Department of Economics and Finance

Economics and Finance Working Paper Series

Working Paper No. 2403

Theo Drossidis, Haroon Mumtaz and Angeliki Theophilopoulou

The Distributional Effects of Oil Supply News shocks

February 2024

http://www.brunel.ac.uk/economics

The Distributional Effects of Oil Supply News shocks*

Theo Drossidis[1], Haroon Mumtaz[2], and Angeliki Theophilopoulou[3]

[2]Queen Mary University of London
[1]Brunel University
[3]Brunel University

This version: February 19, 2024.

Abstract

This paper uses high frequency data on the distribution of US income to investigate the heterogeneous effects of oil supply news shocks. Using a FAVAR with an external instrument, We show that these shocks have large negative effects on the left and right tail of the distribution. For low income individuals, the effect is driven by a decline in wages and proprietor's income, while a fall in corporate profits and interest income drives the effect for affluent individuals.

JEL Classification: C32, E32, Q54
Keywords: Oil shock; income inequality; FAVAR; External instrument identification.

*

1 Introduction

The recent energy crisis has again focused attention on oil prices. A large empirical literature has established the importance of oil shocks for economic fluctuations. In a recent important contribution, Känzig (2021) uses a narrative approach to identify oil supply news shocks, i.e. unexpected fluctuations in current and future oil supply, and shows that these disturbances have a sizeable effect on US industrial production and CPI inflation. By applying this identification approach, Känzig (2021) builds on a large literature that reaches similar conclusions for oil market disturbances (see for e.g. Hamilton (2003), Baumeister and Kilian (2016), Caldara *et al.* (2019)).

One common feature of this literature is the focus on aggregate macroeconomic outcomes. In this paper, we exploit high-frequency data on the distribution of income and its components for the US to investigate the distributional effects of oil supply news shocks. We use a factor augmented VAR (FAVAR) to jointly model the oil market, macroeconomic variables and income in deciles of distribution. The oil supply shock is identified using the external instrument approach of Känzig (2021). The analysis leads to three key findings:

1. While an adverse oil supply shock has the largest effect at the left tail of the income distribution, the income of affluent individuals also declines relative to the median.
2. For individuals at the left tail, the decline in income is driven by a sharp fall in wages and proprietor's income.
3. At the right tail, income declines as the shock pushes down components of capital income such as interest income and corporate profits.

Our paper is related to Berisha *et al.* (2021) who examine the impact of oil production and dependency on the annual Gini coefficient for US states in a reduced-form setting. Our analysis is an extension of Berisha *et al.* (2021), as we identify an oil supply shock taking into account the effect of news and examine how the distribution and components of income are affected rather than focusing one one measure of inequality.
The paper is organised as follows: The data and empirical model is described in Section 2 while Section 3 describes the main results.

2 Empirical model and data

To estimate the impact of oil supply news on income for different groups of the population, we use a factor-augmented VAR (FAVAR) model. The model is defined by the VAR:

$$Y_t = c + \sum_{j=1}^{P} \beta_j Y_{t-j} + u_t \tag{1}$$

where $Y_t = \begin{pmatrix} z_t \\ \hat{F}_t \end{pmatrix}$, where z_t denotes a set of variables pertaining to the oil market: the real price of oil, world oil production and world oil inventories. $\hat{F}_t$ represents factors that summarize information in a panel of macroeconomic and financial series and the individual-level data on income and its components, described below. The factors are estimated using the non-stationary factor model of Barigozzi *et al.* (2021). Denote X_t as the $(M \times 1)$ data matrix that contains the panel of macroeconomic and financial series that summarize information about the economy, and also includes income data at the dis-aggregated level. The observation equation of the FAVAR is defined as:

$$X_t = c + b\tau + \Lambda F_t + \xi_t \tag{2}$$

where c is an intercept, τ denotes a time-trend, F_t are the R non-stationary factors, Λ is a $M \times R$ matrix of factor loadings, and ξ_t are idiosyncratic components that are allowed to be $I(1)$ or $I(0)$. Note that the idiosyncratic components corresponding to the disaggregated income data can be interpreted as shocks that are specific to those groups and also capture possible measurement errors. The shocks to equation (1) represent macroeconomic or common shocks. It is the response to these common shocks that is relevant to our investigation. This ability to estimate the effect of macroeconomic shocks while taking into account idiosyncratic errors via equation 2 is a key advantage of the FAVAR over a VAR, where these two sources of fluctuations are harder to separate (see De Giorgi and Gambetti (2017) and Cantore *et al.* (2023)). Moreover, by incorporating a large data set, the FAVAR reduces the problem of information deficiency (see e.g. Forni and Gambetti (2014)) and shock deformation (see e.g. Canova and Ferroni (2022)).

2.1 Identification of the Oil supply news shock

To identify the oil supply news shock, we use an external instrument approach (see e.g. Stock and Watson (2008) and Mertens and Ravn (2013)). The residuals u_t are related to structural shocks ε_t via:

$$u_t = A_0 \varepsilon_t \tag{3}$$

where $cov\,(u_t) = \Sigma = A_0 A_0'$. We denote the shock of interest as ε_{1t} and the remaining disturbances as ε_{-t}. Identification of ε_{1t} is based on the instrument m_t that satisfies the relevance and exogeneity conditions: $cov\,(m_t, \varepsilon_{1t}) = \alpha \neq 0$ and $cov\,(m_t, \varepsilon_{-t}) = 0$. As discussed in the technical appendix, these conditions can be combined with the covariance restrictions to obtain an estimate of the relevant column of the contemporaneous impact matrix A_0. In our benchmark model, we employ the instrument constructed by Känzig (2021) which is based on variation in oil futures prices around OPEC announcements. Känzig (2021) provides evidence to suggest that the instrument is relevant and exogenous.

2.2 Data and Estimation

As noted above, X includes both aggregate and individual-level data. The aggregate data is taken from the Fred-MD database. This consists of 134 variables covering industrial production, employment, consumer prices, asset prices, interest rates, exchange rates and spreads[1].

The data on individual level income is obtained from the Real Time Inequality database constructed by Blanchet *et al.* (2022). Blanchet *et al.* (2022) construct *monthly* distributions of income, wealth and their components by statistically matching the annual distributional national accounts of Piketty *et al.* (2017) with the current population survey and the survey of consumer finances in order to incorporate demographic information. They then construct monthly variables by re-scaling each component of income and using information on the distribution of wages from monthly and quarterly survey and administrative data. We use *factor income* as our benchmark income measure. Factor income is the sum of labour and capital income.[2]

We define 10 groups based on the deciles of factor income: $P_1, P_2, \ldots, P_{10}$. P_1 includes individuals that fall below the tenth percentile of factor income, P_2 denotes individuals above the tenth percentile but below the twentieth percentile and so on. We construct average factor income, capital and labour income in each of these groups. In addition, we calculate the average of the main components of capital and labour income in each group. All of these income variables are deflated by the national income deflator and included in

[1] A full list of these variables is available on FRED-MD website.

