Not All Oil Supply Shocks are Alike: The Macroeconomic and Distributional Impact of Oil Supply Shocks *

Luis Calderon†

University of Bonn

May 22, 2026

Abstract

Oil supply shocks shape U.S. inequality, and *the type of* oil supply shock matters for who bears the cost. In a Bayesian VAR that embeds the joint distribution of U.S. household income, consumption, and wealth alongside a standard oil-market block, sudden production shortfalls (Baumeister and Hamilton, 2019) raise inequality, whereas oil supply news (Känzig, 2021) mostly reduces it. These differences are not visible in the aggregate dynamics, which appear similar across shocks. A causal mediation analysis (Dufour and Wang, 2024) of these aggregate dynamics shows these shocks operate through distinct channels and it is this mechanism wedge, rather than the aggregate footprint, that determines who bears the cost. The analysis shares the view of Kilian and Lewis (2011) that policy responses should depend on the underlying causes of oil price shocks.

Keywords: *Oil shocks; Income, consumption, and wealth inequality; Household heterogeneity; Functional time-series data; Bayesian VAR; External instruments*

JEL Classification: *C32, C38, D31, E31, E52, Q43*

*All errors are my own.

†Institute for Macroeconomics and Econometrics—University of Bonn, luis.calderon@uni-bonn.de.

1 Introduction

The impact of oil markets on prices and output has long been a central question in macroeconomics, as evidenced by the seminal work of Hamilton (1983). Geopolitical tensions between the U.S. and Iran have renewed interest in the question, with the disruption of the Strait of Hormuz, through which roughly 20% of the world's traded oil passes. This raises the specter of a repeat inflationary episode in the U.S. and globally and has policy makers debating whether they should *look through* these shocks or act proactively. The empirical literature on oil shocks has shown that such movements in oil markets have widespread macroeconomic effects, from Hamilton (1983) through Kilian (2009), Baumeister and Hamilton (2019), and Känzig (2021), but has remained comparatively silent on how these oil-market disturbances propagate to households. Through the lens of a VAR, this paper will show oil supply shocks shape U.S. inequality, and *which type of* oil supply shock matters for who bears the cost.

The present study focuses on two oil supply shocks whose transmission may pass through to households: announcements / news concerning future oil supply and the realized, physical oil supply disruption itself. Känzig (2021) (KZ) identifies oil supply news shocks through high-frequency futures price surprises in narrow windows around OPEC announcements. Along the lines of Beaudry and Portier (2006), Känzig (2021) shows how such news shift the expectations of market participants and drive business cycle fluctuations. Baumeister and Hamilton (2019) (BH) would identify the physical supply shock, identified through informative Bayesian priors on the contemporaneous elasticities of oil supply and demand. A physical supply disruption drives nations to deplete oil inventories to compensate for the sudden shortfall, affecting input costs, industrial production, and eventually the real economy. Both shock series enter the VAR (separately) as internal instruments (Plagborg-Møller and Wolf, 2021).

With that, the paper documents the distributional consequences of both oil supply shock identifications side-by-side. The VAR embeds a parsimonious representation of the joint distribution of household income, consumption, and wealth alongside standard oil-market and macroeconomic variables. In practice, the distributional block will be a set of principal components derived from the Legendre-polynomial coefficients underlying the three-dimensional joint distribution of Bayer, Calderon, and Kuhn (2025). As a consequence of the treatment, I can back out, by linearity, any functional of the joint distribution: group-specific consumption means, robust Gini coefficients, interquantile ratios, joint income–wealth cells, etc. To accommodate the potentially large system, I adopt a Bayesian approach and elect for the asymmetric conjugate prior of Chan (2022), which allows for equation-specific shrinkage—critical to distinguish near-unit-root macroeconomic variables from the stationary distributional factors and the shock instrument itself.

Beyond the impulse responses themselves, I ask *through which channels* oil shocks affect

macroeconomic aggregates. Following the causal mediation framework of Dufour and Wang (2024), I decompose each impulse response into the contributions of individual transmission channels e.g., from real activity, consumer prices, oil market variables, and the distributional block itself. The decomposition can trace, for example, how much of the inflation response is mediated through the price of oil or if monetary policy plays any role in the observed aggregate dynamics (Kilian and Lewis, 2011). Effectively, the decomposition shows over the impulse response horizon when a channel activates, how it propagates, and when it dissipates. This is different than a forecast error variance decomposition, which only informs on the variance contribution of a single shock on each variable. The level decomposition framework of Dufour and Wang (2024), given a single shock, shows the level contribution of *every* channel to *every* impulse response function.

With this richer model in place, I can revisit many conclusions about oil supply shocks and oil shocks in general. First, as most of the literature has already claimed, there is undoubtedly a causal link between oil prices and output: supply-driven oil price increases are contractionary. In response to oil supply shortfalls as defined by Baumeister and Hamilton (2019), real economic activity contracts, but *with a delay*, strengthening the main finding of Kilian (2008b). In contrast to Kilian (2008b), I find precise, persistent positive dynamic effects of headline CPI; however, the *core* CPI response under BH is credibly zero throughout the horizon—supporting Kilian's result that headline inflation is 'flat', shedding light on the importance of relative prices, and is consistent with Hooker (2002), which tested this pass-through from oil prices to inflation under different non-linearities and structural breaks.[1]

Under KZ, by contrast, all prices rise, including core CPI: The news shock is the regime in which oil-shock inflation has bite. I confirm the main findings of Känzig (2021) that oil supply news are indeed inflationary (all prices increase), contractionary, and swift. In the medium-run, for the same oil price impact response, physical disruptions have larger effects on the economy, however, than do oil supply news shocks. Despite the asymmetry in realized inflation dynamics, however, one-year-ahead inflation expectations—measured both for households (Michigan) and professional forecasters (SPF)—respond *identically* to the two shocks, a finding with direct implications for monetary-policy rules that condition on expected inflation.

Oil inventories also tell an interesting story. Under a KZ shock, the increase in inventories that Känzig (2021) interprets as announcement-driven hoarding only materializes after roughly two years, once oil prices ease back toward steady-state values. The channel decomposition clarifies why: the KZ shock channel *does* contribute positive pressure on inventories

[1] I find that CPI owner-occupiers' equivalent rent places considerable downward pressure on CPI inflation, which may be driving the flatness. Recent literature has raised concerns precisely on the measurement of this component of CPI, for which consists of one-fourth the total costs considered in its computation (Chodorow-Reich et al., 2025).

throughout—hoarding intent at the channel level—but it is overwhelmed in the short run by downward contributions from the inflation channel and the inventory channel itself, the latter a characterization of inventory stickiness. The oil-production channel, by contrast, contributes essentially nothing (Figure 23 shows neither OPEC nor non-OPEC production moves upon a KZ shock), so the inventory drawdown is a pure price-response, not a substitution for missing supply. Under a BH shock, the inventory response is consistently negative, in line with depletion of reserves upon a production shortfall. Despite the different signs after year two, inventory responses to both shocks share the same shape—pointing to a common price mechanism: the storage-arbitrage response of Pindyck (2001) and Working (1949), where inventory holders decrease (increase) stocks when spot prices are up (down), irrespective of whether supply is already disrupted (BH) or expected to tighten in the future (KZ).

As for the role of the distribution in the IRFs, I show in Appendix A that distributional factors predict both shocks; thus including the distribution is not merely cosmetic—it gets the aggregate IRFs *right*. Conceptually, the distributional block captures the demand-side information emphasized by Kilian and Lewis (2011)—heterogeneity in expenditure shares and household balance sheets—central to whether these shocks turn out inflationary. The most concrete demonstration of this is the monetary policy response: in the aggregates-only specification, oil supply shocks induce a tightening response of the policy rate; once the household distribution is included, this tightening disappears, consistent with the demand-side channels emphasized by Kilian and Lewis (2011). This has a complementary interpretation in the German context (Broer, Kramer, and Mitman, 2025), where a monetary tightening *does occur* under KZ and accounts for a substantial fraction of employment losses among low-income households. The choice of specification thus carries direct distributional stakes via a monetary channel.

To understand the underlying mechanisms for these results, I decompose the level and the variance of the impulse responses. The variance decomposition results reiterate the point that expectational shifts from oil supply news are more inflationary than realized supply disruptions *despite* similar oil-price responses. Oil supply news explains roughly two to three times the CPI variance share than do physical disruptions, even though the two shocks explain roughly comparable shares of oil-price variance (40% and 30%, respectively, in the medium run). The asymmetry lies in the price-pass-through channel and not in oil-price magnitude. Under BH, the variance contribution to oil prices and inflation is smaller and hump-shaped, implying a build up. Under KZ, they jump strongly on impact and gradually decline. Both shocks drive distributional dynamics similarly, with their FEVD share growing to about 10% over the horizon, consistent with a feedback role for the distribution.

Decomposing the level, I find that under the realized BH shock, the oil market and macroeconomy's responses are mediated primarily by the shock itself and by oil production—a cost-

pull picture in which the physical shortfall (eventually) propagates through real activity. Under the KZ news shock, the response is mediated instead through oil prices and CPI inflation. The contribution of CPI inflation grows before fading slowly, in line with the gradual repricing of staggered price-setters (Calvo, 1983) and the slow updating of inflation expectations (Coibion and Gorodnichenko, 2015). Under KZ, opposite of BH, oil production plays essentially no role, as well as the shock itself, which may be due to OPEC production *and* non-OPEC production virtually not changing over the entire five-year trajectory (Figure 23). As for the U.S. household distribution, it contributes little to the *global* oil-market block, but extending the macro block to more domestic aggregates (Appendix J) reveals it as a medium-run amplifier of certain aggregates. Thus, although the two shocks would produce comparable aggregate IRFs, it would be through *different* mediators—a finding that would motivate different structural models of oil price transmission and naturally different policies.

These different underlying mechanisms will then matter for the distributional outcomes. While both shocks will have strong implications for inequality, the consumption and income dimensions go in opposite directions; wealth inequality, by contrast, moves in the same direction under both shocks, though with very different timing and profile. A sudden physical disruption will raise the Gini across every dimension in the joint distribution; increase the top-10% income share; decrease the bottom-50% consumption share—with asset-rich, income-rich households bearing the cost (in consumption terms) after the one-year lag required for production shortfalls to bind real activity. By contrast, the *asset-rich, income-poor* cells of the joint distribution—low-income households with substantial wealth—systematically benefit in consumption terms at intermediate horizons, as their wealth buffer absorbs the recessionary impulse: in the realized-supply-shock scenario, these are the only winners.

The Känzig (2021) supply news shock finds the opposite at medium horizons: a compressed consumption Gini, a fall in the top-10% income share, and an increase in the bottom-50% consumption share. The mechanism is different: news shocks are priced in opposite directions on impact depending on wealth, with financial-market participants—concentrated at the top of the wealth distribution—repricing expected supply scarcity first, while transfer-dependent households at the bottom are initially shielded by the higher pass-through of energy prices to nominal benefits. The asset-holders in the upper half of the wealth distribution contract consumption preemptively on this repricing—with one exception: at the very top of *joint* income and wealth, capital-income gains from the oil-related repricing dominate the precautionary motive, producing a consumption increase on impact that slowly fades over the horizon, making this group the KZ-side winner. Furthermore, low income and low wealth households decrease their consumption under both shocks in the medium-run, though results are imprecise.

