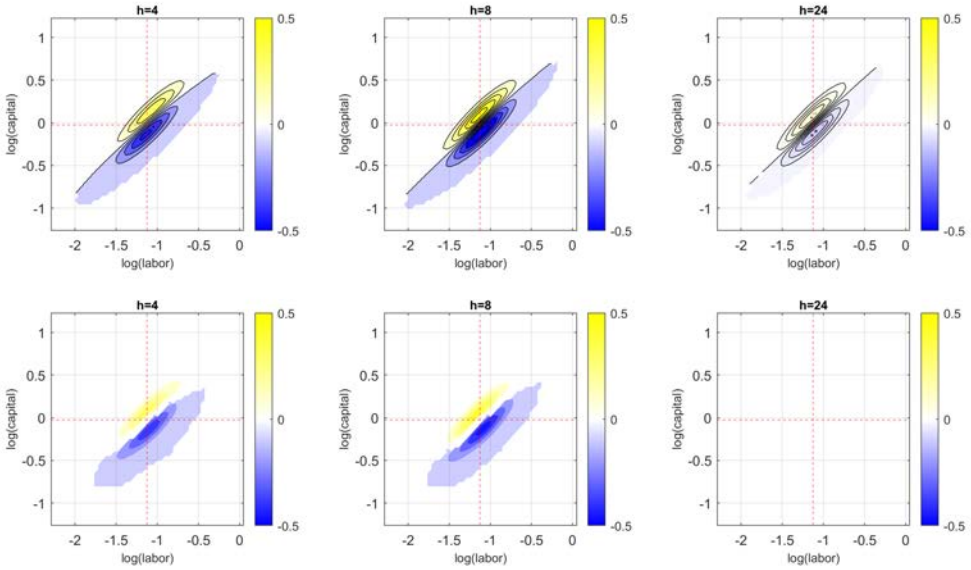


(c) $h = 24$



Notes: The figure shows the bivariate FIRFS for 4,8 and 24 periods following the TFP shock.
On the left hand side the true FIRFS, while on the right hand side the posterior mean estimate from the FunVAR.

Figure 6: Contours from the bivariate FIRF of $f(\log(l), \log(k))$



Notes: The figure shows the contours from the bivariate FIRFs for 4,8 and 24 periods following the TFP shock. The upper panel reports the true FIRF from the heterogeneous firm model, while the lower panel reports value for the posterior mean estimate from the FunVAR only if the $15^{\text{th}} - 85^{\text{th}}$ credible set does not contain zero. Dashed red lines are the steady state mean values in the heterogeneous firm model.

level of capital and labor at the firm level. Concerning the data on the macroeconomic aggregates, these are taken from FRED-QD (Federal Reserve Economic Data Quarterly Dataset). We consider real gross domestic product (`GDPC1`), real personal consumption expenditures (`PCECC96`), non-farm payroll employment (`PAYEMS`), nonresidential real private fixed investment (`PNFIx`) and real hourly non-farm business sector compensation (`COMPRNFB`) and the real interest rate (obtained as `FEDFUND` - $\pi^{\texttt{CPIAUSL}}$).[7] All the variables, excluding the real interest rate, are log-linearly detrended.

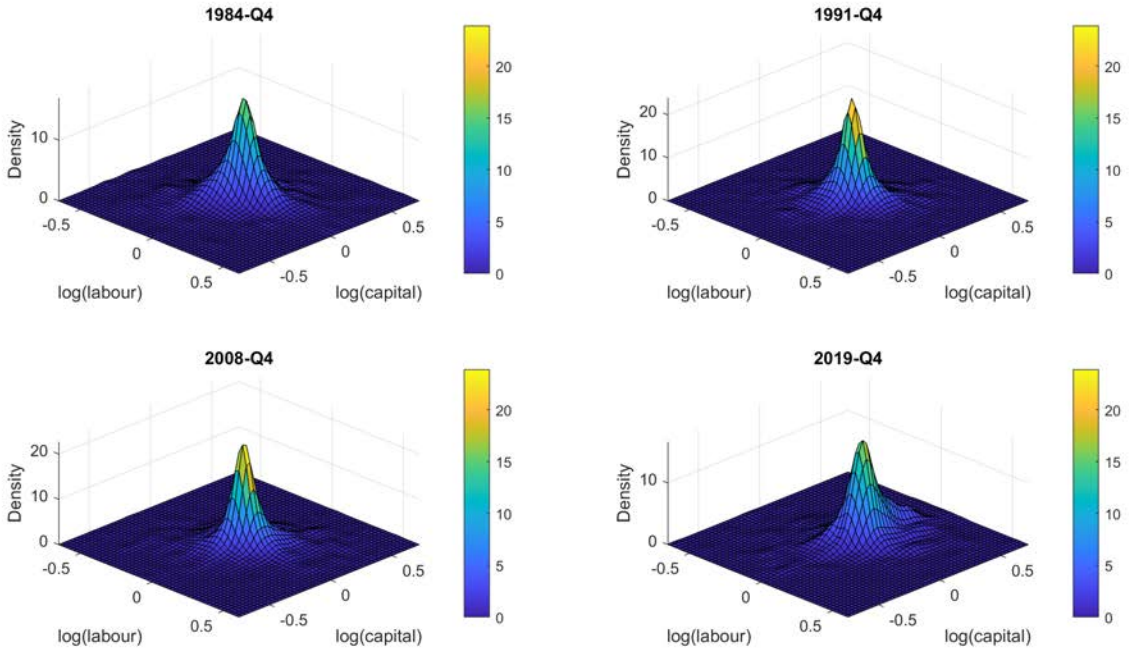
5.2 Aggregate and disaggregated response to TFP shocks

We estimate bivariate density functions from 1985-Q4 to 2019-Q4 for $f(log(K), log(l))$ using the Kernel density estimator in (9) with Gaussian Kernel and bandwidth set according to Silverman's rule (Silverman 1986). Figure 7 shows bivariate kernel density estimate of $f(log(k), log(l))$ for some selected periods in the sample.

To choose among the principal component on the vectorized tensor, bilinear principal component analysis and CP decomposition for approximating the bivariate density functions, we rely on cross validation. Figure 8 shows the average KL divergence to the kernel density estimates on the test

7. $\pi^{\texttt{CPIAUSL}}$ is defined as the yearly inflation rate obtained from the `CPIAUSL` price series.

Figure 7: Bivariate kernel density estimate of $f(log(k), log(l))$ for some selected periods.



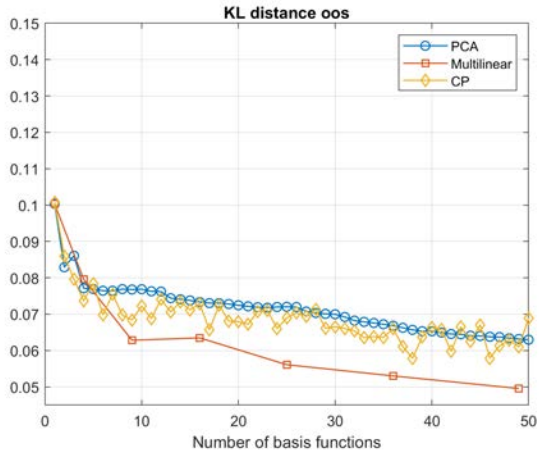
Notes: The figure shows the bivariate kenrnel density estimates for 1985-Q4, 1991-Q4, 2008-Q4 and 2019-Q4.

set across the different methods. In the appendix we report results for the same cross-validation exercise based on the average Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE). Based on the cross-validation we find bilinear principal component analysis to outperform both principal component on the unfolded tensor and CP decomposition and exploit this approach for approximating the bivariate density in the FunVAR. In particular, we consider the bilinear approximation with $K_1 = 5$ and $K_2 = 5$, for a total of 25 factors.