[2] As in Blanchet *et al.* (2022) labour income is defined as the sum of wages and 0.7 times proprietors income. Capital income is the sum of 0.3 times proprietors income, corporate profits, interest income, rental income net of corporate taxes and non-mortgage interest payments.

X. The sample ranges from 1976M1 to 2017M12.[3]

The number of factors in the FAVAR model is chosen via the information criteria of Bai and Ng (2002). This procedure suggests the presence of 15 factors. The lag length is set at 12.[4] The parameters of the VAR model in (1) are estimated using a Bayesian approach. We use a Markov chain Monte-Carlo algorithm to approximate the posterior distributions.[5] We employ $11,000$ iterations, retaining every 10^{th} draw after a burn-in period of 1000.

3 Empirical results

Before turning to the effect of the oil supply news shock on the distribution of income, we show the response of selected aggregate variables to this shock in Figure 1. These results broadly support the conclusions reached by Känzig (2021). A 10% increase in the oil price leads to an increase in oil inventories and the median response of oil production is negative at medium horizons, albeit with large error bands. The shock depresses both global and US industrial production and leads to increase in the US unemployment rate and CPI. The shock has a limited effect on short-term interest rates but affects financial conditions adversely, with the BAA spread increasing and the stock market index declining.

3.1 Impact on the distribution of income

Figure 2 shows our main result. The left panels of the figure show the median response of total income, labour income and capital income, averaged in each group defined by deciles of total income. The right panels present the response of these variables in each decile group, along with 90% error bands at the 2 year horizon. The top row of the figure shows that the oil supply shock has the largest effect on income of individuals on the left tail–for the first decile, total income declines by 1% at the 2 year horizon. The impact is smaller towards the center of the distribution with the income of individuals in groups P_6 and P_7 falling by less than 0.5 %. However, for the top 10 %, the effect of the shock appears to be relatively larger. The second row of the figure shows that the impact on the left tail is driven by the large negative reaction of labour income. In contrast, capital income, that constitutes a larger proportion of income at the right tail, barely reacts significantly below

[3] As discussed in Känzig (2021), the instrument is only available from 1984M4 and the estimation of the A_0 matrix uses this sample.

[4] Our main results are robust to the number of factors and lags

[5] The prior and posterior distributions for the VAR parameters are standard and described in the appendix.

the median at the 2 year horizon. For high income individuals, capital income declines substantially driving the larger reaction of total income observed for this group. Figure 3 shows the reaction of some of the main components of labour and capital income to the oil shock and suggests two key conclusions.[6] First, the shock leads to a decline in labour income at the left tail as both wages and proprietor's income declines. Second, capital income is adversely affected at the right tail – the shock is associated with a fall in interest income for groups P_7 to P_{10} and corporate profits for the top decile, possibly as a result of the rise in corporate spreads and fall in interest rates.

3.2 Robustness

We carry out a number of robustness checks that are presented in detail in the technical appendix. A summary is as follows:

1. Identification: As discussed in Känzig (2021), the oil supply news shock can also be identified under weaker assumptions: i.e. allowing for the possibility that other disturbances occur at the same time as the news shock. One method to accomplish this is identification via heteroscedasticity, which only requires the assumption that the variance of oil supply news shocks increases around OPEC announcements while the variance of other shocks remains unchanged. We show in the technical appendix that we obtain very similar results to benchmark when the oil shock is identified. In particular, the shock has the largest effect on income at the left tail.

2. Specification and model: The results are also preserved for FAVAR models with alternative lag lengths and number of factors. As a further check we estimate the VAR model of Känzig (2021) adding the deciles of income measures one by one to the original set of endogenous variables. As shown in the appendix the results from this model support the benchmark conclusions regarding the distributional effects of the shock.

4 Conclusions

This paper shows that adverse oil supply news shocks have a heterogeneous effect on the US income distribution. While the impact of the shock is largest at the left tail of the distribution, more affluent individuals are also significantly affected by the shock. An

[6]Note that interest income is defined as income from currency bonds and deposits

examination of the components of income suggests that these results are driven by a decline in the labour income of the former group and capital income of the latter.

References

Bai, Jushan and Serena Ng, 2002, Determining the Number of Factors in Approximate Factor Models, *Econometrica* **70**(1), 191–221.

Barigozzi, Matteo, Marco Lippi and Matteo Luciani, 2021, Large-dimensional Dynamic Factor Models: Estimation of Impulse-Response Functions with I(1) cointegrated factors, *Journal of Econometrics* **221**(2), 455–482.

Baumeister, Christiane and Lutz Kilian, 2016, Forty years of oil price fluctuations: Why the price of oil may still surprise us, *Journal of Economic Perspectives* **30**(1), 139–60.

Berisha, Edmond, Carolyn Chisadza, Matthew Clance and Rangan Gupta, 2021, Income inequality and oil resources: Panel evidence from the United States, *Energy Policy* **159**, 112603.

Blanchet, Thomas, Emmanuel Saez and Gabriel Zucman, 2022, Real-Time Inequality, *NBER Working Papers 30229*, National Bureau of Economic Research, Inc.

Caldara, Dario, Michele Cavallo and Matteo Iacoviello, 2019, Oil price elasticities and oil price fluctuations, *Journal of Monetary Economics* **103**, 1–20.

Canova, Fabio and Filippo Ferroni, 2022, Mind the gap! Stylized dynamic facts and structural models, *American Economic Journal: Macroeconomics* **forthcoming**.

Cantore, Cristiano, Filippo Ferroni, Haroon Mumtaz and Angeliki Theophilopoulou, 2023, A tail of labor supply and a tale of monetary policy, *Discussion Papers 2308*, Centre for Macroeconomics (CFM).

De Giorgi, Giacomo and Luca Gambetti, 2017, Business cycle fluctuations and the distribution of consumption, *Review of Economic Dynamics* **23**, 19 − 41.

Forni, Mario and Luca Gambetti, 2014, Sufficient information in structural VARs, *Journal of Monetary Economics* **66**(C), 124–136.

Hamilton, James D., 2003, What is an oil shock?, *Journal of Econometrics* **113**(2), 363–398.

Känzig, Diego R, 2021, The macroeconomic effects of oil supply news: Evidence from OPEC announcements, *American Economic Review* **111**(4), 1092–1125.

Mertens, Karel and Morten O. Ravn, 2013, The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States, *American Economic Review* **103**(4), 1212–47.