Most relevant literature. This paper connects several strands of the macroeconomic literature. I provide a detailed discussion in Section 2. Briefly, my work builds on the empirical oil shock literature initiated by Hamilton (1983) and developed by Kilian (2008b) and Kilian (2009), Baumeister and Hamilton (2019), and Känzig (2021); the nascent literature on oil shocks and inequality (Berisha et al., 2021; Drossidis, Mumtaz, and Theophilopoulou, 2024; Lei, Ludwig, and Ma, 2025; Broer, Kramer, and Mitman, 2025; Del Canto et al., 2023); the functional VAR methodology for distributional data (Chang, Chen, and Schorfheide, 2024; Chang and Schorfheide, 2024; Lenza and Savoia, 2024; Ettmeier, 2023; Bjørnland, Chang, and Cross, 2023); and the growing body of work that models household heterogeneity (Kaplan, Moll, and Violante, 2018; Auclert, 2019; Bayer, Born, and Luetticke, 2020).

To the best of my knowledge, this paper makes three novel contributions. First, I am the first to compare how different types of oil disturbances — physical versus informational — propagate through the macroeconomy and the household sector. Second, I bring the recent econometric advances (Kilian and Zhou, 2023; Baumeister, 2025; Chan, 2022; Plagborg-Møller and Wolf, 2021; De Graeve and Westermark, 2025) to discipline the empirical framework, and apply the decomposition framework of Dufour and Wang (2024) to trace the shock's transmission through individual channels. Finally, within the same framework, I document how oil supply shocks shape the joint distribution of household income, consumption, and wealth — characterizing not just aggregate responses but who bears the cost across income, wealth, and consumption.

Outline. Section 2 reviews more broadly the related literature. Section 3 describes the data. Section 4 presents the empirical framework. Section 5 and Section 6 report the main results for the macroeconomic aggregates and distribution respectively, comparing always the results from realized supply shocks and supply news shocks. Appendix F collects the robustness specifications. Section 7 concludes.

2 Literature

I build on and contribute to several literatures. I organize the discussion around four themes: the macroeconomic effects of oil shocks, identification in these oil-market systems, the distributional consequences of oil and energy price movements, and functional data methods for distributional macroeconomics.

Oil shocks and the macroeconomy. Oil prices have long been viewed as a key driver of macroeconomic fluctuations. Hamilton (1983) argued for a systematic relation between oil

prices and output, noting oil price increases preceded seven of eight postwar U.S. recessions, motivating a large literature on the causal role of energy markets. This led to a central debate on whether oil price movements (which precede the recessions above) are driven primarily by supply or demand shocks. Using a structural VAR, Kilian (2009) decomposes oil price changes into supply, aggregate demand, and oil-specific demand shocks, showing that their macroeconomic effects differ sharply: demand-driven price increases tend to be expansionary, while supply-driven increases are contractionary. In contrast, Kilian (2008a) finds that exogenous supply disruptions have relatively modest effects on output, challenging earlier views that emphasized supply-side channels.

Identification. Early oil-VAR literature (Kilian, 2009; Kilian and Murphy, 2014a) identify oil shocks with timing zero-restrictions on contemporaneous responses (production and activity do not respond to each other within the month) plus sign restrictions on the remaining structural elasticities (supply slopes up, demand slopes down). Baumeister and Hamilton (2019) highlight that this approach is *set-identifying*, not point-identifying: sign restrictions bound structural parameters to an orthant but do not pin them down, and reported "point estimates" are typically arbitrary draws from an under-identified posterior. Because structural impulse responses are non-linear functions of the contemporaneous elasticities, posterior uncertainty in those elasticities propagates non-linearly into the IRFs, and the same data can support widely different conclusions depending on which point in the identified set one selects.

As such, the field has moved toward alternatives: structural Bayesian VARs with informative priors over the contemporaneous elasticities (Baumeister and Hamilton, 2019), external and internal instruments (Stock and Watson, 2018; Mertens and Ravn, 2013; Känzig, 2021; Plagborg-Møller and Wolf, 2021) that use exogenous variation outside the VAR's information set to identify a column of the impact matrix by construction, and heteroskedasticity-based schemes. I follow the internal-instrument route of Plagborg-Møller and Wolf (2021), augmenting the VAR with the proxy ordered first via a Cholesky decomposition. The proxy point-identifies the column of the impact matrix corresponding to the supply shock under standard relevance and exogeneity conditions, sidestepping the under-identification debate.

Oil shocks, inequality, and the distribution. Compared to the vast literature on aggregate effects, remarkably little is known about the distributional consequences of oil shocks. The existing evidence is sparse and largely confined to reduced-form associations. Berisha et al. (2021) document a positive relationship between oil prices and income inequality (as measured by the Gini coefficient) across U.S. states, using panel data methods. Tan and Uprasen (2023) extend this analysis to the ASEAN countries, finding that rising oil prices tend to increase income inequality in oil-importing countries while reducing it in oil-exporting ones. Del Canto

et al. (2023) study the distributional effects of inflationary shocks—a broader category that includes but is not limited to oil—using a feasible-set approach. They find that oil shocks are regressive, with asymmetric effects along the income distribution: households without a college degree require approximately $800 to afford their pre-shock consumption basket, while middle-aged college-educated households gain $833 from the same shock. Drossidis, Mumtaz, and Theophilopoulou (2024) use a factor-augmented VAR with external instruments to study the effects of oil supply news shocks on the U.S. income distribution. They document large negative effects at both tails of the income distribution—low-income households lose through wages and proprietors' income, while high-income households lose through corporate profits and interest income. Their analysis, however, is confined to income deciles and does not consider consumption or wealth. Baumeister et al. (2025) use a nonlinear heterogeneous agent VAR to study how households with different characteristics adjust their inflation expectations in response to oil supply shocks, combining a macro VAR block with a micro panel, allowing for sign and size asymmetries via Bayesian Additive Regression Trees.

Closest to my work are Broer, Kramer, and Mitman (2025) and Lei, Ludwig, and Ma (2025). Broer, Kramer, and Mitman (2025) use 45 years of German administrative data and the Känzig (2021) oil supply news shock series to document that oil shocks reduce earnings at the bottom of the distribution (by 0.75 percentage points two years after a 10% oil-price rise) and barely affect the top, driven primarily by separation-probability increases at the low end. They decompose the income response into a direct supply effect and an indirect effect from the systematic monetary-policy reaction, using the counterfactual-policy framework of Caravello, McKay, and Wolf (2024), and find that the monetary tightening accounts for a substantial part of the employment losses among the poor. Lei, Ludwig, and Ma (2025) study the United States using an IV-SVAR combined with an incomplete markets model and find that oil price increases raise income and wealth inequality persistently, driven by stagflation that erodes the savings of low-wealth households and widens gaps as interest rates recover faster than wages. They also document a strong asymmetry, with oil price increases raising inequality more than decreases reduce it.

My paper departs from all of these contributions in three respects. First, I study the *joint* distribution of income, consumption, and wealth—not any single margin. The joint distribution reveals mechanisms that marginal distributions cannot: in particular, the consumption insurance provided by flow income and wealth buffers. Second, I compare the distributional incidence of two fundamentally different oil disturbances—the realized supply shock of Baumeister and Hamilton (2019) and the supply news shock of Känzig (2021)—showing that whether the shock is physical or informational matters for *who* bears the cost. Third, I decompose each impulse response into the contributions of individual transmission channels via the

causal-mediation framework of Dufour and Wang (2024), applied within a Bayesian VAR with the asymmetric conjugate prior of Chan (2022). This within-model decomposition identifies not just which variable absorbs the shock but the specific channels through which the response is mediated and how each channel's contribution evolves over the horizon—revealing that the BH/KZ wedge in distributional outcomes is fundamentally a wedge in transmission channels, not in aggregate footprints, a finding that variance decompositions or marginal-distribution analyses would miss.

Functional data methods for distributional macroeconomics. My empirical approach builds on recent work incorporating distributional data into VARs while addressing the key challenge of dimensionality. A full joint distribution over income, consumption, and wealth is too large to include directly, yet scalar summaries like Gini coefficients discard most information. Following Chang, Chen, and Schorfheide (2024) and Bayer, Calderon, and Kuhn (2025), I represent distributions as time-varying densities projected onto a finite set of basis functions whose coefficients evolve in a VAR, enabling standard impulse response analysis. I adopt the related framework of Bayer, Calderon, and Kuhn (2025), using Legendre polynomials and a copula-based representation to capture dependence across income, consumption, and wealth. This approach connects to a growing literature applying functional VARs to heterogeneous data (Ettmeier, 2023; Lenza and Savoia, 2024; Bjørnland, Chang, and Cross, 2023), and extends it by modeling a multivariate (three-dimensional) distribution to study the transmission of oil shocks.

Econometric methodology. I estimate the VAR using the asymmetric conjugate prior of Chan (2022), which allows equation-specific shrinkage—critical in my setting where the oil shock equation requires different regularization than the distributional factor equations. Hyperparameters are selected via marginal likelihood following Giannone, Lenza, and Primiceri (2015). I specify the VAR in log levels with priors for the long run (Giannone, Lenza, and Primiceri, 2019) to accommodate mixed persistence without differencing. I use 16 quarterly lags in the baseline, motivated by recent work showing that long-lag Bayesian VARs can reduce both bias and variance of impulse responses under shrinkage (De Graeve and Westermark, 2025). Ludwig (2024) establishes that iterated VAR impulse responses converge to direct local projection estimates as the lag order grows, and Antolín-Díaz and Surico (2025) demonstrate the practical importance of long lags for identifying the long-run effects of government spending. González-Casásus and Schorfheide (2025) develop an impulse response function criterion for jointly selecting the lag order and shrinkage.

3 Data

The analysis spans several data sources, but will primarily rely on five: (1) quarterly functional time-series data on a U.S. household distribution from Bayer, Calderon, and Kuhn (2025), (2) oil market variables from Känzig (2021), (3) macroeconomic aggregates from McCracken and Ng (2021), (4) the structural oil shock series from Känzig (2021) and (5) from Baumeister and Hamilton (2019). Depending on the estimation, the sample period will cover the mid 1970s to 2024, approximately 160-200 quarterly observations. Similarly, in documenting the robustness of the results, variables considered in the model may change. Details on these data and also data treatment are explained in the following sections. Appendix B provides an overview of all data used in this paper, divided into broad categories.

3.1 Distribution of Consumption, Income, and Wealth

Description. To represent the distributional component of the model, I use the quarterly functional time-series data on the joint distribution of U.S. household consumption, income, and wealth constructed by Bayer, Calderon, and Kuhn (2025). Their methodology synthesizes microdata at the household-level with macroeconomic indicators of higher-frequency (McCracken and Ng, 2021) in a Bayesian state-space framework to produce a smooth and high-frequency representation of the joint distribution. The procedure relies on Sklar (1959): any joint distribution $\Xi_t(\boldsymbol{w})$ over $\boldsymbol{w} \in \mathbb{R}^d$ ($d = 3$ for consumption, income, wealth) can be decomposed into its marginal cumulative distribution functions $\Xi_{m,t}$ for $m \in \{c, y, w\}$ and a copula $C_t : [0,1]^d \to [0,1]$ that encodes the dependence structure independently of the marginals. Representing the joint distribution then amounts to estimating the marginal CDFs and a copula.