We estimate our model on the sample of US data that goes from 1984-Q4 to 2019-Q4. The FunVAR comprises seven macroeconomic variables, being the Fernald (2014) TFP series, real personal consumption expenditures, non-farm payroll employment, nonresidential real private fixed investment, real hourly non-farm business sector compensation and real interest rate.[8] In what follows, we present evidences based on the identification of the aggregate TFP shock trough the internal instrument procedure, exploiting the TFP measure developed in Fernald (2014) as our exogenous proxy for the TFP shocks. More in detail, we assume that the TFP proxy is contemporaneously exogenous with respect to the other variables entering the VAR. We exploit the Cholesky decom-

8. We include non-farm payroll employment to match with the micro distribution of number of employees in from Compustat.

Figure 8: Cross-Validated Mean KL Divergence on the Test Set: PCA vs. Multilinear PCA vs. CP

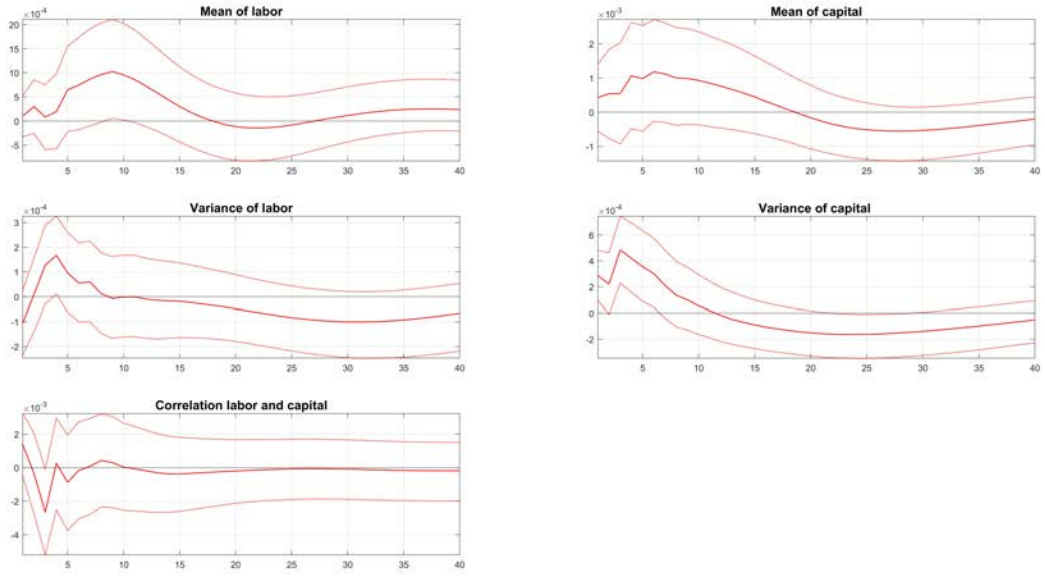


Notes: The figure shows the average KL distance to the Kernel density estimate of the bivariate density function $f(log(k), log(l))$ in the test set in the cross-validation.

position of the variance covariance matrix $\mathbf{\Sigma}$ to identify the column of the impact matrix of the structural VAR, which relates to the effect of the TFP shock both on the aggregate macroeconomic variables and the bivariate labor and capital distributions.

Using the FunVAR framework, we trace how a TFP shock propagates at the firm level by examining its impact on the joint distribution of log labor and log capital. The shock produces a persistent shift in this distribution: resources are reallocated toward both inputs, and the dispersion of capital and labor across firms widens. Figure 9 shows the impulse response functions of selected moments of the bivariate distribution, that are means, variances and correlation coefficient following the one standard deviation TFP shock. After the shock, firms on average expand their endowments of both inputs. The effects of the TFP the shock on the bivariate distributions are persistent, with the increase in the mean of labor distribution peaking after 8 quarters. The dispersion of capital across firms increases on impact, as it does the dispersion of labor after one year, while the effects of the shock does not seem to affect the correlation between capital and labor. To analyze more precisely heterogeneous adjustments in both inputs, we look at the bivariate functional impulse response functions following the one standard deviation TFP shock. Figure 10 reports the posterior median estimate of the steady state distribution of log-labor and log-capital (left upper panel) with the corresponding contour levels (right upper panel). As expected, the steady state bivariate distribution shows a clear positive correlation concerning the endowment of both inputs. The lower panels report the contour plots relative to the change in the mass w.r.t the steady state value in the bivariate density of labor and capital, for different horizons after the TFP shock has occurred.

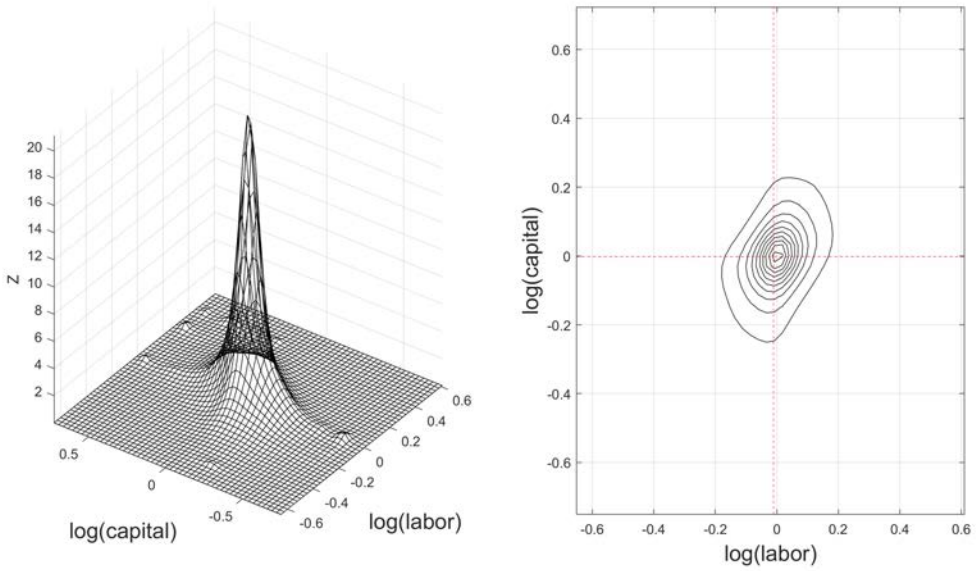
Figure 9: Impulse response functions of means, variances and correlation of the joint labor and capital distribution to a TFP shock



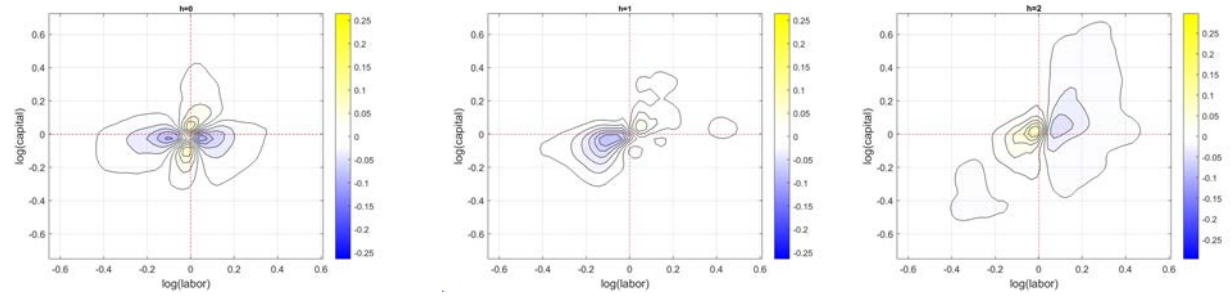
Notes: The figure shows the Impulse Response Functions (IRFs) of the mean, variances and correlation of the joint labor and capital distribution to a one standard deviation TFP shock. In red bold line we report the posterior mean estimate from the FunVAR while in dashed red line we report the 15^{th} and 85^{th} credible sets.

We show the effect on impact $h = 0$, one quarter ahead $h = 1$, two quarters ahead $h = 2$, one year ahead $h = 4$, two years ahead $h = 8$ and 15 quarters ahead $h = 15$. It turns out that after the shock firms expand both capital and labor simultaneously, indicating that they are growing and scaling their operations in response to the productivity gains from the shock. This is reflected by the increase of the mass of firms with both capital and labor above their steady state level. The peak of the effect on the bivariate distribution is between one year and 2 years. After 15 quarters, the effects of the shock gets reabsorbed. In the appendix, Figure 17 and 18 show the functional IRFs associated to the marginal distributions of log-capital and log-labor following the TFP shock. In the appendix, Figure 16, we also report the impulse responses of the aggregate macroeconomic variables to the aggregate TFP shock. The response of the macroeconomic aggregates broadly align with the responses in the real business cycle model. In particular, following the TFP shock, real output, real consumption and real investment all increase on impact, with the effect of the variables peaking after 8 quarters (two years).

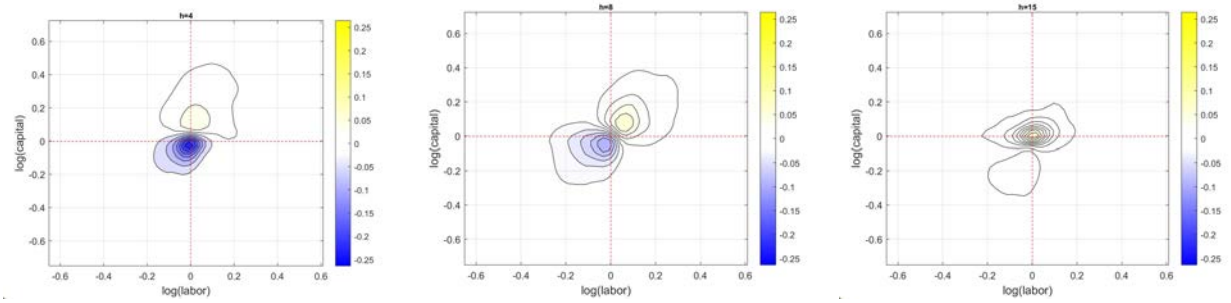
Figure 10: Response of the cross sectional distributions to an aggregate TFP shock



(a) Posterior mean steady state distribution of $\log(l), \log(k)$.