Piketty, Thomas, Emmanuel Saez and Gabriel Zucman, 2017, Distributional National Accounts: Methods and Estimates for the United States*, *The Quarterly Journal of Economics* **133**(2), 553–609.

Stock, James H. and Mark W. Watson, 2008, What's New in Econometrics - Time Series, *Lecture 7*, National Bureau of Economic Research, Inc.

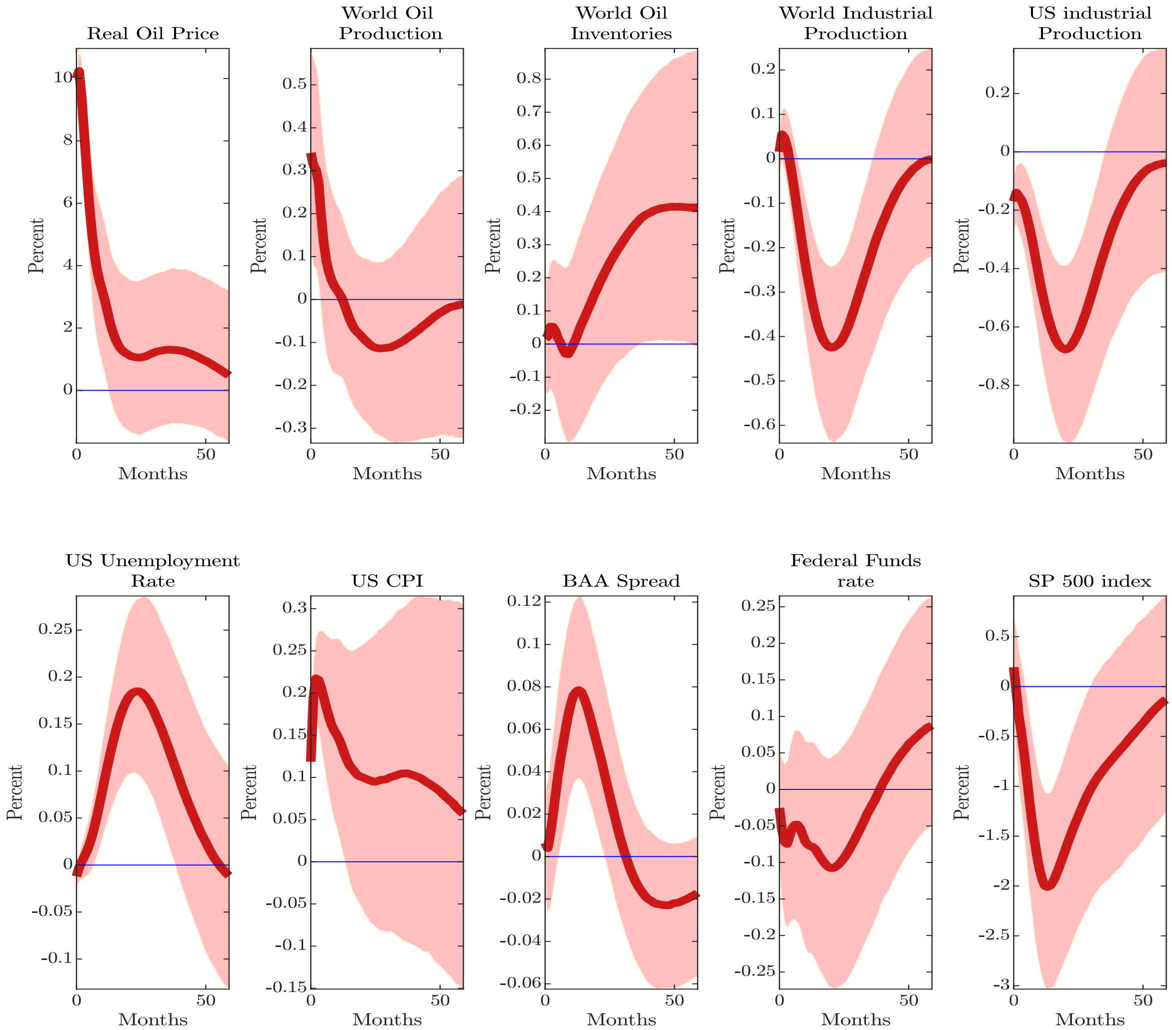

Figure 1: Impulse response functions of selected variables to an oil supply news shock. The shock is normalised to increase the oil price by 10%. The solid lines are the medians while the shaded area represents the 90% error band