Achieving a finite dimensional representation. These distributional objects are infinite-dimensional, making their inclusion in the model infeasible. To handle the infinite dimensionality of these objects, while also maintaining a functional approach, the marginal *quantile functions* $\Xi_{m,t}^{-1}$ and the copula *density* dC_t are projected onto an orthonormal basis of shifted Legendre polynomials $\{Q_o\}_{o\geq 0}$ on $[0,1]$:

$$\Xi_{m,t}^{-1}(u_m) = \sum_{o=0}^{\infty} \xi_{m,o,t}\, Q_o(u_m), \qquad u_m \in [0,1], \tag{1}$$

$$dC_t(u_1, \ldots, u_d) = \sum_{(o_1,\ldots,o_d)\in\mathbb{N}_0^d} \kappa_{(o_1,\ldots,o_d),t} \prod_{m=1}^{d} Q_{o_m}(u_m). \tag{2}$$

By orthonormality of the basis, the polynomial coefficients are inner products between the Legendre polynomials and the corresponding distributional object, be it $dC_t(u_1, \ldots, u_d)$ or $\Xi^{-1}_{m,t}(u_m)$. By infinite dimensionality of the distributional objects, the inner product is defined in the $\mathcal{L}^2$ space, which is the expected value of their product. This implies that the corresponding sample analogue is just a sample average over ranked microdata, shown in Equations 3 and 4:

$$\hat{\xi}_{m,o,t} = N_t^{-1} \sum_i w_{m,i,t}\, Q_o(u_{m,i,t}) \qquad \text{for the quantile function} \tag{3}$$

$$\hat{\kappa}_{(o_1,\ldots,o_d),t} = N_t^{-1} \sum_i \prod_m Q_{o_m}(u_{m,i,t}) \qquad \text{for the copula} \tag{4}$$

$$\boldsymbol{\xi}_{m,t} = \left(\xi_{m,0,t}, \ldots, \xi_{m,O,t}\right)', \tag{5}$$

$$\boldsymbol{\kappa}_t = \left[\kappa_{(o_1,\ldots,o_d),t}\right]_{(o_1,\ldots,o_d)\in\{0,\ldots,O\}^d}, \tag{6}$$

$$\boldsymbol{\theta}_t = \left[\, \boldsymbol{\xi}'_{1,t},\ \ldots,\ \boldsymbol{\xi}'_{d,t},\ \ \mathrm{vec}(\boldsymbol{\kappa}_t)' \,\right]' \in \mathbb{R}^N \tag{7}$$

where $u_{m,i,t}$ is the data rank, derived from the cumulative sample weight at observation i after sorting the dimension m in the data and $w_{m,i,t}$ the data corresponding to the rank: the (treated) quantile for dimension m. Truncating the sums of Equation 1 and 2 at order O provides an approximation of the distributional objects. This truncation leaves us with a finite number coefficients 5 and 6 to be estimated, achieving the finite representation required to estimate the model. Once vertically concatenated, they form a finite vector $\boldsymbol{\theta}_t \in \mathbb{R}^N$, with $N = (O+1)^d + d \cdot (O+1)$ entries (quantile coefficients across d dimensions plus copula coefficients). These polynomial coefficients $\boldsymbol{\theta}_t$ carry the business-cycle fluctuations of the joint distribution; the basis functions Q_o are known and invariant over time. For the interested reader, Appendix C restates the decomposition in greater detail.

Representation in BVAR. What will be important is how these data are represented in the Bayesian VAR. At the moment, to estimate the model with the coefficients $\boldsymbol{\theta}_t$ directly, although finite, would result in far too many equations. To reduce the dimensionality of the system, I exploit the correlational structure of $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_T]'$ and extract the principal components of the coefficient matrix via PCA. From Bayer, Calderon, and Kuhn (2025), I know the data generation process of the joint distribution(s) estimated here and thus also know the first 8 principal components, $\boldsymbol{F}_t = (F_{1t}, \ldots, F_{8t})'$, capture *all* the business cycle variation in the consumption, income, and wealth distribution. It is precisely these factors that enter the baseline Bayesian VAR. The treatment of the distributional data thereafter ensures impulse reponses $\boldsymbol{\Psi}^F_h$ still re-

tain its properties as a distribution.[2] Appendix C explains how to then recover distributional moments of interest from these factor responses $\mathbf{\Psi}_h^F$, be it consumption of different household groups, inequality indices, and other conditional moments.

Robustness. I compare results where only the marginals of the distribution are included in the VAR. This amounts to only retaining coefficients $\boldsymbol{\xi}_{m,t}$ and in terms of variation, would be in line with the pre-dominant specification(s) found in the empirical inequality literature, allowing for a direct comparison. See Appendix D for more results.

Operationalization of concepts. The distributional factor estimates $\boldsymbol{F_t}$ rely on U.S. micro-data, U.S. macro data, a state-equation dictating the law of motion for the distribution, and a measurement equation which optimally weights the contribution of each data source to each estimate at each point in time, and which also reconciles differences in how each underlying micro-source operationalizes consumption, income, and wealth. In that sense, the operationalization of each concept — consumption, income, wealth — is a convex combination of these data sources and what the model believes is the best estimate for the unobserved distributional object.

Transforming these factors to observables, however, requires de-standardizing the coefficients with means and standard deviations estimated from a specific micro-dataset. This choice of de-standardization dataset enters only through the steady-state mean and unconditional volatility of each variable; the dynamics of the responses I report below are unaffected, since the measurement equation has already reconciled the operationalization differences across the underlying micro-sources. Since I am interested in the consumption, income, and wealth distribution, I use means and standard deviations from the PSID. To nonetheless provide some idea of the underlying variation constituting these estimates, I provide a description of the operationalization of concepts used in the PSID.

In the PSID, I measure **consumption** as containing food, rent (for renters) or housing rental equivalence (6% of home market value, for homeowners), utilities, health care, public transportation, education, childcare, and gas/vehicle repairs. **Income** contains labor earnings, public transfers, professional/self-employment income, rental income, dividends, and business/farm income, but is not net of taxes. **Wealth** is total assets minus total debt. Assets comprise liquid assets (checking, savings, CDs, money market, T-bills) plus illiquid assets (housing

[2] Specifically, the reconstruction from factor estimates to distributional data preserves the properties of the underlying distributional objects across the impulse-response horizon $h \in 0, \ldots, H$: the reconstructed marginal quantile functions $\hat{\Xi}_{m,h}^{-1}$ are weakly monotone in $u \in [0,1]$ for each m; the reconstructed copula density $d\hat{C}_h$ is non-negative on the trimmed cube $[\varepsilon, 1-\varepsilon]^d$; and the copula integrates to unity, $\int [0,1]^d d\hat{C}_h = 1$. Uniform marginals follow from the orthonormality of the Legendre basis. Monotonicity and non-negativity are verified ex post on a G^d tensor grid; deviations from the strict constraints are below 10^{-6} and treated as numerical noise.

net of property debt, real estate, automobiles, retirement plans, life insurance, stocks, mutual funds, business equity). Debt comprises housing debt (mortgages, home equity lines of credit) plus personal debt (car loans, student loans, credit cards, medical, legal).

3.2 Oil Market and Macroeconomic Variables

Description. Alongside the distributional factors, the baseline VAR includes six oil market and macroeconomic variables: the real price of crude oil (West Texas Intermediate deflated by CPI-U), world crude oil production, a global industrial production index (OECD plus six major economies), U.S. industrial production, the consumer price index, and a proxy for OECD crude oil inventories. This block follows the standard specification in the oil shock literature (Kilian, 2009; Baumeister and Hamilton, 2019) to capture the global oil market and following Känzig (2021), I augment the global oil market structure with U.S. industrial production and U.S. consumer prices to capture domestic real activity and the inflation channel. Again, this defines the baseline of the model. I perform a battery of robustness checks to address concerns related to the above selection of variables. They are discussed below.

As my baseline inventory measure, I use the OECD log-inventory series constructed by cumulating Baumeister and Hamilton (2019)'s monthly Δi_t data into a quarterly log-inventory series. The underlying source is the same Kilian and Murphy (2014a) OECD series, also used by Hamilton (2009), Kilian and Murphy (2014b), and Känzig (2021), but the deseasoning and aggregation pipeline differs and is, by the argument of BH (Section III), less contaminated by measurement error, accounting for attenuation of its own estimates and general contamination of the system. I anchor the cumulated series at the level (of the noisy inventories variable) in 1975Q1 so that levels are comparable. BH's replication archive ends in 2019Q4; I extend the series through 2024Q2 by replicating their construction on current EIA inputs (U.S. crude, OECD petroleum, U.S. petroleum stocks, global production). The replication agrees with BH's published Δi_t to a correlation of 0.997 over the 2000–2019 overlap window; details and the data-vintage caveat are in Appendix E. The results with the "noisy" series are provided for comparison in Appendix F.1.

Representation in BVAR. All oil market and macroeconomic variables enter in log levels rather than first differences and scaled by 100 for interpretability. First-differencing imposes that all structural shocks have permanent effects on the level of every variable, since a shock to Δy, the first difference, implies a level-shift forever. This would be a (unecessary) restriction at odds with economic theory, which holds that some shocks (e.g., monetary policy) have only transitory effects, while others (e.g., technology shocks) do not; especially since prop-

erly modeling the VAR in levels does not affect inference.[3] More on how my Bayesian setup accommodates these variables in Section 4. No further data treatment is taken.

Robustness. Here I consider robustness (and comparisons) using alternative oil or macro specifications, all detailed alongside their data sources in Table 3 of Appendix B.

Forward-looking information. Motivated by Mori and Peersman (2024), I address the potential predictability of both shocks by augmenting the baseline with a set of financial variables to capture forward-looking information. Their assessment on the predictability of the shocks relied on monthly data, so I computed my own Granger-causality tests for the quarterly data I use here (see Appendix A). The results show that the excess bond premium (EBP) of Gilchrist and Zakrajšek (2012) predicts the KZ shocks and the interest rates (term spread and short-term rates) predict the BH shocks. I have a robustness check that augments the baseline with precisely these channels. For completeness, I also report the Mori-Peersman block (VIX, S&P500, 1Y interest rate, EBP) that confounds monthly estimates, see Appendix F.2.

FRED-QD common factors. In a more brute force approach, I replace or augment the macro variables with principal component factors extracted from the FRED-QD quarterly database (McCracken and Ng, 2021)—a large panel of 113 macroeconomic series. I exclude from the PCA input panel any series that *already* appears as a named VAR observable; see Table 3, so the factor loadings cannot say *double-count* regressors. I use the levels factors $\boldsymbol{F}_t^{\text{lev}}$ in the spirit of Bai (2004), extracted from the log-level series. Factor count is selected following Freyaldenhoven (2022). Appendix A shows indeed these factors also predict the instrument. Does this predictable variation overlap with the financial information set above? To see, I run a regression of each financial variable on the level factors. Appendix A show the level factors explain 92–99 % of variation in the rate-and-equity Granger-causing block above, 78% the term spread, but around 20% for EBP and 14% for VIX. Full impulse responses augmenting the baseline VAR with the levels factors are reported in Appendix F.7.