$h = 0$ $h = 1$ $h = 2$



$h = 4$ $h = 8$ $h = 15$

(b) Contours of the posterior mean estimate of the bivariate labor and capital distribution FIRF following an aggregate TFP shock.

Notes: The panel above shows the estimated posterior mean steady state distribution of $\log(l)$ and $\log(k)$. The panel below shows the contours of the posterior mean estimate of the bivariate FIRF following an aggregate TFP shock for 1, 4, 8, and 16 periods ahead. The dashed line reports the posterior mean steady state values.

6 The distributional effects of monetary policy shocks

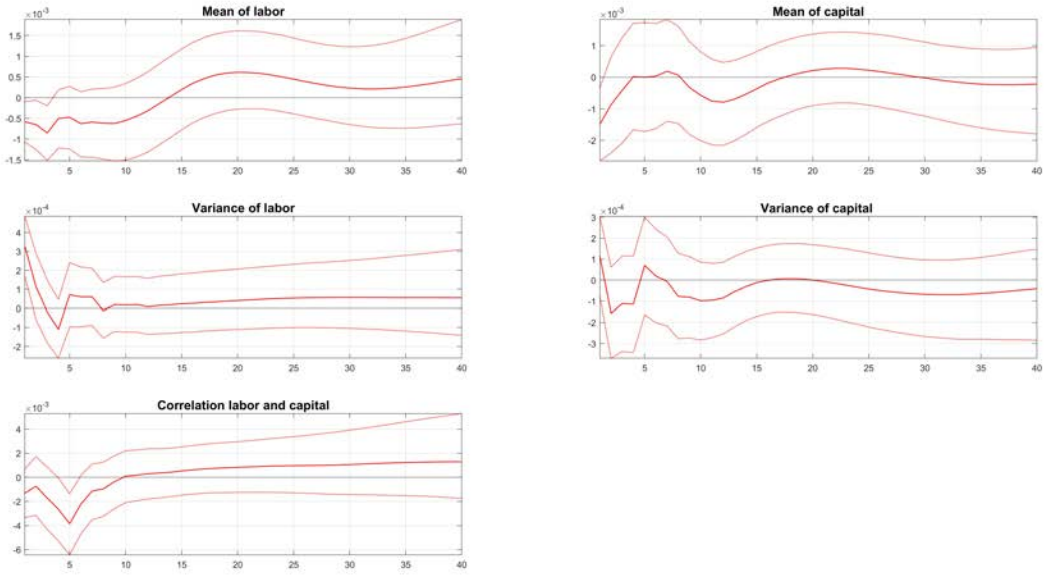
We also exploit our funVAR model to study the propagation of monetary policy shocks in the U.S. economy, analyzing both the effects on the macroeconomic aggregates and on the joint firm-level distribution of capital and labor. As in the previous section, we consider a functional VAR model for the bivariate capital and labor distribution together with macroeconomic aggregates. Concerning the macroeconomic aggregates, we consider industrial production, the unemployment rate, commodity price index, CPI index, the excess bond premium (Gilchrist et al. 2012) and the two-year Treasury yield. All variables exempt the excess bond premium are log-linearly detrended. For the identification of the monetary policy shocks, we leverage an internal instrument, being the orthogonalized series of high-frequency U.S. monetary-policy surprises by Bauer et al. (2023). This series isolates the unexpected part of 30-minute eurodollar-futures rate jumps around Fed events by removing variation predictable from pre-announcement macro-financial data, yielding a clean proxy for exogenous monetary-policy shocks.

Figure 11 traces the effects of a restrictive monetary policy shock on the bivariate log-labor and log-capital distribution. Figure 12, instead, shows the response of the bivariate capital and labor distribution to monetary policy shocks, allowing the detection of heterogeneous adjustments in the two inputs. Following the shock, firms on average reduce both labor and capital. The effects of the restrictive monetary policy shock on the bivariate distribution are found to be less persistent compared to the effects of TFP shocks, with the impact vanishing after one year. In addition to reducing both labor and capital, the shock also attenuate their co-movement. In fact, the labor–capital correlation falls sharply in the short run and only gradually normalizes as employment and investment realign. In the appendix, Figures 20 and 21 report the functional IRFs associated with the two marginal distributions. The aggregate effects of the restrictive monetary policy shock are also reported in the appendix, in Figure 19, revealing a sudden increase in the excess bond premium following the monetary policy shock and a contraction in industrial production peaking after one year.

7 Cross-sectional uncertainty and macroeconomic fluctuations

In this last section, we exploit our funVAR model to leverage micro-economic distributional data on firms output, labor and capital to identify the effects of cross-sectional uncertainty shocks on aggregate macroeconomic fluctuations. Dating back to the seminal contribution of Bloom (2009), a

Figure 11: Impulse response functions of means, variances and correlation of the joint labor and capital distribution to a monetary policy shock

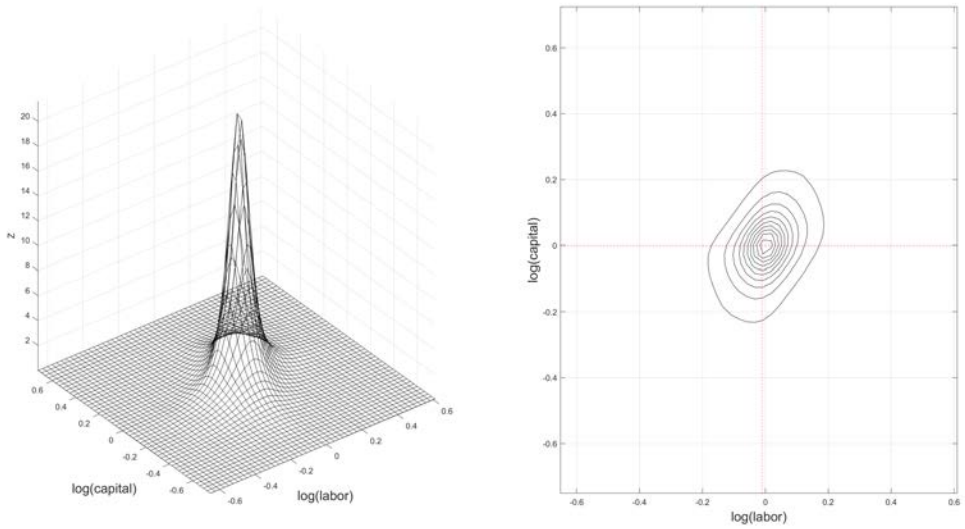


Notes: The figure shows the Impulse Response Functions (IRFs) of the mean, variances and correlation of the joint labor and capital distribution to a one standard deviation TFP shock. In red bold line we report the posterior mean estimate from the FunVAR while in dashed red line we report the 15^{th} and 85^{th} credible sets.

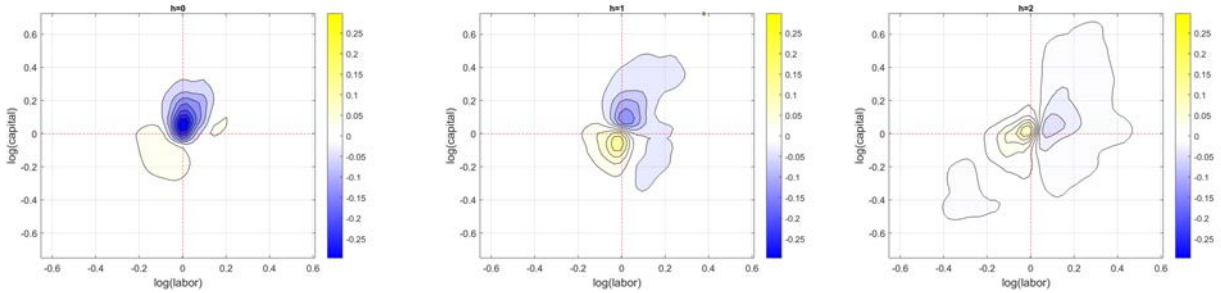
substantial body of literature has examined the macroeconomic effects of uncertainty shocks both from a theoretical and empirical perspective. Within this strand of literature, empirical studies have predominantly focused on aggregate uncertainty, aiming to measure it and trace its business cycle implications (Jurado et al. 2015; Carriero et al. 2018). While much of the empirical work emphasizes the effects of aggregate uncertainty shocks, a significant number of structural models highlight the importance of cross-sectional uncertainty as a key driver of macroeconomic fluctuations (Bloom 2009; Christiano et al. 2014; Bloom et al. 2018; Arellano et al. 2019). Indeed, several theoretical contributions point to the cross-sectional component of uncertainty as the critical force behind business cycle dynamics. For instance, Christiano et al. (2014) show that fluctuations in the dispersion of firm-level productivity—so-called risk shocks—play a central role in driving macroeconomic outcomes by influencing the availability of credit to entrepreneurs, reducing entrepreneurial investment and thereby depressing output, consumption, and employment.