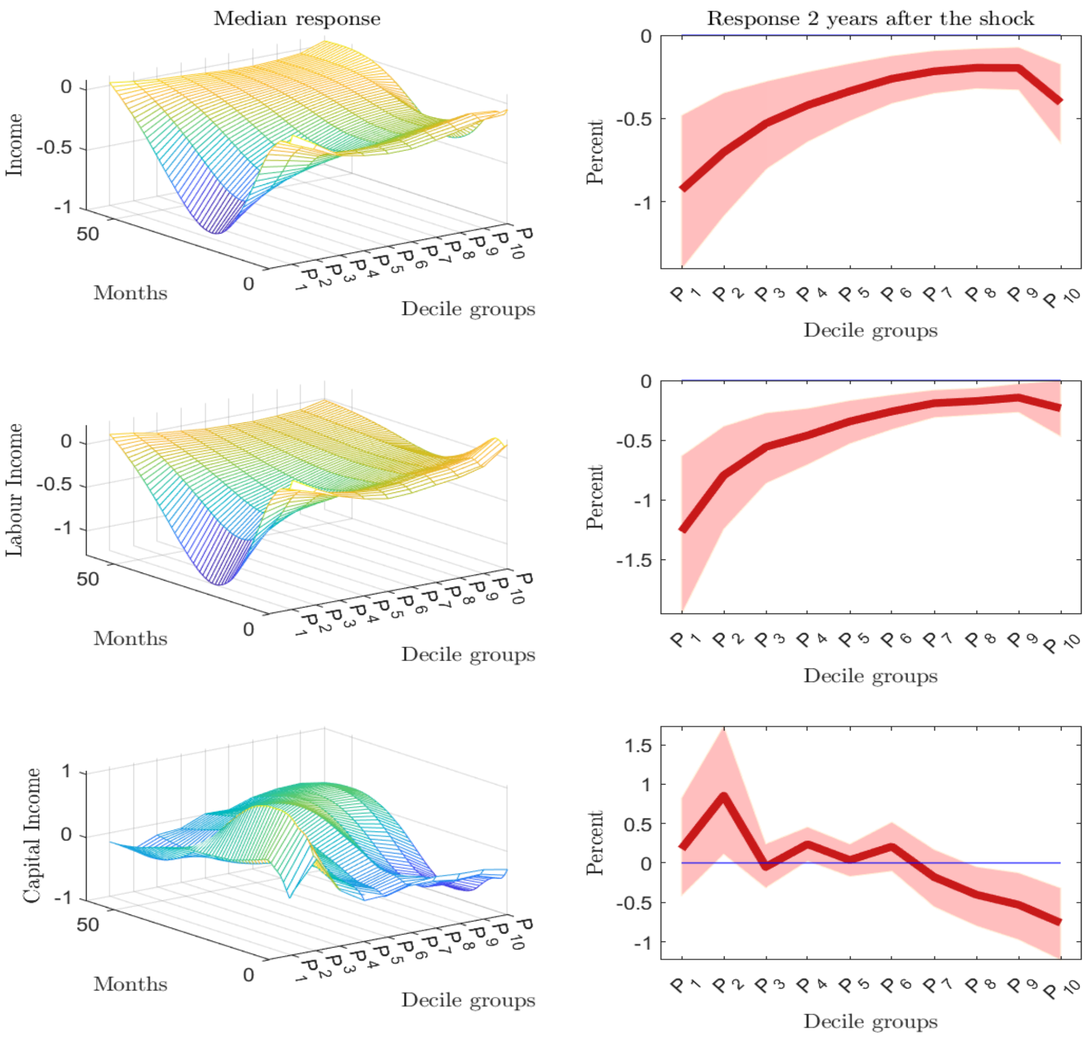

Figure 2: Impulse response functions of total income, labour income and capital income. The shock is normalised to increase the oil price by 10%. The first column shows the median response. The right panel shows the response at the 2 year horizon. The solid lines are the medians while the shaded area represents the 90% error band. $P_1, P_2, \ldots, P_{10}$ denotes the decile groups.

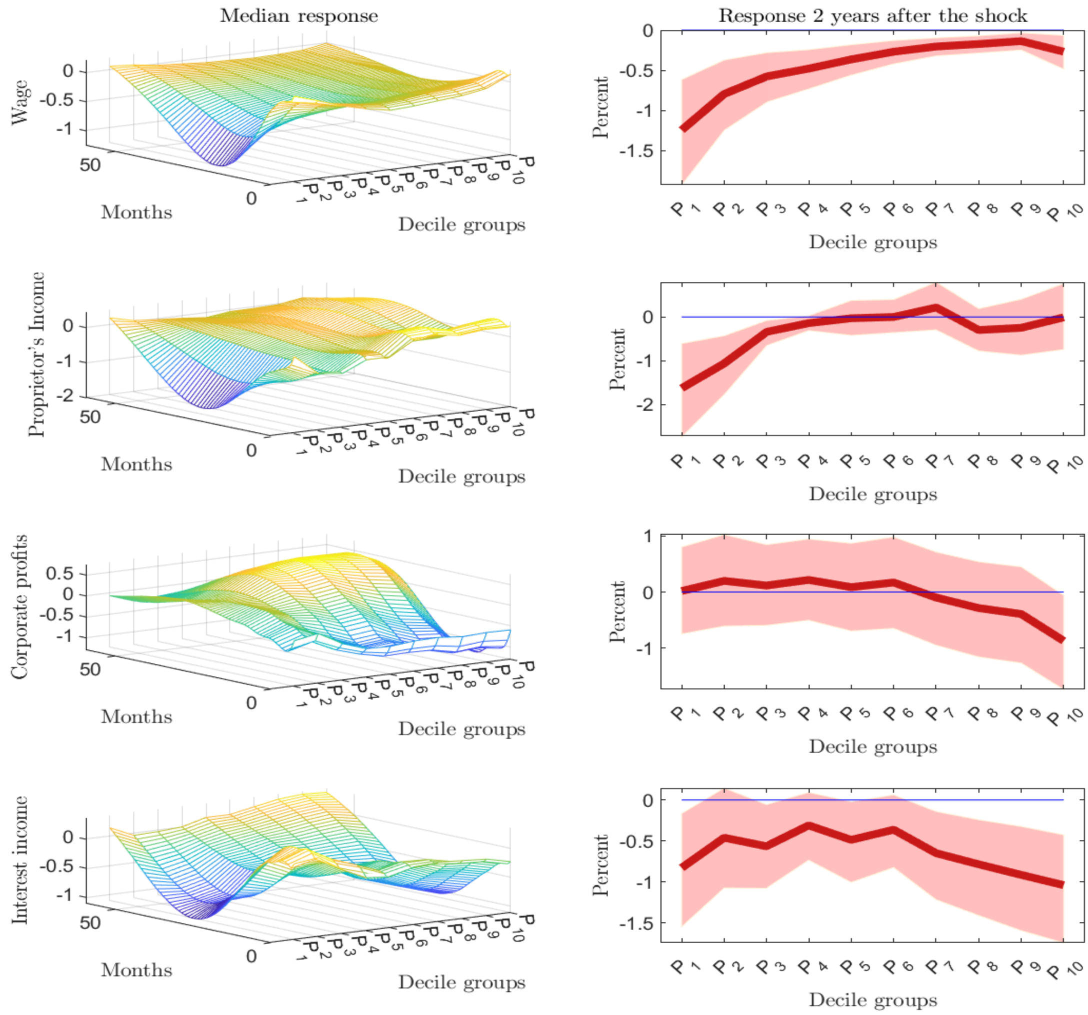

Figure 3: Impulse response functions of the components of labour and capital income The shock is normalised to increase the oil price by 10%. The first column shows the median response. The right panel shows the response at the 2 year horizon. The solid lines are the medians while the shaded area represents the 90% error band. $P_1, P_2, \ldots, P_{10}$ denotes the decile groups.

The Distributional Effects of Oil Supply News shocks. Appendix

Theo Drossidis[1], Haroon Mumtaz[2], and Angeliki Theophilopoulou[3]

[2]Queen Mary University of London

[1]Brunel University

[3]Brunel University

This version: February 19, 2024.

1 Model estimation

The FAVAR model is defined by the following equations:

$$X_{it} = c_i + b_i\tau + \Lambda_i F_t + \xi_{it} \tag{1}$$

$$Y_t = c + \sum_{j=1}^{P} \beta_j Y_{t-j} + u_t \tag{2}$$

$$cov\left(u_t\right) = \Sigma = A_0 A_0' \tag{3}$$

Where $Y_t = \underbrace{\begin{pmatrix} R_t \\ F_t \end{pmatrix}}_{N\times 1}$, R_t denotes the 1 year interest rate, $i = 1, 2, .., M$ denotes the cross-sectional dimension of the panel data-set X_{it} while $t = 1, 2, .., T$ is the dimension. As described in Barigozzi *et al.* (2021), the factors can be consistently estimated using a principal components (PC) estimator. In particular, the factor loadings are estimated via PC analysis of the first differenced data ΔX_{it}. With these in hand, the factors are estimated as $\hat{F}_t = \frac{1}{M}\left(\hat{\Lambda}'\tilde{X}_t\right)$. Here,$\Lambda$ is the matrix of factor loadings, $\tilde{X}_t$ is given by $(x_{1t}, x_{2t}, ..., x_{Mt})$ where $x_{it} = X_{it} - \hat{c}_i - \hat{b}_i\tau$ Note that Barigozzi *et al.* (2021) describe a procedure to check if the *ith* series contains a linear trend and that $\hat{b}_i$ is different from zero.