Oil-price measure. I replace WTI with two alternatives: the U.S. refiners' acquisition cost (RAC) of imported crude and the real-log Brent front-month price; construction details in Appendix B. This follows concerns raised by Kilian and Zhou (2023), where they argue WTI (and similar *U.S.* domestic prices) shouldn't proxy the *global* oil price post-1974 because U.S. price regulation until the early 1980s and 2010–2015 transportation bottlenecks broke arbitrage with world

[3] Sims, Stock, and Watson (1990) show that OLS estimation of a VAR in levels produces consistent parameter estimates and asymptotically valid impulse responses regardless of the integration order of the variables, provided the model includes an intercept, so that pre-testing for unit roots is unnecessary. Toda and Yamamoto (1995) shows hypothesis testing on potentially mixed order systems as ones handled in Sims, Stock, and Watson (1990) is made possible via lag-augmentation, where the additional lags pick up the nonstationary nuisance parameters.

markets and that one should use the U.S. refiners' acquisition *cost* of imports (or Brent, post-mid-1980s); however Känzig (2021) shows in Appendix A.18-19 that results do not change when using the real refiner acquisition cost as oil price indicator, so I adopt the same baseline measure. I have run a robustness comparing impulse responses under these alternative measures and the qualitative conclusions are unchanged; figures available on request.

Capturing global demand. The baseline VAR uses world industrial production from Baumeister and Hamilton (2019) (OECD plus six major economies). Following Känzig (2021), as a robustness, I substitute in Kilian's (2009) Real Economic Activity (REA) index—a dry-cargo single-voyage shipping rate—motivated by the short-run identifying assumption that vessel capacity is fixed, so rate fluctuations track demand rather than supply (Kilian, 2019; Funashima, 2020). Two caveats apply: Kaufmann (2011) documents no statistically measurable relation between the shipping index and oil consumption over 1968–2008, weakening its interpretation as an oil-demand proxy; and Kolodzeij and Kaufmann (2014) flag that pairing the shipping index with RAC, as would be implied by Kilian (2009) and Kilian and Zhou (2023), as the price measure produces mechanical regression artifacts because transportation costs appear on both sides of the relevant equations. Neither concern appears in the baseline (World IP and real WTI); REA and RAC enter only as separate robustness checks on the activity and price measures, respectively.

Inventory measure. The baseline VAR uses the BH-cumulated inventory series described above; in a robustness (more of a comparison) I revert to the series of Kilian and Murphy (2014a) and Känzig (2021), which Baumeister and Hamilton (2019)'s Section III flags as carrying substantial measurement error. Their level correlation is 0.997, yet their $\Delta \log$ correlation is only 0.27, indicating that approximately 73 % of quarter-on-quarter variance is non-shared "construction noise." Reverting to the noisier series gives an explicit reading of how much of the baseline result is sensitive to the measurement-error treatment. Comparison of results are in Appendix F.1; the more detailed comparison of the two series is in Appendix E.

Production aggregation. The baseline VAR uses world crude oil production. Following Kolodzeij and Kaufmann (2014), who argue that OPEC and non-OPEC producers operate under different output-setting criteria, I report a robustness that decomposes world oil production into OPEC and non-OPEC components. This is an interesting exercise, especially for the Känzig oil supply news shock because it reveals how non-OPEC nations react to these news, but also whether OPEC follows through on their announcements. See Appendix F.3.

All robustness exercises are estimated with the same prior and identification strategy as the

baseline.

3.3 Oil Shock Series

To provide a thorough understanding of how oil supply disruptions propagate through the macroeconomy and household distributions, I provide a baseline estimation covering two empirically identified structural oil supply shock series.

Realized structural supply shock. The first shock is the realized structural oil supply shock of Baumeister and Hamilton (2019), identified through informative Bayesian priors on the structural impact matrix of a four-variable oil market VAR.[4] This shock captures *sudden* physical supply disruptions—*actual* shortfalls in oil production.

Oil supply news shock. The second shock are oil futures price surprises of Känzig (2021), constructed from a composite measure of WTI crude oil futures price changes in a narrow window around OPEC announcements, spanning the first-year term structure. This shock captures news regarding the future supply and operates through the *anticipation* channel. The announcements on oil production may or may not materialize later on—a key difference from Baumeister and Hamilton (2019).

The two identifications therefore isolate different stages of the same process. The KZ shock measures the market's *initial* response to OPEC's stated intentions—the repricing of future supply expectations on announcement day. The BH shock measures the *realized* supply outcome—what actually happened to production, regardless of what was announced. The wedge between the two shocks reflects the difference between OPEC announcements and realized production.[5]

3.4 Sub-samples

I re-estimate the baseline VAR on three alternative subsamples that correspond to changes in the macro and oil-market regime. (1) The *post-1982* sample, which drops periods before

[4]The BH identification rests on prior beliefs about the contemporaneous structural elasticities of oil supply and demand—in particular, a tight prior centered on a small short-run price elasticity of demand ($\alpha_{pq} \approx -0.1$), disciplined by the micro-evidence on gasoline and oil demand at the pump (Hamilton, 2009). BH (their Section IV.D) argue that the wider, more "agnostic" priors used by Kilian and Murphy (2014a) are informative *against* the established micro evidence and effectively put non-trivial mass on values inconsistent with that literature. My use of their shock therefore inherits this prior structure: the supply-vs-demand decomposition that produces the realized supply shock is identified up to BH's prior over the four contemporaneous elasticities. The Känzig (2021) shock, identified by 30-minute announcement-window futures surprises, sidesteps this dependency by isolating supply news without taking a stand on demand elasticities, and serves as my comparison shock.

[5]Realized production at the world level mixes OPEC compliance with non-OPEC production; Appendix F.3 reports a robustness that disentangles the OPEC and non-OPEC components, which directly provides some evidence of commitment (does the OPEC follow through), as well as whether non-OPEC offsets in any way. Baumeister (2023) show that OPEC compliance varies substantially over the sample period considered here.

1982Q1, corresponds to an institutional regime change in OPEC: formal production quotas were introduced in March 1982, so the post-1982 sample is the period in which OPEC announcements correspond to quota commitments—the institutional setup the Känzig (2021) identification implicitly assumes. As Nakov and Pescatori (2010) emphasize, the period around 1981 was a time of dramatic change in world oil markets, in domestic energy production and consumption, and in U.S. monetary policy and the inflation environment; Hooker (2002) documents that the oil-price-to-inflation pass-through itself weakens substantially after this regime break, so the subsample doubles as a check that the inflation responses I report are not artifacts of pre-Volcker passthrough. (2) The *pre-Covid* sample truncates at 2019Q4, dropping the pandemic and its inflation aftermath. (3) The *pre-shale* sample truncates at 2010Q4, before the U.S. tight-oil boom materially raised the medium-run elasticity of global oil supply. Subsample IRFs are reported in Appendix F.4.

4 Bayesian Framework and Identification

This section presents the empirical framework for generating causal effects of oil supply shocks on macroeconomic outcomes and household distributions. This consists of a VAR that embeds both distributional and macroeconomic variables, a prior on the parameter space to discipline the model responses, and an identification strategy based on internal instruments. Each of these are discussed in turn.

4.1 Bayesian VAR

I embed the distributional factors alongside macroeconomic aggregates and oil market variables in a reduced-form VAR:

$$\boldsymbol{Y}_t = \boldsymbol{c} + \boldsymbol{A}_1 \boldsymbol{Y}_{t-1} + \cdots + \boldsymbol{A}_p \boldsymbol{Y}_{t-p} + \boldsymbol{u}_t, \qquad \boldsymbol{u}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}), \tag{8}$$

where $\boldsymbol{Y}_t = [z_t, \boldsymbol{M}_t', \boldsymbol{F}_t']'$ collects the oil supply instrument $z_t \in \mathbb{R}$, macroeconomic and oil market variables $\boldsymbol{M}_t \in \mathbb{R}^m$ following Känzig (2021), distributional factors $\boldsymbol{F}_t \in \mathbb{R}^d$ from Bayer, Calderon, and Kuhn (2025), and the reduced-form residual u_t . For the baseline, $\boldsymbol{M}_t$ includes the real price of crude oil (WTI deflated by CPI-U), world crude oil production, a global industrial production index (OECD plus six major economies), U.S. industrial production, the consumer price index, and the OECD crude oil inventories of Baumeister and Hamilton (2019). The number of lags is denoted by p and is equal to 16. With 15 variables and 16 lags, each equation contains 241 regressors (including constant). To address the resulting high dimensionality, I adopt a Bayesian framework.

Asymmetric conjugate prior. The standard Bayesian estimator of (8) stacks the reduced-form coefficient matrix $\boldsymbol{A} = [\boldsymbol{c}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_p]$ and places a conjugate Normal–Inverse-Wishart (NIW) prior:

$$\operatorname{vec}(\boldsymbol{A}') \,\big|\, \boldsymbol{\Sigma} \;\sim\; \mathcal{N}\big(\operatorname{vec}(\boldsymbol{m}'),\; \boldsymbol{\Sigma} \otimes \boldsymbol{\Omega}\big), \qquad \boldsymbol{\Sigma} \;\sim\; \mathcal{IW}(\boldsymbol{\Psi}, \nu). \tag{9}$$

where $\boldsymbol{m}$ is the prior mean coefficient matrix (Minnesota-style: zero on every entry except the own first lag of each variable, set to a persistence parameter δ shared across all equations); $\boldsymbol{\Omega}$ is a positive-definite matrix encoding lag-decay shrinkage on the regressors; and $\boldsymbol{\Sigma}$ is the covariance matrix of the reduced-form residuals $\boldsymbol{u}_t$, on which one can place an Inverse-Wishart prior with scale matrix $\boldsymbol{\Psi}$ ($n \times n$, positive definite) and degrees of freedom $\nu > n - 1$.

The Kronecker structure $\boldsymbol{\Sigma} \otimes \boldsymbol{\Omega}$ ties the prior covariance *across* equations: the same lag-decay matrix $\boldsymbol{\Omega}$ governs every equation's coefficients, scaled only by the contemporaneous covariance $\boldsymbol{\Sigma}$. In a Minnesota parameterization this collapses to a single scalar *own*-lag tightness κ_1, a single *cross*-lag tightness κ_2, and a *single* persistence parameter δ shared by every variable. This is restrictive: the shock instrument z_t should be a priori unpredictable ($\delta = 0$, very tight κ_1); macroeconomic aggregates are near-unit-root ($\delta \approx 1$, looser κ_1); distributional factors are stationary, smooth, and numerous, and benefit from aggressive cross-equation shrinkage ($\kappa_2 \to 0$). One $(\kappa_1, \kappa_2, \delta)$ triple cannot accommodate all three.