Despite these theoretical insights, empirical evidence on the dynamic macroeconomic effects of cross-sectional uncertainty shocks remains scarce. A notable exception is the recent work by Dew-Becker et al. (2023), who develop a forward-looking measure of cross-sectional uncertainty using

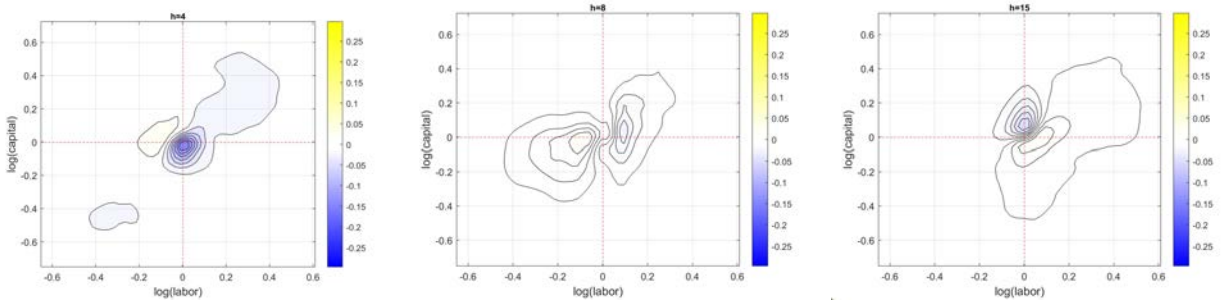
Figure 12: Response of the cross sectional distributions to a monetary policy shock



(a) Posterior mean steady state distribution of $(\log(l), \log(k))$.



$h = 0$ $h = 1$ $h = 2$



$h = 4$ $h = 8$ $h = 15$

(b) Contours of the posterior mean estimate of the bivariate labor and capital distribution FIRF following a monetary policy shock.

Notes: The panel above shows the estimated posterior mean steady state distribution of $\log(l)$ and $\log(k)$. The panel below shows the contours of the posterior mean estimate of the bivariate FIRF following a contractionary monetary policy shock for 1,2, 4, 8, and 15 periods ahead. The dashed line reports the posterior mean steady state values.

stock options data from individual firms. Their analysis reveals a nuanced relationship between cross-sectional uncertainty and aggregate activity, primarily based on correlation and predictive regressions.

Here, we directly leverage micro-economic distributional data on firms output, labor and capital to identify the effects of cross-sectional uncertainty shocks on the macroeconomic aggregates, by considering the following FunVAR model:

$$\boldsymbol{y_t} = \boldsymbol{c_y} + \sum_{s=1}^{p} \boldsymbol{B}_{l,yy}\boldsymbol{y_{t-s}} + \sum_{s=1}^{p} \int \boldsymbol{B_{s,yl}}(\boldsymbol{x}) l_{t-s}(\boldsymbol{x}) d\boldsymbol{x} + \boldsymbol{u_{y,t}} \ , \tag{33}$$

$$l_t(\boldsymbol{x}) = c_l(\boldsymbol{x}) + \sum_{s=1}^{p} \boldsymbol{B}_{s,ly}(\boldsymbol{x})\boldsymbol{y}_{t-s} + \sum_{s=1}^{P} \int B_{ll}(\boldsymbol{x}, \boldsymbol{x}') l_{t-s}(\boldsymbol{x}') d\boldsymbol{x}' + u_{l,t}(\boldsymbol{x}) \ . \tag{34}$$

where $\boldsymbol{x} = [log(output), log(k), log(l)]$ is the vector of firms characteristics and $l(\boldsymbol{x})$ is the centered-log-ratio transformation of the trivariate distribution of firm's output, capital and labor, while $\boldsymbol{y}_t$ is a vector of macroeconomic aggregates. The finite dimensional approximation of the FunVAR model is a reduced form factor augmented VAR model with transition equation given by:

$$\boldsymbol{w}_t = \boldsymbol{\Phi}\boldsymbol{x}_t + \boldsymbol{u}_t \ , \qquad \boldsymbol{u}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}) \ , \tag{35}$$

where $\boldsymbol{w}_t = [\boldsymbol{y}_t, \boldsymbol{\beta}_t']'$. We assume that the relationship between the VAR residuals and the cross-sectional uncertainty shocks ε_t^{CSU}, is as follows:

$$\boldsymbol{u}_t = \boldsymbol{\Sigma_{tr}}\boldsymbol{q}\varepsilon_t^{CSU} \tag{36}$$

where $\boldsymbol{\Sigma_{tr}} = chol(\boldsymbol{\Sigma})$ is the lower triangular Cholesky factor of the reduced form residuals variance covariance matrix and $\boldsymbol{q}$ is an orthonormal vector. The contemporaneous effect of cross-sectional uncertainty shocks on $\boldsymbol{w}_t$ is therefore given by

$$\begin{bmatrix} \boldsymbol{a}_{macro} \\ \boldsymbol{a}_{micro} \end{bmatrix} = \boldsymbol{\Sigma_{tr}}\boldsymbol{q} \tag{37}$$

where $\boldsymbol{a}_{macro}$ is the effect of the uncertainty shocks on the macroeconomic aggregates, while $\boldsymbol{a}_{micro}$ is the effect of the uncertainty shocks on the scores $\boldsymbol{\beta}_{t;K}$ and consequently on the joint distribution of firm level capital and labor. Following Bloom (2009) and Christiano et al. (2014), we think of uncertainty shocks as the main drivers of cross-sectional dispersion of firm productivity. In terms

of the observables, these shocks manifest as exogenous changes in the cross-sectional dispersion of firms' output, conditional on their capital and labor inputs.[9] Accordingly, we label a cross-sectional uncertainty shock as the structural shock which maximizes the unexpected variation of the cross-sectional dispersion of firms sales on impact, conditionally on firm's contemporaneous level of capital and labor. In practice, cross-sectional uncertainty shocks are constructed as distributional shocks that maximize the variation in the conditional variance of firm's output on impact. These shocks are operationally identified by finding the orthogonal rotation $\boldsymbol{q}$ that maximizes the change in the variance of output given capital and labor on impact. The FunVAR comprises the same variables considered in the empirical application concerning the effects of TFP shocks in Section 5. We focus on investigating the effect of the cross-sectional uncertainty shocks on real GDP, consumption, investment hours worked and real wages. Figure 13 shows the response of the macroeconomic aggregates to the cross-sectional uncertainty shocks. The time series of the estimated cross-sectional uncertainty shocks is instead shown in the appendix, Figure 22.

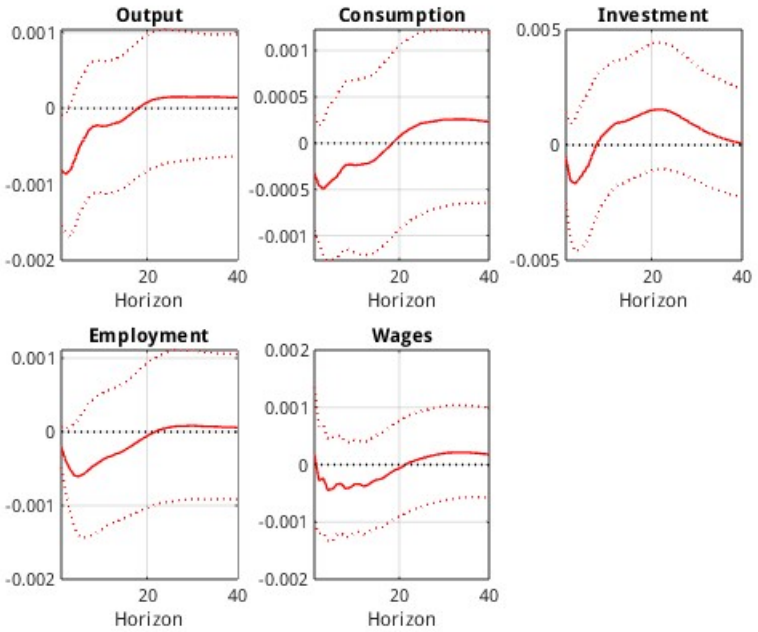
The posterior mean estimates of the impulse response function, show that cross-sectional uncertainty shocks generate contraction in investment, output, employment and consumption. This is in line with the theoretical predictions in Bloom (2009) and Christiano et al. (2014), where risk shocks generate investment falls and a decline in the purchase of goods, output, consumption, and employment. As a caveat, while the 15th–85th credible interval for the effect of cross-sectional uncertainty on output is unambiguously contractionary and excludes zero, the credible intervals for the effects on the other aggregates are relatively wide, reflecting less precise estimates of these effects.