Given the estimated factors, the VAR in equations 2 is estimated using a Bayesian methods.

1.1 Priors

Denote the var coefficients as $B = vec([\beta_1, \beta_2, .., \beta_P, c])$. We follow Banbura *et al.* (2007) and use a Natural Conjugate prior implemented via dummy observations. The priors are implemented by the dummy observations y_D and x_D that are defined as:

$$y_D = \begin{bmatrix} \frac{diag(\gamma_1 s_1 ... \gamma_n s_n)}{\kappa} \\ 0_{N\times(P-1)\times N} \\ diag(s_1 ... s_n) \\ .... \\ 0_{EX\times N} \end{bmatrix}, \qquad x_D = \begin{bmatrix} \frac{J_P \otimes diag(s_1 ... s_n)}{\kappa} & 0_{NP\times EX} \\ \multicolumn{2}{c}{..............} \\ \multicolumn{2}{c}{0_{N\times(NP)+EX}} \\ \multicolumn{2}{c}{..............} \\ 0_{EX\times NP} & I_{EX}\times 1/c \end{bmatrix} \tag{4}$$

where $J_P = diag(1, 2, ..., P)$, γ_1 to γ_n denote the prior mean for the parameters on the first lag obtained by estimating individual AR(1) regressions, s_1 to s_n is an estimate of the variance of the endogenous variables obtained individual AR(1) regressions, κ measures the tightness of the prior on the autoregressive VAR coefficients, and c is the tightness of the prior on the remaining regressors. We set $\kappa = 0.2$ and $c = 1000$. We also implement priors on the sum of coefficients (see Banbura *et al.* (2007)). The dummy observations for this prior are defined as:

$$\tilde{y}_D = \frac{diag(\gamma_1\mu_1 ... \gamma_n\mu_n)}{\tau}, \tilde{x}_D = \begin{pmatrix} (1_{1\times P}) \otimes \frac{diag(\gamma_1\mu_1 ... \gamma_n\mu_n)}{\tau} & 0_{N\times EX} \end{pmatrix} \tag{5}$$

where μ_i is the sample average of the *ith* variable. As in Banbura *et al.* (2007) we set $\tau = 10\kappa$. The total number of dummy observations is T_D.

1.2 MCMC algorithm

Banbura *et al.* (2007) show that posterior distribution can be written as:

$$g(\Sigma|Y) \sim iW\left(\bar{\Sigma}, T_D + 2 + T - K\right) \tag{6}$$

$$g(B|\Sigma, Y) \sim N\left(\bar{B}, \Sigma \otimes (X_*' X_*)^{-1}\right) \tag{7}$$

where iW denotes the inverse Wishart distribution, K denotes the number of regressors in each equation of the VAR model. Note that $Y_* = \begin{pmatrix} Y \\ y_D \\ \tilde{y}_D \end{pmatrix}$ and $X_* = \begin{pmatrix} X \\ x_D \\ \tilde{x}_D \end{pmatrix}$, X collects the regressors, and

$$\begin{aligned} \tilde{B} &= \left(X_*'X_*\right)^{-1}\left(X_*'Y_*\right) \\ \bar{B} &= vec\left(\tilde{B}\right) \\ \bar{\Sigma} &= \left(Y_* - X_*\tilde{B}\right)'\left(Y_* - X_*\tilde{B}\right) \end{aligned}$$

Posterior draws can be easily generated by drawing Σ from the marginal distribution in 6 and then b from the conditional distribution in equation 7. We set the number of draws to 10,000 with a burn-in of 1,000.

2 IV Identification

For a given draw of B, Σ and u_t, we obtain the first column of $A_0$by using the procedure proposed by Mertens and Ravn (2013). We assume that the instrument is relevant and exogenous:

$$\begin{aligned} cov\left(m_t, \varepsilon_{1t}\right) &= \alpha \\ cov\left(m_t, \varepsilon_t^-\right) &= 0 \end{aligned}$$

where ε_{1t} denotes the structural shock of interest that is ordered first for convenience, while ε_t^-represent all remaining shocks and $\varepsilon_t = \begin{pmatrix} \varepsilon_{1t} & \varepsilon_t^- \end{pmatrix}$. Re-writing the relevance and exogeneity conditions in vector form:

$$E(m_t\varepsilon_t') = [\alpha \quad 0] \tag{8}$$

$$E(m_t\varepsilon_t'A_0') = [\alpha \quad 0]A_0' \tag{9}$$

$$E(m_t u_t') = \alpha a_0 \tag{10}$$

where a_0 is a $(1 \times R)$ vector corresponding to the first row of A_0' (hence first column of A_0). An estimate of $E(m_t u_t') = \begin{pmatrix} E(m_t u_{1t}') \\ E(m_t u_{2t}') \\ . \\ . \\ E(m_t u_{Nt}') \end{pmatrix}'$ can be easily obtained by using a linear regression. However, α on the RHS of equation 10 is unknown. This parameter can be eliminated by normalising the left and the right hand side by dividing by the first element of $E(m_t u_t')$ and a_0, respectively. Therefore the impulse vector to a unit shock is given by

$$\tilde{a}_0 = \begin{pmatrix} 1 \\ \frac{E(m_t u_{2t}')}{E(m_t u_{1t}')} \\ . \\ . \\ \frac{E(m_t u_{Nt}')}{E(m_t u_{1t}')} \end{pmatrix}'.$$

3 Robustness

3.1 Identification by Heteroscedasticity

Following Känzig (2021), we also consider identifying the oil supply news shock using an approach based on heteroscedasticity. The shock is identified by assuming that the variance of the oil supply news shock increases on OPEC announcement days, while the variance of all other disturbances remains the same. As described in Känzig (2021), this assumption can be used to estimate the impact vector. we use this identification as an alternative to the instrument-based approach in the FAVAR. As shown in Figure 5, the impulse responses of the distribution of income and components are similar to the benchmark case.

3.2 Model

We conduct further tests to ensure the reliability of our results. We provide insights from a proxy-BVAR model, similar to the one used by Känzig (2021). Here, we employed Bayesian methods to estimate the model. The variables included in our estimation are the same as those used by Känzig (2021). The shock is normalised to produce a 10% increase in the oil price. We set the lag length to 12, calculate 11000 iterations and save every second draw after burning the first 1000. The findings confirm the asymmetric effects along the distribution presented in the baseline estimation.