Triangular reparameterization. Chan (2022) replaces the symmetric NIW with a per-equation Normal–Inverse-Gamma prior by reparameterizing (8) into recursive structural form. Decompose the reduced-form innovation covariance as

$$\boldsymbol{\Sigma} \;=\; \boldsymbol{A}_0^{-1}\, \boldsymbol{D}\, \boldsymbol{A}_0^{-\top}, \qquad \boldsymbol{A}_0 \;=\; \boldsymbol{I} - \boldsymbol{L}, \tag{10}$$

where $\boldsymbol{L}$ is strictly lower triangular (zero diagonal, free entries a_{ij} for $i > j$) and $\boldsymbol{D} = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_n^2)$. Pre-multiplying (8) by $\boldsymbol{A}_0$ transforms the system to

$$\boldsymbol{A}_0\, \boldsymbol{Y}_t \;=\; \boldsymbol{A}_0\, \boldsymbol{c} \;+\; \boldsymbol{A}_0\, \boldsymbol{A}_1\, \boldsymbol{Y}_{t-1} \;+\; \cdots \;+\; \boldsymbol{A}_0\, \boldsymbol{A}_p\, \boldsymbol{Y}_{t-p} \;+\; \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \;\sim\; \mathcal{N}(\boldsymbol{0}, \boldsymbol{D}), \tag{11}$$

with structural innovations $\boldsymbol{\varepsilon}_t = \boldsymbol{A}_0\, \boldsymbol{u}_t$ that are mutually orthogonal by construction. Reading off equation i gives the recursive form

$$y_{i,t} \;=\; \underbrace{-\sum_{j<i} a_{ij}\, y_{j,t}}_{\text{contemporaneous block}} \;+\; \boldsymbol{x}_t'\, \boldsymbol{\beta}_i \;+\; \varepsilon_{i,t}, \qquad \varepsilon_{i,t} \;\sim\; \mathcal{N}(0, \sigma_i^2), \tag{12}$$

where $\boldsymbol{x}_t = [\boldsymbol{Y}'_{t-1}, \ldots, \boldsymbol{Y}'_{t-p}, 1]'$ collects the lagged regressors and the constant. Conditional on the contemporaneous coefficients $\{a_{ij}\}_{j<i}$, equation i is a single-equation linear regression with Gaussian errors; the structural innovations $\varepsilon_{i,t}$ are independent across equations, so equation-by-equation priors are now coherent. This is the step that breaks the Kronecker constraint of (9).

Stacking the unknowns of equation i as $\boldsymbol{\theta}_i = [a_{i,1}, \ldots, a_{i,i-1}, \boldsymbol{\beta}_i']'$ (contemporaneous coefficients first, then lag coefficients and constant), Chan (2022) places independent Normal–Inverse-Gamma (NIG) priors:

$$\boldsymbol{\theta}_i \,\big|\, \sigma_i^2 \;\sim\; \mathcal{N}\big(\boldsymbol{m}_i,\; \sigma_i^2\, \boldsymbol{V}_i\big), \qquad \sigma_i^2 \;\sim\; \mathcal{IG}\Big(\tfrac{\nu_i}{2},\; \tfrac{S_i}{2}\Big), \qquad i = 1, \ldots, n. \tag{13}$$

The hyperparameters (ν_i, S_i) govern the marginal Inverse-Gamma prior on the structural-shock variance: $\mathbb{E}[\sigma_i^2] = S_i/(\nu_i - 2)$ for $\nu_i > 2$. I set $\nu_i = 3$ and $S_i = \hat{\sigma}_i^2$ so that $\mathbb{E}[\sigma_i^2] = \hat{\sigma}_i^2$, again anchored on the AR(p) residual variance. The Inverse-Gamma is the univariate analog of the Inverse-Wishart, conjugate to the equation-i Gaussian likelihood, so the posterior of $(\boldsymbol{\theta}_i, \sigma_i^2)$ is again Normal–Inverse-Gamma and draws are available in closed form.

The prior mean $\boldsymbol{m}_i$ is zero on every entry except the coefficient on the own first lag of variable i, which is set to δ_i (the equation-specific persistence prior). The prior variance is diagonal with the Minnesota structure of Litterman (1980) and Doan, Litterman, and Sims (1984):

$$V_i^{(j,\ell)} \;=\; \begin{cases} \kappa_{1i} \,/\, (\ell^{2\lambda_3}\, \hat{\sigma}_i^2) & \text{own lag } (j = i), \\[2mm] \kappa_{2i} \,/\, (\ell^{2\lambda_3}\, \hat{\sigma}_j^2) & \text{cross lag } (j \neq i), \\[2mm] \kappa_{2i} \,/\, \hat{\sigma}_j^2 & \text{contemporaneous coefficient } a_{ij}, \\[2mm] \kappa_4 & \text{intercept}, \end{cases} \tag{14}$$

where $\hat{\sigma}_j^2$ is the residual variance from a univariate AR(p) for variable j, ℓ indexes the lag order, λ_3 governs lag decay, and κ_4 is a fixed loose prior on the constant. The crucial difference from (9) is that each equation now has its own $(\kappa_{1i}, \kappa_{2i}, \delta_i)$.

Equation-specific hyperparameter selection. Conjugacy of (13) delivers a closed-form equation-level marginal likelihood. I exploit this to select $(\kappa_{1i}, \kappa_{2i}, \delta_i)$ *per equation* by grid search:

$$(\kappa_{1i}^*, \kappa_{2i}^*, \delta_i^*) \;=\; \arg \max_{(\kappa_1, \kappa_2, \delta) \,\in\, \mathcal{G}} \; \log\, p\big(\{y_{i,t}\}_{t=1}^T \,\big|\, \kappa_1, \kappa_2, \delta;\; \mathcal{H}_{-i}\big), \tag{15}$$

where $\mathcal{H}_{-i}$ holds the system-wide hyperparameters fixed at their current values. The persistence grid is restricted as

$$\delta_i \in \begin{cases} \{0\} & \text{if } i \leq n_s \quad \text{(shock equations: dogmatic white-noise prior)}, \\ [0,\, 1] & \text{otherwise}, \end{cases} \tag{16}$$

where n_s is the number of shock instruments. The shock equation is pinned at $\delta_i = 0$ to encode the assumption that z_t is unpredictable from its own past; every other equation selects δ_i^* from the data, so I(0) factors typically receive small δ_i^*, near-I(1) macro variables receive $\delta_i^* \approx 1$, and stationary-but-persistent variables select intermediate values. The shared hyperparameters—λ_3 and the long-run prior tightness (Section 4.1)—are optimized at the system level using the marginal-likelihood approach of Giannone, Lenza, and Primiceri (2015).

The end result: every equation receives its own shrinkage $(\kappa_{1i}^*, \kappa_{2i}^*)$ and its own persistence prior δ_i^*, all chosen by the data, in contrast to the single $(\kappa_1, \kappa_2, \delta)$ imposed by the symmetric NIW prior in (9). The oil shock equation requires little shrinkage; the distributional factor equations, smooth and numerous, receive aggressive cross-equation regularization. Posterior draws are independent across equations, conditional on $(\kappa_{1i}^*, \kappa_{2i}^*, \delta_i^*)$, and available in closed form.

Long-run priors. All oil market and macroeconomic variables enter in log levels. To discipline long-run behavior without imposing unit roots, I augment the likelihood with the *prior for the long run* (PLR) of Giannone, Lenza, and Primiceri (2019), which regularizes the long-run multiplier matrix $\boldsymbol{A}(1) = \boldsymbol{I} - \boldsymbol{A}_1 - \cdots - \boldsymbol{A}_p$ and lets the data determine which linear combinations of variables share common stochastic trends, rather than hard-coding cointegration rank ex ante. The construction is data-adaptive (ADF-based stationarity classification + auto-detected cointegrating pairs), and the long-run tightness is optimized jointly with the Minnesota hyperparameters via marginal likelihood. Full implementation details are in Appendix G.

Lag length. I set $p = 16$ quarterly lags. Kilian and Zhou (2023) and Hamilton and Herrera (2004) show longer lags are needed to allow for the building and contraction of oil market cycles. This choice is also motivated by recent evidence that, under Bayesian shrinkage, long-lag VARs can simultaneously reduce both the bias and variance of impulse response estimates (De Graeve and Westermark, 2025). Relatedly, Ludwig (2024) shows that iterated VAR impulse responses converge to direct local projection estimates as the lag order grows with the horizon, implying that long-lag VARs inherit the misspecification robustness of local projections while

retaining the efficiency of the parametric model. Following comments from Baumeister (2025), which shares the view of González-Casásus and Schorfheide (2025), I do not select the lag order via AIC or BIC, which target forecast loss and can yield specifications that are suboptimal for structural inference. Instead, I wish to speak on short- to medium-run dynamics and thus, p is set to reflect that. In Appendix H, I show how impulse responses change from varying the lag order p, which will take values $\in \{2, 4, 8, 12, 20\}$ around the baseline $p = 16$.

Posterior simulation. Following Chan (2022), the posterior of each equation is available in closed form as a NIG distribution, and draws are independent. I draw 5,000 independent coefficient vectors with a stable companion matrix (eigenvalues within the unit circle), discarding explosive draws, and compute structural impulse responses for each. Credible bands are constructed as pointwise 68% posterior intervals; 90% bands and posterior point summaries other than the median (mean and mode) are available on request and yield qualitatively identical conclusions.

4.2 Identification

To identify oil supply shocks, I use the internal instrument approach of Plagborg-Møller and Wolf (2021). The shock series z_t is included directly in the VAR as the first variable, ordered before the real oil price and all other endogenous variables.

Relevance and exogeneity in a Bayesian context. Identification is then based on treating each external proxy separately as an instrument, z_t, that satisfies the relevance and exogeneity condition regarding its underlying structural shock (Stock and Watson (2012) and Mertens and Ravn (2013))[6]:

$$cov(z_t, \varepsilon_{1t}) = \alpha \neq 0 \tag{17}$$

$$cov(z_t, \varepsilon_{-t}) = 0, \tag{18}$$

where ε_{1t} denotes the first shock in the system and ε_{-t} the remaining $n - 1$ shocks. Condition (17) is the relevance condition of the instrument and is testable. Condition (18) is the exogeneity of the instrument. Because ε_{-t} is unobserved, this condition is not *directly* testable, but its observable implications can be examined—for example, orthogonality of z_t to pre-determined macro variables and to alternative external instruments for non-target shocks. I present such diagnostics in Appendix A for both shocks.

[6]Sufficient lags absorb the lag-exogeneity component by whitening the residuals; the lead-exogeneity component is an assumption about the construction of the proxy, which in my case follows from KZ's OPEC-announcement timing / BH's sign-restriction design.

Under the Bayesian internal instrument approach employed here, relevance governs the informativeness of the likelihood about the structural impulse responses. A weak instrument implies a flat likelihood in the direction of the shock of interest, so that the posterior is dominated by the prior: credible bands remain wide and centered on prior-implied dynamics rather than being updated by the data. I provide evidence that this concern does not apply in my setting: the FEVD in Section 5 will show that the Baumeister and Hamilton (2019) shock explains a non-trivial share of the forecast error variance of all variables, and under the oil supply news instrument, the impulse responses largely align with those obtained in Känzig (2021), for which relevance is well-established.

Fundamentalness. A further consideration is the fundamentalness of the instrument relative to the VAR's information set. A shock is fundamental if it can be recovered from current and past observables in the system; non-fundamentalness arises when agents possess information not spanned by the econometrician's VAR. Miranda-Agrippino and Ricco (2023) show that, under non-fundamentalness, the external SVAR-IV approach produces biased impulse responses, and Bruns and Lütkepohl (2025) formally establish that the external and internal approaches are not equivalent precisely in this case.

This is what makes the internal instrument approach of Plagborg-Møller and Wolf (2021) attractive. It resolves the point mechanically by augmenting the VAR with z_t directly: the system's information set is enlarged to span the instrument, and fundamentalness holds by construction within the augmented system. The inclusion of the instrument, as opposed to the external-instrument approach, does affect inference. It may produce wider credible bands, since the model propagates the instrument's full parameter uncertainty—its own AR dynamics, its loadings on the other variables, and the associated rotation of the structural impact matrix — rather than only the first-stage impact uncertainty as with the external instrument approach (see Caldara and Herbst, 2019).[7] In this sense my inference is conservative. This pattern of small bands in external vs. large bands in internal is visible in Degasperi, Figure C.4, panel (a).