9. In particular, given a standard firm level production function:

$$log(output_{it}) = f(log(k_{it}), log(l_{it})) + \sigma_t \varepsilon_{it} \tag{38}$$

where ε_{it} is an idiosyncratic productivity shock with $\mathbb{E}[\varepsilon_{it}] = 0$ and $var(\varepsilon_{it}) = 1$, cross-sectional uncertainty shocks are driving exogenous shifts in σ_t.

Figure 13: Impulse response functions to cross-sectional uncertainty shocks



Notes: The figure shows the Impulse Response Functions (IRFs) of the macroeconomic aggregates to a cross-sectional uncertainty shock. In red bold line we report the posterior mean estimate from the FunVAR while in dashed red line we report the 15^{th} and 85^{th} credible sets

8 Conclusion

We develop a Functional Augmented Vector Autoregression (FunVAR) model to explicitly incorporate firm-level heterogeneity observed in more than one dimension and study its interaction with aggregate macroeconomic fluctuations. We remark the importance of modeling the joint distribution, rather than just the marginal distributions of the micro observables, for qualitatively assessing the effects of economic shocks on the micro-level distributions and for determining which scenario, potentially aligned with a specific structural model, the data support more. We address the challenge of approximating a multidimensional distribution using data reduction techniques for tensor data objects. We use these methods for approximating multiple cross-sectional distributions of micro-variables, which dynamically interact with the macroeconomic aggregates in a Functional VAR framework. We use the model to study the transmission of aggregate TFP shocks and monetary policy shocks both on the cross-sectional distribution of firm-level capital and labor and the macroeconomic aggregates. Then, we exploit a functional VAR model which links the joint distribution of firms output, capital and labor to identify the effects of cross-sectional uncertainty shocks

on the macroeconomic aggregates. We find that cross-sectional uncertainty shocks—defined as distributional shocks that maximize the variation in the cross-sectional dispersion of firms' output, conditional on the contemporaneous level of firm capital and labor—are associated with declines in investment, output, employment, and consumption consistent with theoretical predictions.

A Appendix

A.1 Factor augmented approximation of the Functional VAR model

Thanks to the finite dimensional approximation (6), and taking $\boldsymbol{s}_K(\boldsymbol{x'})$ as a K-dimensional vector of functional basis such that

$$\int \boldsymbol{s}_K(\boldsymbol{x})\boldsymbol{h}_K(\boldsymbol{x})'d\boldsymbol{x} = \boldsymbol{C_\beta} \,, \tag{39}$$

we can rewrite the function augmented VAR expanding the unknown functions of $\boldsymbol{x}$ as follows:

$$\begin{aligned} c_l(\boldsymbol{x}) &= \boldsymbol{h}_K(\boldsymbol{x})'\tilde{c}_{l,t} \\ \boldsymbol{B}_{s,yl}(\boldsymbol{x}) &= \boldsymbol{B_{s,yl}}\boldsymbol{s}_K(\boldsymbol{x}) \\ \boldsymbol{B}_{s,ly}(\boldsymbol{x}) &= \boldsymbol{h}_K(\boldsymbol{x})'\boldsymbol{B_{s,ly}} \\ B_{s,ll}(\boldsymbol{x},\boldsymbol{x'}) &= \boldsymbol{h}_K(\boldsymbol{x})'\boldsymbol{B_{s,ll}}\boldsymbol{s}_K(\boldsymbol{x'}) \\ \boldsymbol{u}_l(\boldsymbol{x}) &= \boldsymbol{h}_K(\boldsymbol{x})'\tilde{u}_{l,t} \,, \end{aligned} \tag{40}$$

where $\boldsymbol{B_{s,ll}}$ is a $K \times K$ matrix, $\boldsymbol{B_{s,ly}}$ is a $K \times n_y$ matrix, and $\boldsymbol{B_{s,yl}}$ is an $n_y \times K$ matrix. Plugging in (33) and (34) we get:

$$\boldsymbol{y_t} = \boldsymbol{c_y} + \sum_{s=1}^{p} \boldsymbol{B}_{s,yy}\boldsymbol{y_{t-s}} + \sum_{s=1}^{p} \boldsymbol{B_{s,yl}}\boldsymbol{C_\beta}\boldsymbol{\beta}_{t-s;K} + \boldsymbol{u_{y,t}} \tag{41}$$

$$\boldsymbol{h}_K(\boldsymbol{x})'\boldsymbol{\beta}_t = \boldsymbol{h}_K(\boldsymbol{x})'\tilde{c}_{l,t} + \sum_{s=1}^{p} \boldsymbol{h}_K(\boldsymbol{x})'\boldsymbol{B_{s,ly}}\boldsymbol{y}_{t-s} + \boldsymbol{h}_K(\boldsymbol{x})' \sum_{s=1}^{P} \boldsymbol{B_{s,ll}}\boldsymbol{\beta}_{t-s;K} + \boldsymbol{h}_K(\boldsymbol{x})'\tilde{u}_{l,t} \,, \tag{42}$$

which becomes:

$$\boldsymbol{\beta}_t = \tilde{c}_{l,t} + \sum_{s=1}^{p} \boldsymbol{B_{s,ly}}\boldsymbol{y}_{t-s} + \sum_{s=1}^{P} \boldsymbol{B_{s,ll}}\boldsymbol{C_\beta}\boldsymbol{\beta}_{t-s;K} + \tilde{u}_{l,t} \,. \tag{43}$$

A.2 Missing densities and mixed frequency

In applied work, the macroeconomic aggregate variables $\boldsymbol{y}$ and the distribution of firm's micro level characteristics $f(\boldsymbol{x})$ are often sampled at different frequencies. We adapt the standard factor augmented VAR framework, to allow for the missing densities. Our model becomes:

$$l_t(\boldsymbol{x})^{obs} = \begin{cases} \boldsymbol{H}\boldsymbol{\beta_t} + \boldsymbol{\varepsilon_t}, & \text{if } t = m, 2m, 3m, \ldots, Tm \\ \varnothing, & \text{otherwise} \end{cases} \tag{44}$$

$$\begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{\beta}_t \end{bmatrix} = \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \begin{bmatrix} \boldsymbol{y}_{t-1} \\ \boldsymbol{\beta}_{t-1} \end{bmatrix} + \ldots + \boldsymbol{\Phi}_p \begin{bmatrix} \boldsymbol{y}_{t-p} \\ \boldsymbol{\beta}_{t-p} \end{bmatrix} + \begin{bmatrix} \boldsymbol{u}_{y,t} \\ \tilde{u}_{q,t} \end{bmatrix} , \tag{45}$$

where m is the number of high-frequency observations contained within the span of one low-frequency observation period (for example in the case annual and quarterly observations $m = 4$), and the full sample size is T in low frequency and Tm in high frequency. The observation equation (44) assumes that the firm-level variables on labor and capital are reporting their end of the quarter values. This would be for example the case of both output, labor and capital as reported in the Compustat dataset.[10] To estimate the model we proceed again in two steps. First, we estimate the basis functions $\boldsymbol{H}$ using either principal component analysis or multilinear principal component analysis or CP decomposition. In particular, for a sufficient large number of observed low-frequency densities, we can still consistently estimate $\boldsymbol{H}$ by performing principal component analysis on the sample of observed low frequency densities. Then, we estimate the factor augmented VAR model using the Gibbs Sampler described in the previous section, which needs just to be adapted by considering the following observation equation

$$\boldsymbol{l} = \boldsymbol{M}\boldsymbol{b} + \boldsymbol{\epsilon} , \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{TN^{grid}}) , \tag{46}$$

where now $\boldsymbol{M} = \boldsymbol{Q} \otimes \boldsymbol{H}$ where $\boldsymbol{Q} = \boldsymbol{I}_{T/m} \otimes diag(\boldsymbol{0}_{1\times m-1}, 1)$ is a $T \times T$ matrix selecting the factors according to the observation equation (44). The posterior distribution of $\boldsymbol{b}$ is then just adjusted with this new definition of $\boldsymbol{M}$.