3.3 Specification

In this section, we present insights obtained by modifying the baseline model in different ways. Specifically, we reduced the number of factors to 10 and the lag length of the FAVAR to 6. The results of these exercises are presented in Figures 3 and 4, respectively. We find that the outcomes are consistent with previous findings, and altering the number of lags or factors does not significantly impact the results.

4 Data

We use Household data from the real-time inequality database. Following Blanchet *et al.* (2022) this database creates monthly income distributions, available within a few hours after the publication of official high-frequency national accounts aggregates. It uses only publicly available data sources and combines monthly and quarterly survey data with corresponding monthly and quarterly national account statistics in a unified framework. Income aggregates such as factor, pretax, disposable and posttax income as well as their components are derived at a monthly frequency. Hereby, the authors use monthly and quarterly national accounts provided by the Bureau of Economic Analysis. In the next step, the authors make use of the annual data provided by Piketty *et al.* (2018), who combine IRS tax microdata, surveys and national accounts data and construct annual distributions of income and wealth that are consistent with the national accounts aggregates. Converting annual data into monthly data involves two main challenges. First, the existing annual income distributions are adjusted to match monthly totals, which is done through normalisation. However, as pointed out by the authors, income components may change differently every month, leading to redistribution effects. To account for these high-frequency changes in the income components, the authors need to update the files monthly to ensure a representative distribution. The second step involves incorporating these changes into the dataset by adjusting the month-to-month evolution of income components. Our study uses the provided monthly files from January 1976 to December 2017 to derive income data for each decile of the distribution.

Figure 5 shows how each income component contributes to the total income of each decile group over time[1]. We use factor income (code: princ) as our measure of total gross income, which in turn adds up to national income. The observations and distribution are scaled on the household level. Total income is defined as the sum of labour income

[1]Since some income components take negative values for some months, we were forced to eliminate these and only focused on the main components that define total income to derive the figure.

and capital income. To calculate labour income, we follow Blanchet *et al.* (2022) and add the compensation of employees (code: flemp) and 70% of proprietors' income (code: proprietors). For capital income, we add 30% of proprietors' income with corporate profits (code: profits) and interest income (code: fkfix). The latter captures the income from currency, deposits, and bonds. Moving up the distribution, it becomes apparent that the percentage of capital income increases. Corporate profits and interest income are the main contributors to capital income, which is consistent with the explanations of Blanchet *et al.* (2022), who created these datasets.

References

Banbura, Martha, Domenico Giannone and Lucrezia Reichlin, 2007, Bayesian VARs with large panels, *CEPR Discussion Papers* **6326**.

Barigozzi, Matteo, Marco Lippi and Matteo Luciani, 2021, Large-dimensional Dynamic Factor Models: Estimation of Impulse-Response Functions with I(1) cointegrated factors, *Journal of Econometrics* **221**(2), 455–482.

Blanchet, Thomas, Emmanuel Saez and Gabriel Zucman, 2022, Real-Time Inequality, *NBER Working Papers 30229*, National Bureau of Economic Research, Inc.

Känzig, Diego R, 2021, The macroeconomic effects of oil supply news: Evidence from OPEC announcements, *American Economic Review* **111**(4), 1092–1125.

Mertens, Karel and Morten O. Ravn, 2013, The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States, *American Economic Review* **103**(4), 1212–47.

Piketty, Thomas, Emmanuel Saez and Gabriel Zucman, 2018, Distributional National Accounts: Methods and Estimates for the United States, *The Quarterly Journal of Economics* **133**(2), 553–609.

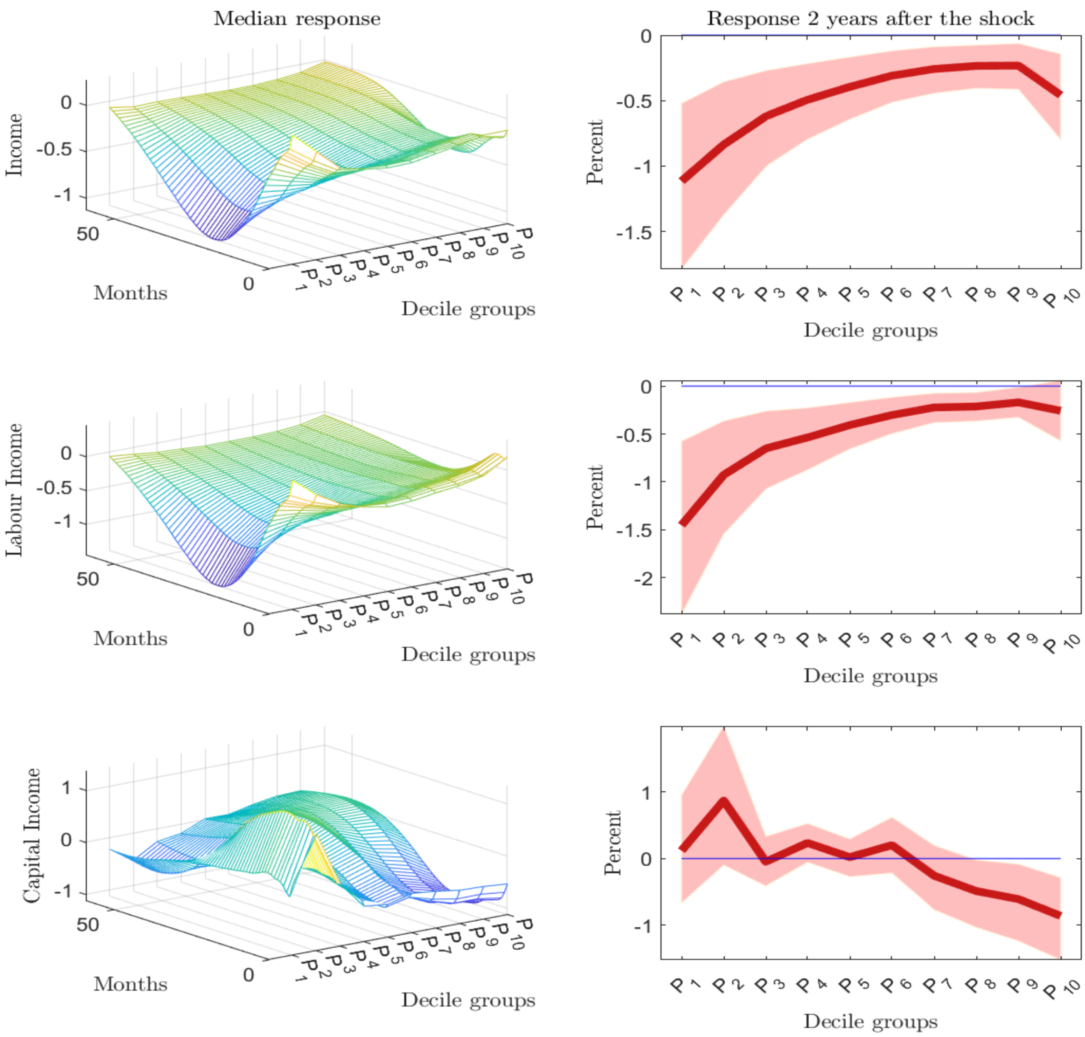

Figure 1: Impulse response functions of total income, labour income and capital income. Shock identified using heteroscedasticity. The shock is normalised to increase the oil price by 10%. The first column shows the median response. The right panel shows the response at the 2 year horizon. The solid lines are the medians while the shaded area represents the 90% error band. $P_1, P_2, \ldots, P_{10}$ denotes the decile groups.