Structural impulse responses. Consider the mapping between the reduced form errors u_t and the structural shocks ε_t of Equation (8):

$$u_t = \boldsymbol{B}_0 \varepsilon_t \tag{19}$$

[7] This widening is more pronounced when the instrument is weak, because a fully Bayesian treatment internalizes the weak-signal case as part of the posterior, whereas the external approach takes the first-stage estimate as a *point* and builds bands conditional on it, so the instrument's own uncertainty does not propagate.

where $\text{cov}(u_t) = \boldsymbol{\Sigma} = \boldsymbol{B}_0\boldsymbol{B}_0'$. I identify *oil supply* (Baumeister and Hamilton, 2019) and *oil supply news* (Känzig, 2021) shocks based on the internal instrument approach motivated by Plagborg-Møller and Wolf (2021) and order the analogous instrument z_t first in the following Cholesky ordering:

$$\boldsymbol{Y}_t = [z_t, \boldsymbol{M}_t', \boldsymbol{F}_t']'. \tag{20}$$

Ordering the instrument first ensures that the first structural shock, ε_{1t}, is the innovation to z_t that is orthogonal to lagged information in the system. The first column of $\boldsymbol{B}_0$ gives the impact responses of all variables to this shock. The remaining $n-1$ shocks are indeterminate up to an orthogonal rotation (Stock and Watson, 2018). As shown by Plagborg-Møller and Wolf (2021), this approach is robust to non-invertibility, abstracts from a necessary first stage regression outside the VAR and instrument relevance is reflected in the width of the posterior credible bands. The shock is normalized to produce a 10 percent increase in the real price of oil on impact.

The Wold moving average coefficients Ψ_h are computed recursively from the estimated VAR coefficients:

$$\Psi_0 = I, \qquad \Psi_h = \sum_{\ell=1}^{\min(p,h)} A_\ell \, \Psi_{h-\ell}, \quad h \geq 1. \tag{21}$$

The structural impulse response of variable i at horizon h to the identified oil shock is then

$$\Theta_{i,h} = e_i' \Psi_h B_0 \, e_1, \tag{22}$$

where e_i and e_1 are selector vectors. Since the oil shock is ordered first, the first column of B_0 gives the on-impact responses, and Ψ_h propagates these forward through the VAR dynamics.

5 Macroeconomic Effects of Oil Supply Shocks

This section presents the first part of the empirical results, with the focus on macroeconomic aggregates. The section opens with the aggregate impulse responses.[8] I then present forecast error variance decompositions, to provide some evidence on the relevance of the oil supply shocks on the uncertainty of the variables in my system. Next, I use the framework of Dufour and Wang (2024) to decompose the *level* of the IRFs and uncover the different channels driving the observed responses using the framework of Dufour and Wang (2024). Throughout the results section, I compare the Baumeister and Hamilton (2019) physical oil supply shocks with

[8] Appendix F.5 provides a comparison between my VAR augmented with distributional factors and those without as typically examined in the literature.

the Känzig (2021) oil supply news shock, to distinguish the two economic channels: sudden physical disruptions versus expectational shifts in oil supply.

To complement these aggregate responses, Appendix I runs a *rotating* Factor-Augmented Vector Autoregression (FAVAR). This amounts to running many Bayesian VARs, where one observable is *rotated* out for each Bayesian VAR, with the idea of obtaining evidence on the transmission of both shocks across the many parts of the macroeconomy. This strategy has been similarly adopted by Känzig (2021) and Lei, Ludwig, and Ma (2025), among others, and a quick way to obtain *a sense* of the transmission mechanisms one can draw from these shocks.[9]

5.1 Aggregate Impulse Responses

Figure 1 reports the impulse responses of key macroeconomic and oil market variables to the two shocks. To allow for a clean comparison, both shocks are normalized to produce a 10 percent increase in the real price of oil on impact.[10] Given they are otherwise identical systems, the normalization allows us to even the field and ask how an oil price hike transmits in both settings. All impulse responses are estimated to horizon $h = 20$.

First, Panel A of Figure 1 shows the classical oil supply shock: a sudden production shortfall, coupled with a depletion of oil inventories. The model finds sudden disruptions in the availability of oil *increases* general prices on impact; which remain elevated for about two years. Industrial production, both for the U.S. and abroad, show a rather delayed effect, with U.S. industrial production decreasing near 1% two years later. The initial rise in World and U.S. IP seems to be mechanical. The geography of World IP is OECD + 6 non-member economies, many of which are oil-producing nations that benefit from the OPEC production cuts, taking prices as given. You can see from U.S. IP that they are a contributor of the rise. Eventually, demand side factors do drive production downward. I have a robustness specification in Appendix F.6 which replaces World IP with Kilian's (2009a) measure of real economic activity (REA), deflated dry-cargo shipping rate. There, it is free of this geographic artifact, showing a much more muted rise and quicker decline.[11]

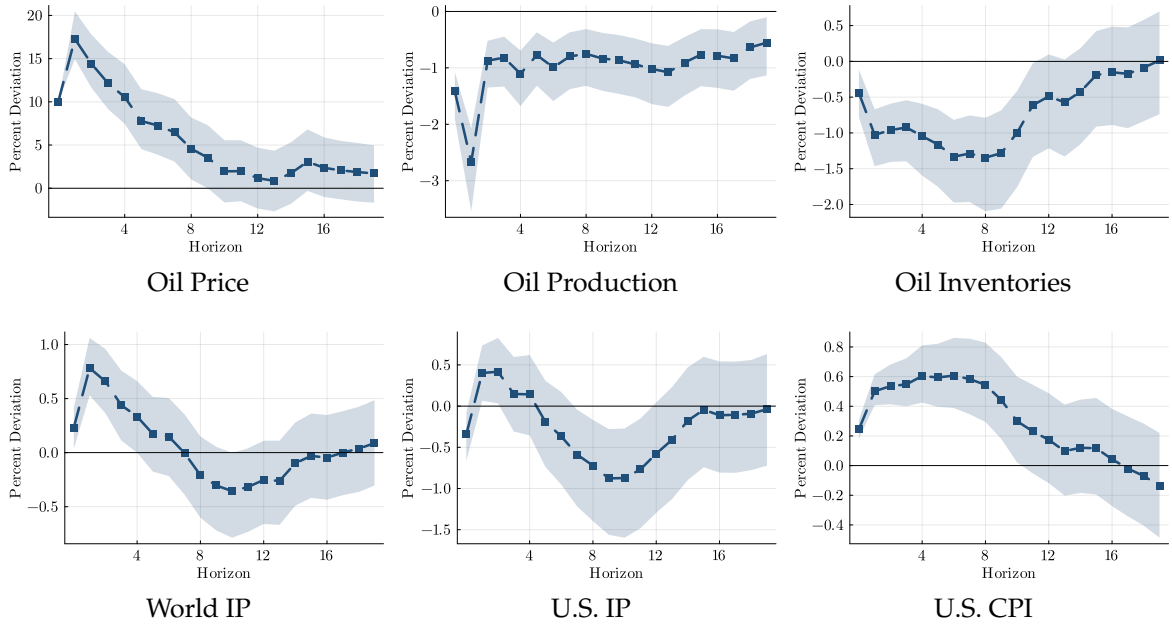
[9] I generate IRFs for variables on real economic activity, labor markets, income, fiscal policy, consumer prices, consumption, interest rates, financial conditions, credit supply, household debt, and others, for the U.S. economy. For each variable, I plot two IRFs: the observable IRF with aggregate factors identified with the McCracken and Ng (2021) FRED-QD and another where distributional factors are added on top.). These responses were only meant to inform my hypothesis of the main results, which *do* model the oil structure, which I have argued is important for identification. The rotating FAVARs *do not* have the oil variables. Nevertheless, the sensitivity I raised thus far on modeling the oil variables concerned only oil inventories. The decomposition I show later on provides a formal identification of these mechanisms within a single model.

[10] The variable normalized on (the oil price) responds contemporaneously under both shocks, so the contemporaneous unit-effect normalization is appropriate (Stock and Watson, 2018, §2.1.3); the variables that respond with a delay under KZ (production, inventories) are correctly traced as IRFs from this normalization, not used as the normalizing object.

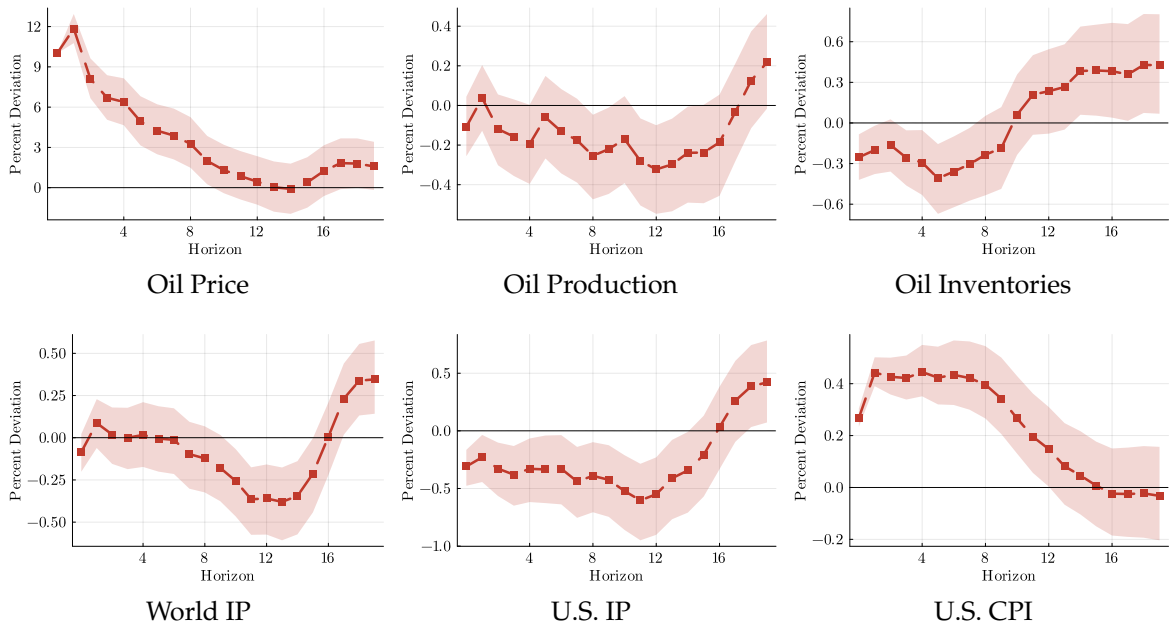
[11] Kaufmann (2011) documents that the dry-bulk shipping index — the basis of the Kilian REA — has no statistically measurable relation to oil consumption over 1968–2008, weakening its interpretation as a global oil-demand proxy. The baseline VAR uses worldip (OECD + 6 industrial production); REA enters only as a robustness on the

Figure 1: Aggregate impulse responses to oil supply shocks

Panel A: Baumeister & Hamilton (2019) realized supply shock



Panel B: Känzig (2021) supply news shock



Notes: Impulse responses to oil supply shocks normalized to produce a 10% increase in the real price of oil on impact. Solid lines: posterior median. Shaded areas: 68% credible sets. Horizon in quarters.