A.3 Details on Compustat data

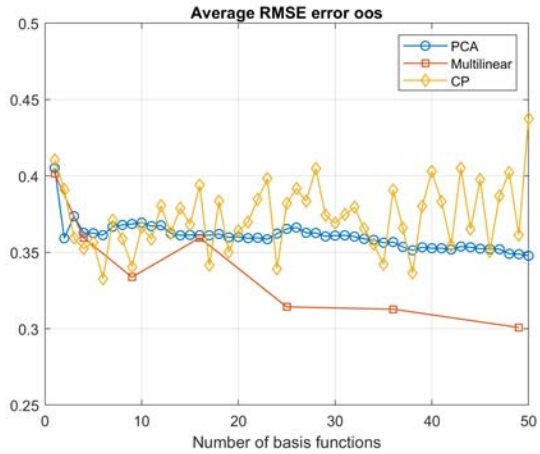
We mostly follow Ottonello et al. (2020b) for what concerns data cleaning and variable transformations. In our analysis, we initialize the assessment of each firm's investment behavior by setting the initial value of capital stock for the next period, $k_{j,t+1}$, based on the reported level of gross plant, property, and equipment from Compustat (`ppegtq`, item 118). We continue by tracking the evolution of this capital stock through changes in net plant, property, and equipment (`ppentq`, item 42), which usually offers more observations and incorporates adjustments for depreciation. When encountering missing data points in `ppentq` between two periods, linear interpolation is employed using the nearest available values to estimate the missing entry. However, we avoid interpolation

10. Accounting for flow variables in this framework would be more complicated, since the observed distribution would most likely be a convolution of the unobserved missing distributions as noted in Marcellino et al. (2024).

when the dataset exhibits two or more consecutive missing observations to maintain accuracy. Our empirical analysis implements several exclusion criteria to ensure data quality and relevance. Initially, we exclude firms operating within specific sectors deemed non-representative of typical corporate investment behaviors. These sectors include finance, insurance, and real estate, which are categorized within Standard Industrial Classification (SIC) codes ranging from 6000 to 6799, as well as utilities with SIC codes from 4900 to 4999. Additionally, non-operating establishments (SIC code 9995) and industrial conglomerates (SIC code 9997) are also omitted from our study. Moreover, our analysis is restricted to firms that are incorporated in the United States, excluding any firm-quarter observations that fail to meet this criterion. Within the retained data, we further refine our sample by excluding observations that display characteristics of extreme financial behavior or data anomalies. This includes observations with negative capital or assets, and those involving significant acquisitions, defined as acquisitions where the acquired assets are greater than 5% of the firm's total assets. We also exclude observations where the investment rate appears anomalously high or low, falling in the top or bottom 0.5% of the distribution, as well as those where the investment spell is shorter than 40 quarters. Additional exclusions apply to quarters with real sales growth exceedingly high above 1 or significantly low below -1. Observations with negative sales or liquidity are similarly omitted to maintain the integrity and reliability of our analysis. We transform the nominal firm level capital into real firm level capital dividing by the implicit price deflator of the nonresidential gross private domestic fixed investment available from FRED (`A008RD3Q086SBEA`). We log-linearly detrend the level of capital at the firm level to concentrate on cyclical fluctuations. For the labor variable we consider `emp` from the Annual dataset and interpolate quarterly observations from annual observations.

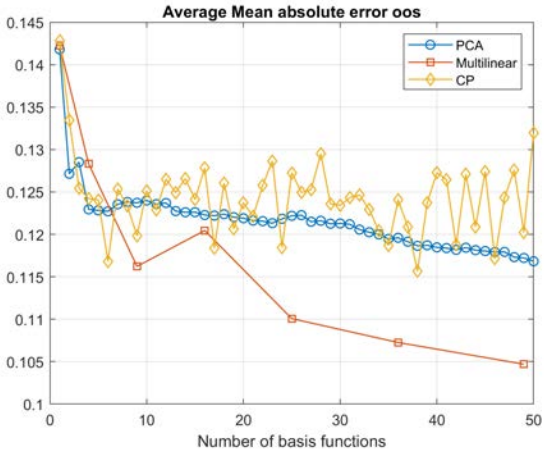
A.4 Additional results from cross-validation

Figure 14: Cross-Validated RMSE on the Test Set: PCA vs. Multilinear PCA vs. CP



Notes: The figure shows the RMSE to the kernel density estimate of the bivariate density function $f(log(k), log(l))$ in the test set in the cross-validation.

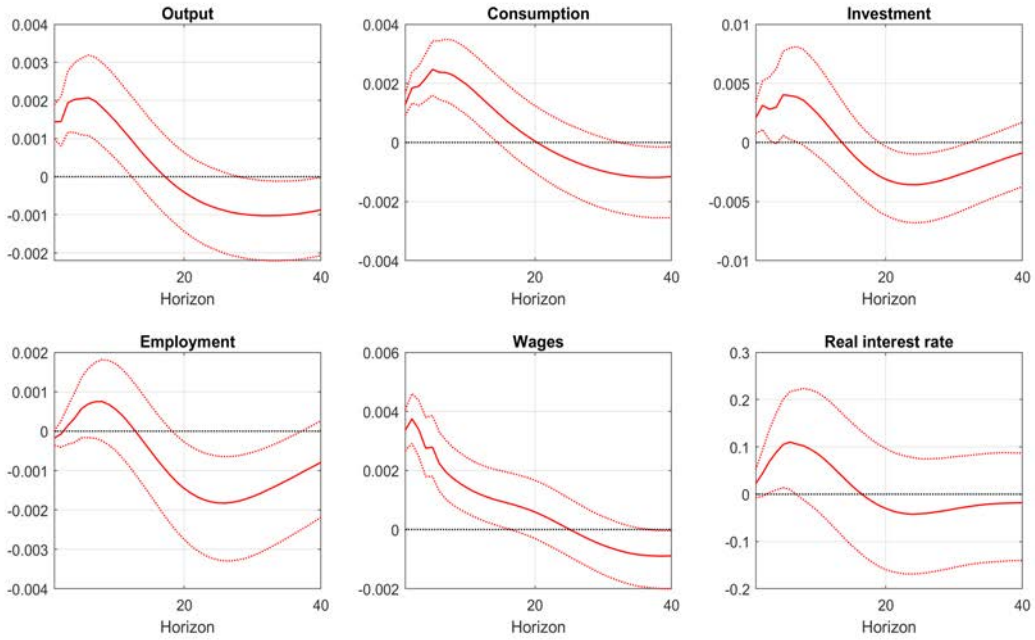
Figure 15: Cross-Validated MAE on the Test Set: PCA vs. Multilinear PCA vs. CP



Notes: The figure shows the MAE distance to the kernel density estimate of the bivariate density function $f(log(k), log(l))$ in the test set in the cross-validation.

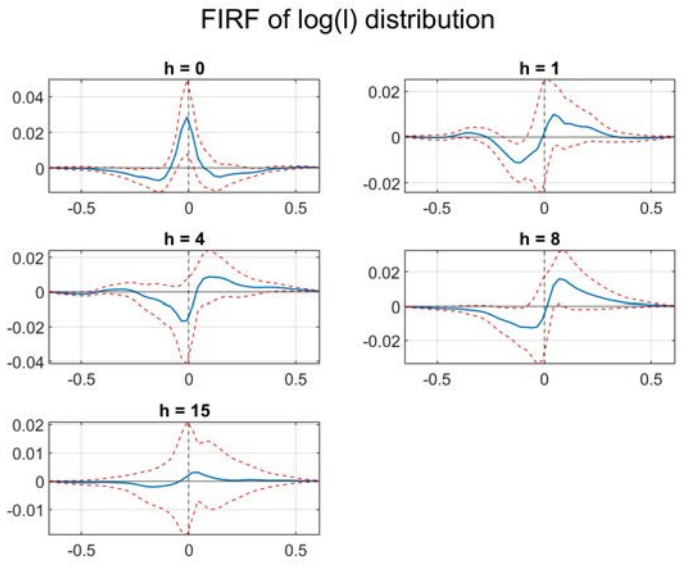
A.5 Additional results: The distributional effects of TFP shocks

Figure 16: Responses to a TFP shock



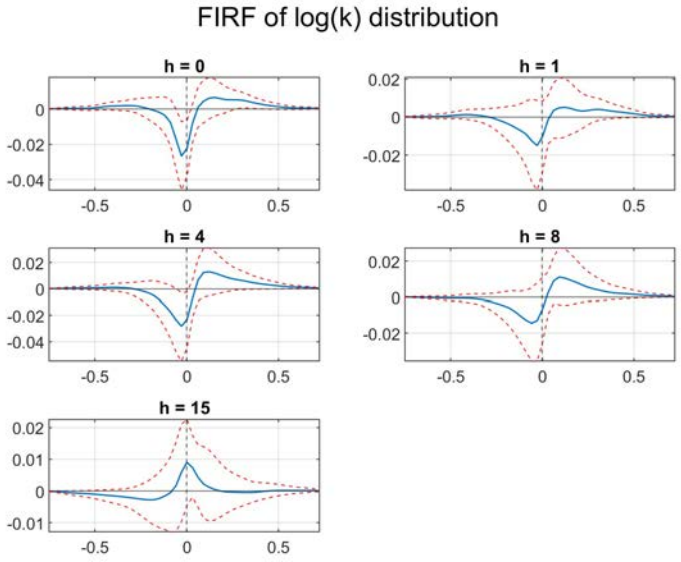
Notes: The figure shows the Impulse Response Functions (IRFs) of the macroeconomic aggregates to a one standard deviation TFP shock. In red bold line we report the posterior mean estimate from the FunVAR while in dashed red line we report the 15^{th} and 85^{th} credible sets.