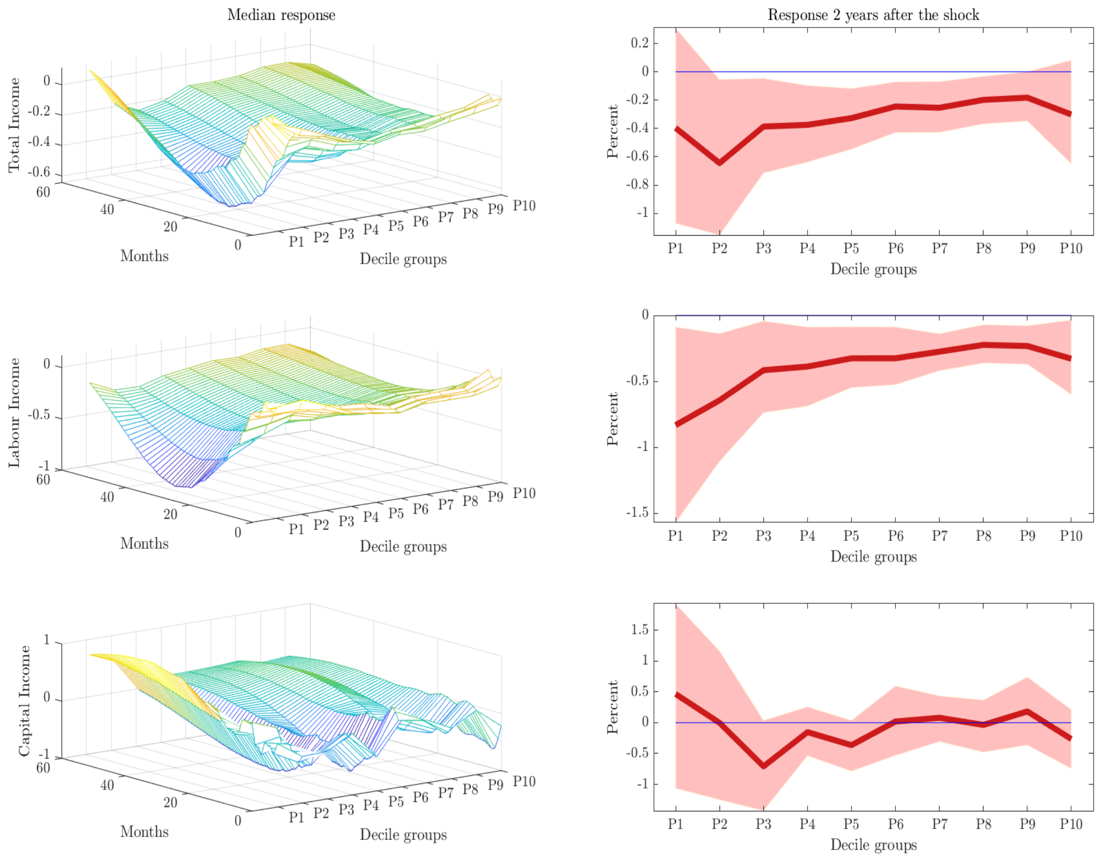

Figure 2: Impulse Response functions from a Bayesian proxy-VAR estimation. The figure displays the 90% confidence bands to an oil price shock. The shock is normalised to produce a 10% increase in the oil price. We used a loose prior for the estimation.

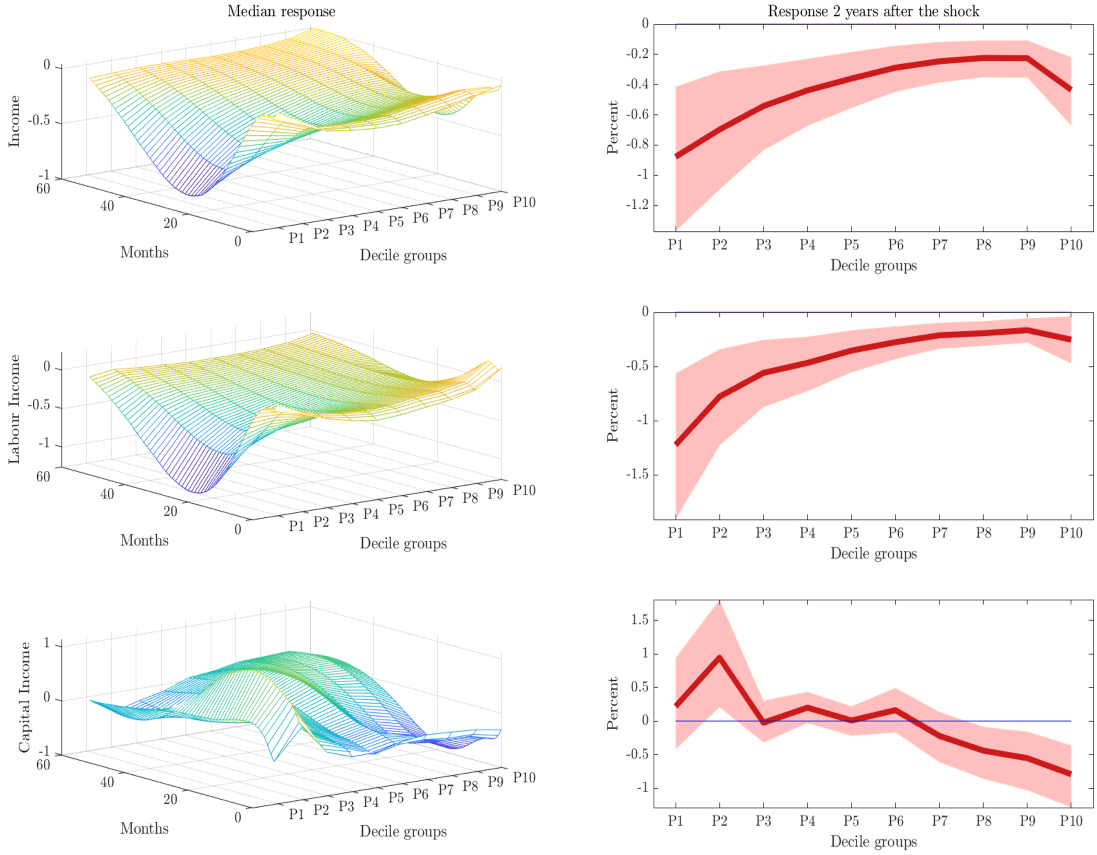

Figure 3: Impulse Response functions from a FAVAR estimation with 6 lags. All other settings equal the baseline specification.