To continue, Panel B of Figure 1 shows the macroeconomic responses under an oil supply news shock. For prices, the same pattern emerges: prices jump on impact and remain elevated for two years before gradually returning to their steady state levels. Interestingly, industrial production falls on impact and for three years, is negative. The response of World IP is quite muted relative to the BH shock and similarly decreases with a delay as in the BH shock, albeit without the strong decline in oil production. The IRFs suggest that the anticipation channel dominates the oil-price-driving extraction effect: in mere anticipation of a lower oil supply following OPEC's announcement, industrial firms reduce production. I confirm this in Figure 4 which shows the channel decomposition. Another interesting phenomenon of these news shock is the recovery: World IP, U.S. IP, and oil production observe a significant rise around year five, returning and exceeding pre-shock levels, as opposed to the BH shock. This feature is robust to several specifications.

Discussion on oil inventories. Under the oil supply news shock, I can only confirm the baseline response of oil inventories Känzig (2021) finds for the medium, but not for the short term. Whereas Figure 3 of KZ documents the characteristic precautionary "piling up" of inventories on impact—consistent with the interpretation that news shocks shift inventory demand in anticipation of future supply shortfalls—my estimates show inventories *declining* initially (about one-third of the BH response), with an accumulation emerging only after approximately two years. This result is present in virtually all my robustness checks and is not an artifact of the different inventories variable I use here (see Figure 17). Other checks show their result is sensitive to primarily truncation bias (Figure 34) and similarly, omitted variable bias (Figure 35 and 32). Examining the KZ results, I find two drivers of the deviation: the absence of an oil-market structure in the VAR and additional conditioning variation that absorbs macroeconomic dynamics unrelated to oil markets.[12]

I do not interpret my findings as an absence of the speculative-demand mechanism, but rather for oil inventories, there may be evidence that other forces operate in the opposite direction and dominate at the announcement horizon. For example, *storage arbitrage* may exist, where the impact spot-price spike compresses the carry premium and pushes optimal inventories *down* (Pindyck, 2001). Second, *imperfect "cartel" discipline*, for which there is substantial evidence. OPEC members routinely produce above their assigned quotas: historical compliance rates range from 40 to 80 percent, varying with the price cycle and internal cohesion. This

activity measure.

[12] Känzig (2021) runs a robustness check with FRED-MD factors layered on his 6-variable VAR but does not report the result. Running the baseline (oil + distributional factors) with FRED-QD macroeconomic factors on top (Figure 32) shows a compromise: oil inventories decline on impact, but remain stable for two years (credible interval contains zero), then increase afterwards similarly to my IRF. I thus identify a sensitivity of the KZ result to truncation bias (Figure 34) and omitted variable bias (Figure 32).

is why several suthors have stated that OPEC is not a cartel (Colgan, 2014; Ratti and Vespignani, 2015; Molchanov, 2003; Baumeister and Kilian, 2016). To the extent that OPEC production announcements affect oil prices (conditional on non-members beliefs), the probability of deviation necessarily compresses the revision in $\mathbb{E}_t[P_{t+1}]$ (from a change in expected supply) that would otherwise motivate hoarding; and lastly, perhaps the *post-shale supply elasticity* weakens the precautionary motive over the post-2010 part of my sample, which would be roughly 1/4 of the entire sample. This last channel is assessed Figure 26 and indeed shows less depletion of oil inventories in the first year. [13]

What conclusions can we draw? First, the realized BH shock and the supply news KZ shock produce remarkably similar oil-price dynamics and aggregate dynamics, but the timing of the general dynamics is worth noting: The *real*-side transmissions are fundamentally different. The BH shock operates through a *cost-pull* channel—actual production shortfalls raise input costs on impact (Appendix Figure 45 and 38), but real activity (Appendix Figure 36), labor markets (Appendix Figure 37) and the SP500 (Appendix Figure 41) contract with a one year lag required for those costs to bind, strengthening the result of Kilian (2008b). CPI energy and CPI transportation (Figure 38) do rise on impact and remain elevated for roughly two years, consistent with energy-cost pass-through into both upstream producer prices and the energy/transport components of consumer prices; but, CPI-core does not—the credible interval contains zero throughout the entire IRF horizon—again in line with Kilian (2008b), who find this similar profile but for CPI inflation, and with Hooker (2002), who documents that the pass-through from oil prices to inflation weakens substantially after the early-1980s regime break. In contrast, the corresponding KZ panels show that input costs rise, but real activity, labor markets, and the SP500 all respond on impact and the contractions are persistent. Furthermore, KZ is inflationary—all prices rise.

A striking related observation concerns *inflation expectations*. Using the University of Michigan household one-year-ahead inflation expectations and the SPF median one-year CPI forecast (Figure 44), the response of expected inflation to BH and KZ shocks is essentially identical. Both surveys—the household survey and the professional-forecaster survey—treat the two shocks the same, even though their underlying inflation dynamics differ: BH raises headline but not core, KZ raises all prices including core. Households and forecasters either do

[13] Signs of this result also appear in Känzig (2021). In fact, a closer look shows the early negative response and subsequent rise more closely resembles the inventories response to oil supply *surprises* in KZ's two-shock specification (Appendix Figure A.13), and aligns with his LP-IV results (Appendix Section A.2.2, Figure A.4), which shows the strong negative response when one adds more lags/more controls.[14] They also align with the Figure 3 oil inventories response of Mori and Peersman (2024), which as mentioned already, deal with a specific form of contamination in the instrument.[15] The result suggests the presence of other forces that matter for the anticipation channel in oil inventories, which actually relieves concerns that this is a special storage demand shock as stated by Kilian and Zhou (2023).

not distinguish between physical and informational oil disturbances, or they do not have the information set the VAR uses to decompose them. This has direct monetary-policy implications: a central bank that responds to inflation expectations sees the *same* signal regardless of shock type, even though the warranted policy response is plausibly different (more muted under BH given the absent core response; more active under KZ given the persistent broad pass-through).

What about monetary policy? The KZ inflation response induces a somewhat stronger monetary policy reaction on impact, but the effect is not credibly different from zero. Beyond impact, policy is largely unresponsive over the first two years and subsequently eases in response to recessionary pressures; the BH shock exhibits a similar profile. Taken together, these results suggest that monetary policy does not systematically tighten in response to oil supply shocks, even if they may be inflationary, consistent with Kilian and Lewis (2011), who emphasizes the importance of demand-side channels. Notably, excluding the household distribution yields the opposite pattern—an increase in policy rates—indicating that incorporating the distribution dampens the case for a tightening response.

What about the medium-run? As a result of the cost-pull channel, the BH shock generates larger effects despite the delay. Except for World IP, aggregate responses are larger in the medium run under a BH shock, indicating the presence of adverse general equilibrium effects kicking in. Rotating FAVAR results show this is a systematic response. The decomposition later on will reveal the key drivers of these effects.

And what about recovery? The year-five recovery and overshoot under KZ, absent under BH, is itself interesting: when expectations of supply scarcity prove only partially realized or even *not* realized (consistent with the OPEC non-compliance evidence in the inventories discussion above and also Figure 24), the implicit precautionary contraction unwinds—capital, inventory, and saving decisions made under the initial belief gradually reverse over the horizon and real activity rebounds toward—and in some cases beyond—its pre-shock level. Both findings echo Kilian (2009)'s seminal point that *not all oil price shocks are alike*—but here the two shocks fall on opposite sides of the same OPEC decision, depending on whether identification picks up the *announcement* or the *realization*.[16]

[16]The BH shock concerns *world* oil production, so, not exactly opposites of the same OPEC decision and indeed, Figure 24 shows both OPEC and non-OPEC oil production decline on impact, but *only* OPEC production is negative and different than zero across nearly all horizons—the change in non-OPEC oil production is indistinguishable from zero.

5.2 Forecast Error Variance Decompositions

Description. Figure 2a and Figure 2b present the forecast error variance decomposition (FEVD) of the identified shocks.[17] The BH shock explains approximately 60 percent of oil production variance near impact, gradually declining to below 40% over the long horizon; about 30 percent of oil price variance, and 10–20 percent of CPI variance. In the distributional block, it accounts for little on impact, but incrementally grows past 10% at horizon $h = 20$, a non-trivial share, suggesting that the joint distribution contains non-negligible information about oil shock transmission that grows gradually. Oil inventories would follow a similar profile. Under the current specification, the shock explains relatively little of industrial production variance, both in the U.S. and abroad—echoing Kilian (2009)'s finding that exogenous oil-supply disruptions contribute only modestly to real-activity fluctuations.[18]

In comparison, the KZ news shock explains the largest share of forecast variance for the two macro variables most associated with oil episodes: the oil price and CPI. About 45% of oil price variance is explained by the shock on impact, declining modestly to 40% by horizon $h = 20$, which is larger than the BH shock (30% stable). CPI variance follows a similar trajectory, jumping to 40%, but declines quicker. Unlike the BH shock, oil production starts near zero and rises to just under 10% — the shock doesn't move production immediately (it is news about future supply), but gradually explains more as anticipated production adjustments perhaps materialize. The distributional block's FEVD share grows similarly to under the BH shock—small on impact, 10% by horizon $h = 20$—confirming that the distribution carries non-trivial information about transmission under both identifications. Oil supply shocks drive less the uncertainty of the other variables, but nonetheless the contribution of the shocks are distinguishable and non-trivial, providing evidence of their relevance.

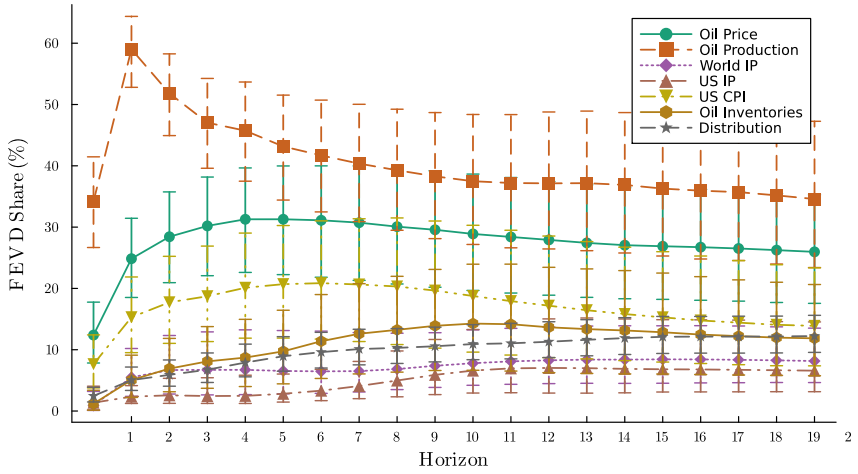
Main finding. Expectational shifts from oil supply news are markedly more inflationary than realized supply disruptions *despite* the similar oil-price magnitude in Figure 1. The KZ shock explains roughly two to three times the CPI variance share than the BH shock does, even though the two shocks explain roughly comparable shares of oil-price variance (40% and 30%, respectively, in the medium run). The asymmetry lies in the price-pass-through channel and not in oil-price magnitude. Under BH, the contribution to oil prices and inflation is smaller and hump-shaped, implying a build up. Under KZ, they jump on impact and gradually decline. Regarding the distribution, both shocks drive distributional dynamics similarly, with their FEVD share growing to about 10% over the horizon, consistent with a feedback role for the distribution, which I further example through a within-model decomposition in the next

[17] Under Stock and Watson (2018), these are identified, since the IRFs are identified and the system is, by construction, invertible.