Figure 17: Functional impulse response functions of the marginal log-labor distribution to the TFP shock



Notes: The figure shows the functional impulse response function of the marginal $log(l)$ distribution following the TFP shock. In plain blue the posterior median estimates, while in dashed red we report the 15^{th} and 85^{th} credible sets.

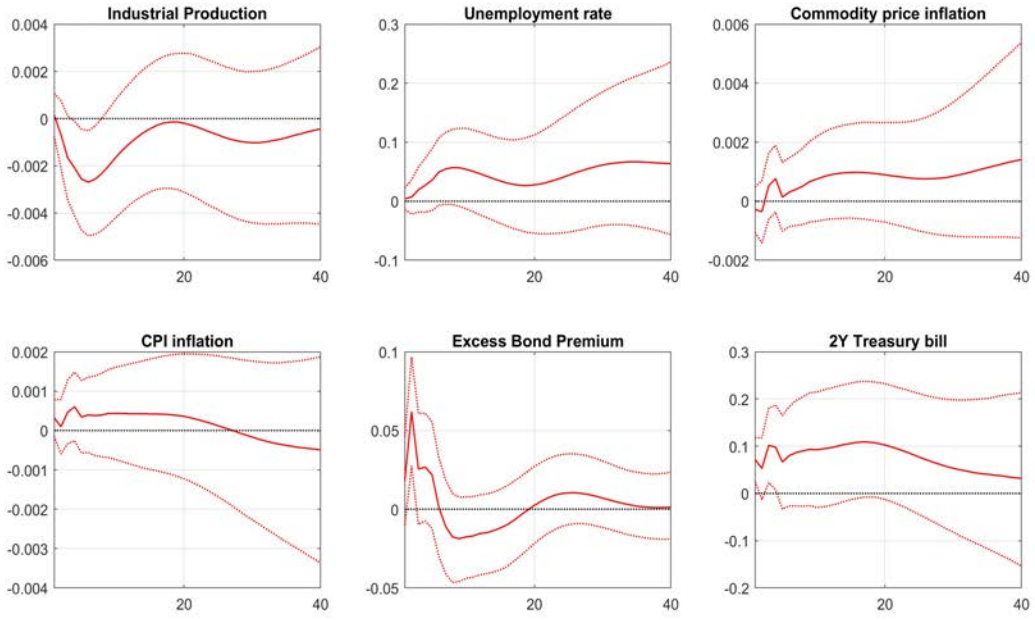
Figure 18: Functional impulse response functions of the marginal log-capital distribution to the TFP shock



Notes: The figure shows the functional impulse response function of the marginal $log(k)$ distribution following the TFP shock. In plain blue the posterior median estimates, while in dashed red we report the 15^{th} and 85^{th} credible sets.

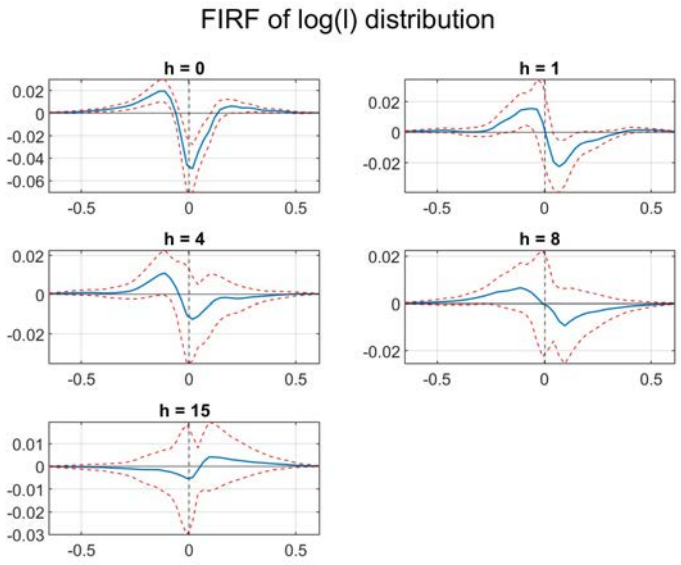
A.6 Additional results Monetary policy shocks

Figure 19: Responses to a monetary policy shock



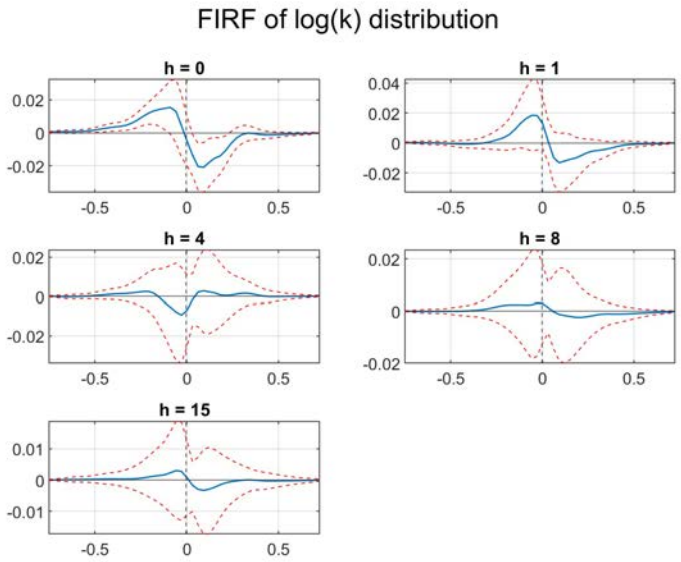
Notes: The figure shows the Impulse Response Functions (IRFs) of the macroeconomic aggregates to a one standard deviation MP shock. In red bold line we report the posterior mean estimate from the FunVAR while in dashed red line we report the 15^{th} and 85^{th} credible sets.

Figure 20: Functional impulse response functions of the marginal log-labor distribution to the monetary policy shock



Notes: The figure shows the functional impulse response function of the marginal $log(l)$ distribution following the monetary policy shock. In plain blue the posterior median estimates, while in dashed red we report the 15^{th} and 85^{th} credible sets.

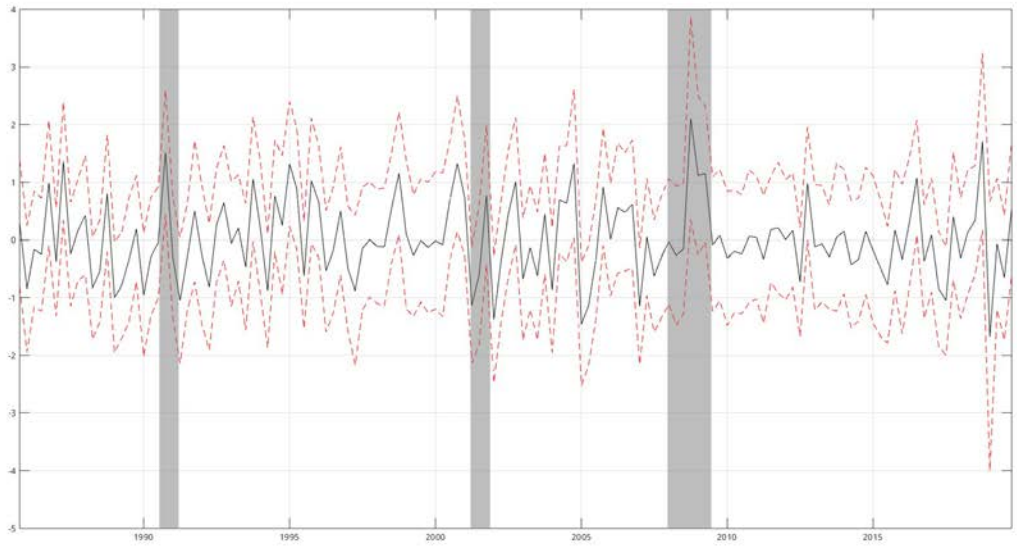
Figure 21: Functional impulse response functions of the marginal log-capital distribution to the monetary policy shock



Notes: The figure shows the functional impulse response function of the marginal $log(k)$ distribution following the monetary policy shock. In plain blue the posterior median estimates, while in dashed red we report the 15^{th} and 85^{th} credible sets.