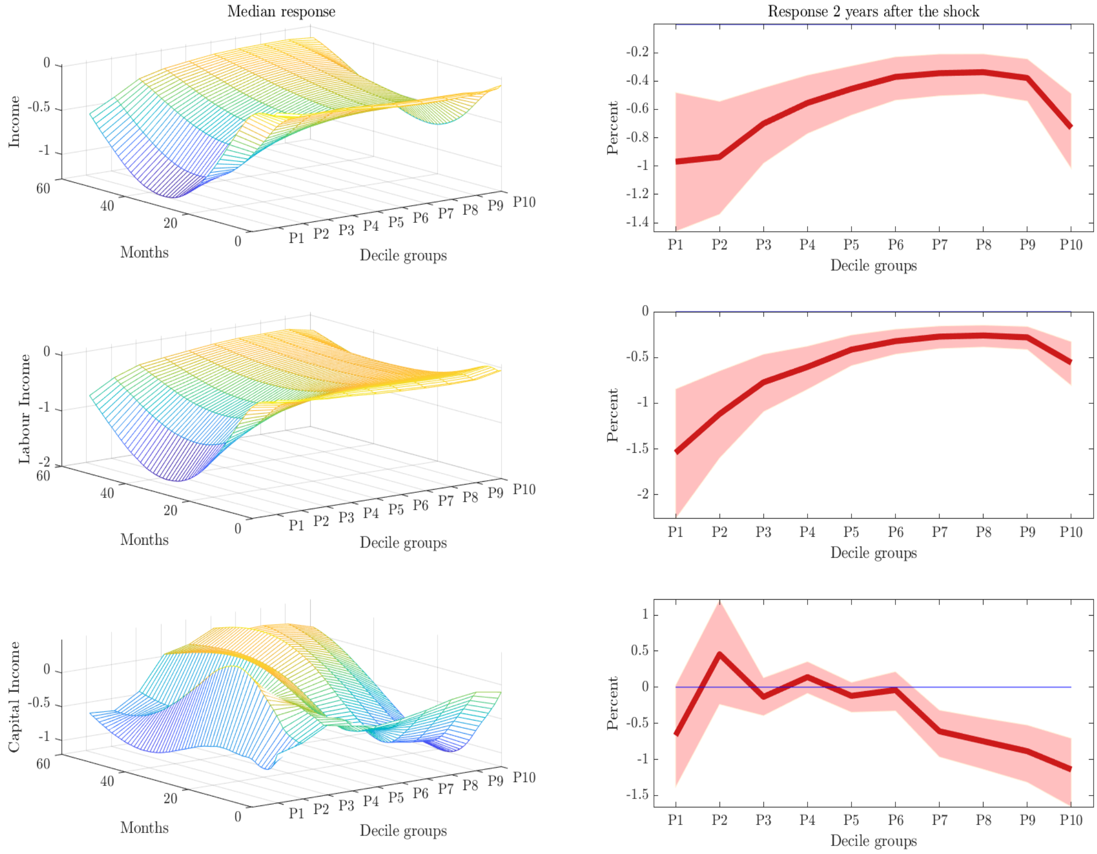

Figure 4: Impulse response functions obtained from the benchmark FAVAR model with 10 factors. All other settings equal the baseline specification.

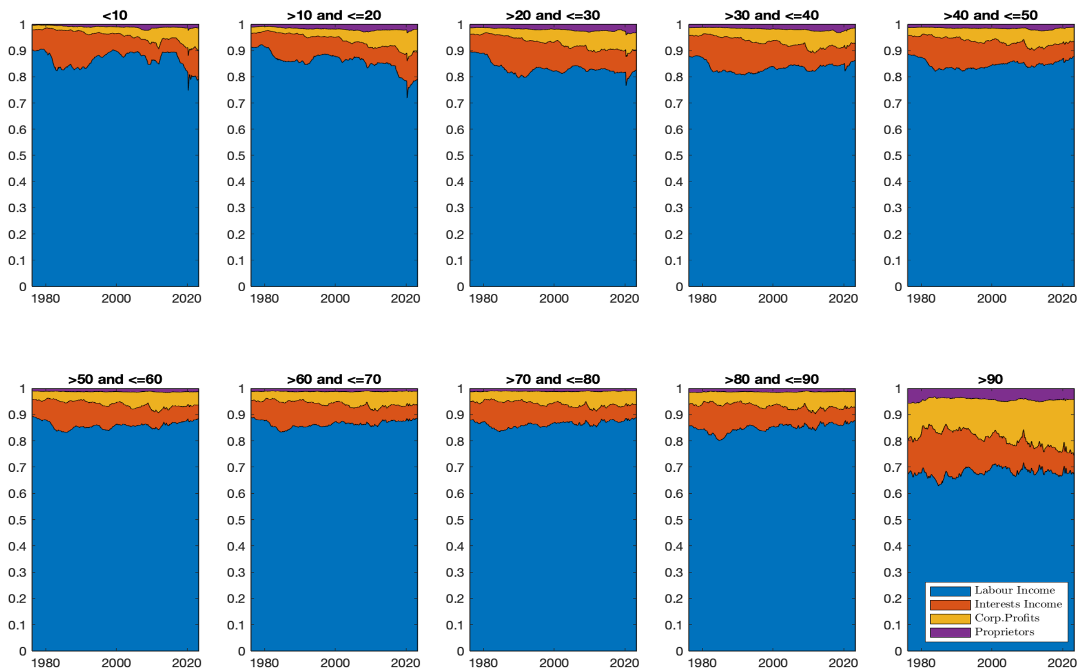

Figure 5: Income Decomposition for every decile group of the income distribution. The figure displays real income values. Labour income is defined as the sum between 0.7*proprietors' income and the compensation of employees. Interest income is the total income generated from currency, deposits, and bonds. Corporate profits are defined as income from businesses, while proprietors' income is the share of income earned by proprietors, classified as capital income (i.e., 30%). Hence, the total of the upper components (interest-, corporate- and proprietors' income) represents capital income.