A.7 Additional results uncertainty shocks

Figure 22: Estimated cross sectional uncertainty shocks in the functional VAR



Notes: The figure shows the time series of the estimated cross sectional uncertainty shocks in the funVAR. In plain black the posterior median estimates, while in dashed red we report the 15^{th} and 85^{th} credible sets. In shaded gray NBER recession periods.

References

Aiyagari, S Rao. 1994. "Uninsured idiosyncratic risk and aggregate saving." *The Quarterly Journal of Economics* 109 (3): 659–684.

Arellano, Cristina, Yan Bai, and Patrick Kehoe. 2019. "Financial Frictions and Fluctuations in Volatility." *Journal of Political Economy* 127 (5): 2049–2103.

Babii, Andrii, Eric Ghysels, and Junsu Pan. 2024. "Tensor PCA for Factor Models."

Bańbura, Marta, Domenico Giannone, and Lucrezia Reichlin. 2010. "Large Bayesian vector auto regressions." *Journal of Applied Econometrics* 25 (1): 71–92.

Bauer, Michael D., and Eric T. Swanson. 2023. "An Alternative Explanation for the "Fed Information Effect"." *American Economic Review* 113, no. 3 (March): 664–700.

Bayer, Christian, Ralph Luetticke, Lien Pham-Dao, and Volker Tjaden. 2019. "Precautionary savings, illiquid assets, and the aggregate consequences of shocks to household income risk." *Econometrica* 87 (1): 255–290.

Beyeler, Simon, and Sylvia Kaufmann. 2018. *Factor augmented VAR revisited: A sparse dynamic factor model approach.* Technical report. Working paper.

Bilbiie, Florin O, Giorgio Primiceri, and Andrea Tambalotti. 2023. *Inequality and business cycles.* Technical report. National Bureau of Economic Research.

Bjørnland, Hilde C, Yoosoon Chang, and Jamie Cross. 2023. "Oil and the stock market revisited: A mixed functional var approach." *Available at SSRN.*

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3): 623–685.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J Terry. 2018. "Really uncertain business cycles." *Econometrica* 86 (3): 1031–1065.

Carriero, A., G. Kapetanios, and M. Marcellino. 2009. "Forecasting exchange rates with a large Bayesian VAR." Forecasting Returns and Risk in Financial Markets using Linear and Nonlinear Models, *International Journal of Forecasting* 25 (2): 400–417.

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino. 2018. "Measuring Uncertainty and Its Impact on the Economy." *The Review of Economics and Statistics* 100, no. 5 (December): 799–815.

Carroll, J. D., and J.-J. Chang. 1970. "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition." *Psychometrika* 35 (3): 283–319.

Carter, C. K., and R. Kohn. 1994. "On Gibbs Sampling for State Space Models." *Biometrika* 81 (3): 541–553.

Chan, Joshua C.C., Aubrey Poon, and Dan Zhu. 2023. "High-dimensional conditionally Gaussian state space models with missing data." *Journal of Econometrics* 236 (1): 105468.

Chan, Joshua CC. 2022. "Asymmetric conjugate priors for large Bayesian VARs." *Quantitative Economics* 13 (3): 1145–1169.

Chang, Minsu, Xiaohong Chen, and Frank Schorfheide. 2021. *Heterogeneity and aggregate fluctuations.* Technical report. National Bureau of Economic Research.

Chang, Minsu, and Frank Schorfheide. 2022. "On the effects of monetary policy shocks on earnings and consumption heterogeneity."

Chang, Yoosoon, Soyoung Kim, and Joon Y Park. 2025. "How Do Macroaggregates and Income Distribution Interact Dynamically? A Novel Structural Mixed Autoregression with Aggregate and Functional Variables CAMA Working Paper 7/2025."

Christiano, Lawrence J., Roberto Motto, and Massimo Rostagno. 2014. "Risk Shocks." *American Economic Review* 104, no. 1 (January): 27–65.

Dew-Becker, Ian, and Stefano Giglio. 2023. "Cross-Sectional Uncertainty and the Business Cycle: Evidence from 40 Years of Options Data." *American Economic Journal: Macroeconomics* 15, no. 2 (April): 65–96.

Doz, Catherine, Domenico Giannone, and Lucrezia Reichlin. 2011. "A two-step estimator for large approximate dynamic factor models based on Kalman filtering." Annals Issue on Forecasting, *Journal of Econometrics* 164 (1): 188–205.

Fernald, John. 2014. "A quarterly, utilization-adjusted series on total factor productivity." Federal Reserve Bank of San Francisco.

Gilchrist, Simon, and Egon Zakrajšek. 2012. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review* 102, no. 4 (June): 1692–1720.

Harshman, Richard A. 1970. "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis." *UCLA Working Papers in Phonetics* 16:1–84.

Hron, K., A. Menafoglio, M. Templ, K. Hrůzová, and P. Filzmoser. 2016. "Simplicial principal component analysis for density functions in Bayes spaces." *Computational Statistics & Data Analysis* 94:330–350.

Huber, Florian, Massimiliano Marcellino, and Tommaso Tornese. 2024. "The Distributional Effects of Economic Uncertainty." Manuscript.

Hung, Hung, Peishein Wu, Iping Tu, and Suyan Huang. 2012. "On multilinear principal component analysis of order-two tensors." *Biometrika* 99 (3): 569–583.

Jurado, Kyle, Sydney C Ludvigson, and Serena Ng. 2015. "Measuring uncertainty." *American Economic Review* 105 (3): 1177–1216.

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

Khan, Aubhik, and Julia K. Thomas. 2008. "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics." *Econometrica* 76 (2): 395–436.

Koby, Yann, and Christian Wolf. 2020. "Aggregation in heterogeneous-firm models: Theory and measurement." *Manuscript, July.*

Kolda, Tamara G., and Brett W. Bader. 2009. "Tensor Decompositions and Applications." *SIAM Review* 51 (3): 455–500.

Kydland, Finn E, and Edward C Prescott. 1982. "Time to build and aggregate fluctuations." *Econometrica: Journal of the Econometric Society* 50 (6): 1345–1370.

Lenza, Michele, and Ettore Savoia. 2024. *Do we need firm data to understand macroeconomic dynamics?* Technical report. Sveriges Riksbank Working Paper Series.

Liu, Laura, and Mikkel Plagborg-Møller. 2023. "Full-information estimation of heterogeneous agent models using macro and micro data." *Quantitative Economics* 14 (1): 1–35.

Marcellino, Massimiliano, Andrea Renzetti, and Tommaso Tornese. 2024. "Nowcasting Distributions: A Functional MIDAS Model." Manuscript draft, *Unpublished Manuscript* (November).

Ottonello, Pablo, and Thomas Winberry. 2020a. "Financial heterogeneity and the investment channel of monetary policy." *Econometrica* 88 (6): 2473–2502.

———. 2020b. "Financial Heterogeneity and the Investment Channel of Monetary Policy." *Econometrica* 88 (6): 2473–2502.

Petersen, Alexander, Chao Zhang, and Piotr Kokoszka. 2022. "Modeling Probability Density Functions as Data Objects." *Econometrics and Statistics* 21 (C): 159–178.

Ramey, V.A. 2016. "Chapter 2 - Macroeconomic Shocks and Their Propagation," edited by John B. Taylor and Harald Uhlig, 2:71–162. Handbook of Macroeconomics. Elsevier.

Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis.* London: Chapman & Hall/CRC.

Tucker, Ledyard R. 1966. "Some mathematical notes on three-mode factor analysis." *Psychometrika* 31 (3): 279–311.

Winberry, Thomas. 2018. "A method for solving and estimating heterogeneous agent macro models." *Quantitative Economics* 9 (3): 1123–1151.

———. 2021. "Lumpy investment, business cycles, and stimulus policy." *American Economic Review* 111 (1): 364–396.

Ye, Jieping. 2004. "Generalized low rank approximations of matrices." In *Proceedings of the twenty-first international conference on Machine learning,* 112.