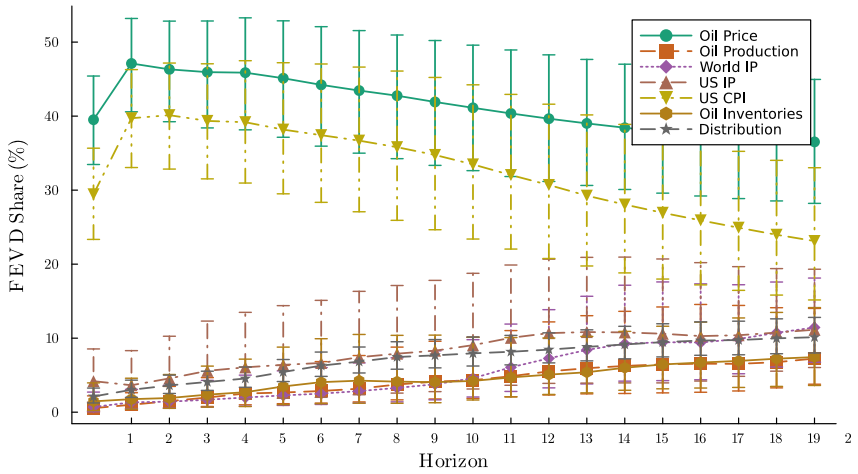
[18] This result is also robust to using the measure of Kilian (2009). The result is stored and available upon request.

section, Section 5.3. To my knowledge, this is the first FEVD of identified oil supply shocks for distributional aggregates.

Figure 2: Forecast error variance decomposition (FEVD): oil supply shocks



(a) Realized oil supply shock (Baumeister and Hamilton, 2019)



(b) Oil supply news shock (Känzig, 2021)

Notes: Share of forecast error variance explained by the identified shocks for each variable in the VAR, including the distributional block. Markers: posterior median. Error bars: 68% credible sets. Horizon in quarters.

5.3 What Drives Aggregate Responses?

The aggregate responses of CPI to both shocks are quantitatively similar; however they are the net effect of several channels and from solely the IRFs, it is not clear what drives the CPI response for each shock. To understand the underlying mechanisms, I apply the generalized impulse response (GIR) framework of Dufour and Renault (1998) and the causal mediation interpretation of Dufour and Wang (2024).[19] This methodology decomposes each impulse re-

[19]This is *not* to be confused with the generalized impulse responses of Pesaran and Shin (1998), which only captures the causal effect of a structural shock, possibly in a nonlinear setting, but does not capture the *mediation*

sponse into variable-level contributions, revealing not only *which* channels drive the response but also *when* each channel's influence materializes and how large this contribution is.[20,21]

The Dufour and Wang (2024) decomposition. The key insight is that the total impulse response $\Theta_{i,h}$ of variable i at horizon h can be written as a sum of *channel contributions*:

$$\Theta_{i,h} = \sum_{m=1}^{n_y} C_{m\to i}^{(h)}, \qquad C_{m\to i}^{(h)} = \sum_{n=0}^{h} c_{m\to i}^{(n,h)}, \tag{23}$$

where $C_{m\to i}^{(h)}$ is the total contribution of variable m to the response (the IRF) of variable i at horizon h. The impulse response above would then be the sum of $C_{m\to i}^{(h)}$ for all n_y variables, each a possible mediator/channel m. To compute the contribution of a specific channel, $C_{m\to i}^{(h)}$, requires estimating $c_{m\to i}^{(n,h)}$, which is the contribution made at some intermediate date n, for each n until we reach horizon h. One can then interpret (n,h) in superscript as the contribution from n to h. This contribution $c_{m\to i}^{(n,h)}$ can be decomposed into a product of two forces:

$$c_{m\to i}^{(n,h)} = \sum_{\ell=1}^{\min(n+1,\,p)} \underbrace{[\varphi_{\ell}^{(h-n)}]_{i,m}}_{\substack{\text{Propagation:}\\ \text{how } m \text{ at lag } \ell\\ \text{reaches } i \text{ over}\\ h-n \text{ periods}}} \times \underbrace{\Theta_{m,\,n+1-\ell}}_{\substack{\text{Activation:}\\ \text{how strongly}\\ \text{the shock moves}\\ m \text{ at horizon } n+1-\ell}} \tag{24}$$

For intuition on Equation 24, I provide the first two terms of $C_{m\to i}^{(h)}$, complemented with Figure 3, to better understand what exactly is being measured. The first term is $c_{m\to i}^{(0,h)}$, which corresponds to $n=0$, the date of the shock's impact. The term quantifies the contribution of channel m that occurrs at $n=0$ (if activated) and how this channel's response $\Theta_{m,0}$ propagates from $n=0$ to $n=h$. So, the shock hits at $n=0$, moves m on impact, m then propagates to variable i over h periods. For $n=1$, we only need to consider the propagation from $n=1$ to $n=h$, hence the $(h-1)$, where the superscript on the φ terms can be interpreted as the length of propagation. Then it's about measuring how the impact $n=0$ response contributed to this in-between horizon of $n=1$ and $n=h$ and how the $n=1$ response contributed to the same in-between horizon. The remaining terms for $n=2$ to $n=h$ are assessed analogously.

effects (the channels), the object of interest here.

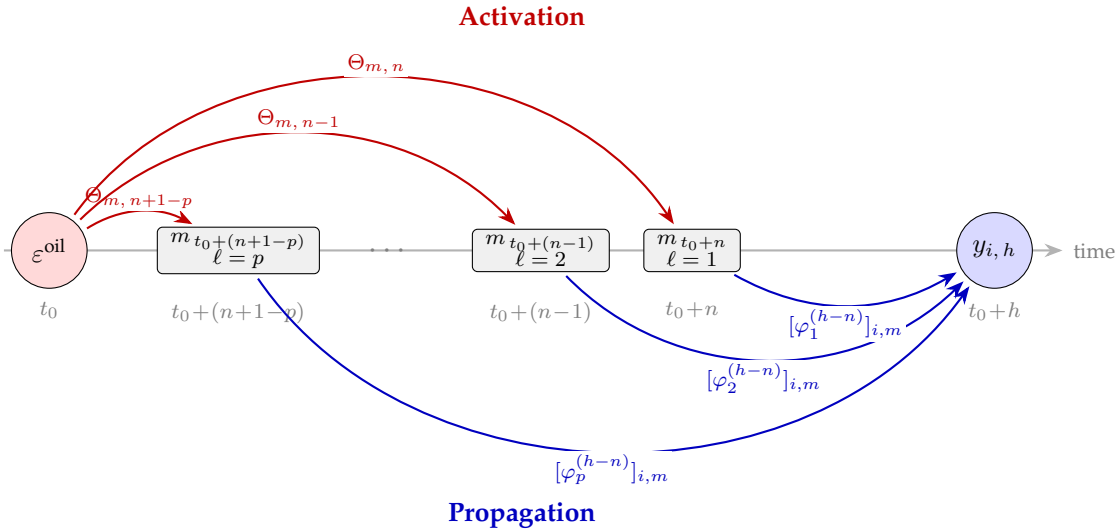
[20] An important disclaimer: this is not a counterfactual exercise and more a descriptive accounting exercise i.e., an ex-post attribution of the responses to different channels, given the DGP; however, the decomposition is an important empirical step to identify an active channel which can in turn be used to discipline a structural model that then microfounds that channel. One can then do actual policy counterfactuals in the structural model.

[21] The decomposition is order-invariant as the moments used to identify the contributions are also order-invariant.

$$c_{m\to i}^{(0,h)} = \underbrace{[\varphi_1^{(h)}]_{i,m}}_{\text{propagation over } h \text{ periods}} \times \underbrace{\Theta_{m,0}}_{\text{impact response of } m} \tag{25}$$

$$c_{m\to i}^{(1,h)} = \underbrace{[\varphi_1^{(h-1)}]_{i,m}\,\Theta_{m,1}}_{m \text{ at } n=1 \text{ propagated over } h-1 \text{ periods}} + \underbrace{[\varphi_2^{(h-1)}]_{i,m}\,\Theta_{m,0}}_{m \text{ at } n=0 \text{ entering as a 2-period lag}} . \tag{26}$$

Figure 3: Anatomy of a single channel contribution $c_{m\to i}^{(n,h)}$



Notes: Visualization of equation (24) for a single mediating variable m, evaluation horizon n, and outcome horizon h. The oil supply shock ε^{oil} fires at time t_0 and *activates* the channel variable m at each of the time points $t_0 + (n + 1 - \ell)$ for $\ell = 1, \ldots, \min(n+1, p)$ (red arcs above the time axis). Each activated value of m then *propagates* forward to outcome i at time $t_0 + h$ (blue arcs below). Arc length is drawn proportional to time-distance: more-recent m values have long activation arcs and short propagation arcs, older m values the reverse, with shock-to-outcome spans summing to h in every case. The $n = 0$ case in equation (25) corresponds to keeping only the rightmost arc pair ($\ell = 1$); the $n = 1$ case in equation (26) adds the next pair ($\ell = 2$), and so on up to $\min(n + 1, p)$.

Measuring propagation. $\varphi_\ell^{(h)}$ measures the amount of propagation over h periods and acts as a dynamic multiplier that scales the activation into the outcome IRF. The subscript ℓ identifies which lag of channel m is being propagated, so at each intermediate date n, I have the total contribution from m aggregates $\varphi_\ell^{(h-n)}$ across each of its active lags $\ell = 1, \ldots, \min(n+1,\, p)$. These are the GIR coefficients of Dufour and Renault (1998), computed via the recursion $\varphi_\ell^{(h+1)} = \varphi_{\ell+1}^{(h)} + \Psi_h A_\ell$ with initial conditions $\varphi_\ell^{(1)} = A_\ell$. Element $[\varphi_\ell^{(h)}]_{i,m}$ captures how a perturbation in variable m at lag ℓ propagates to variable i over h periods, accounting for all intermediate feedback through every other variable in the system.

The decomposition allows to examine many questions such as the role of inflation on the transmission of such shocks; the role of financial conditions; or the role of the distributional in shaping the aggregate responses themselves; all within a single model.

Why not two VARs? ? I investigate the impact of oil supply shocks through the lens of a *single* model, as opposed to running two VARs, where one is some baseline system and the other removes some channel of interest, keeping everything else the same. A naive econometrician would attribute the difference in aggregate impulse responses to the omitted block / the channel of interest; but the two-VAR gap conflates omitted-variable bias in the smaller model, genuine mediation through the distribution, and identification drift (impact vector not invariant to information set)—which cannot be separated in a two-VAR comparison. The GIR decomposition I adopt in this section is immune to all three issues: it fixes the model, fixes the identification, and analytically attributes the identified impulse responses above to each mediator through the Wold representation (and not through omission). This setup permits to make quantitative statements like "channel m accounts for $X\%$ of the macro response"—not possible with two VARs.

Results. Figure 4 presents the channel decomposition for the aggregate responses for both shocks. The stacked bars show the posterior median contribution of each channel at each horizon, where each channel is identified by a separate color and the color scheme is identical across panels, so the meaning of purple for example is the same across panels. The dashed white line and outlined white circles correspond to the median point estimates from Figure 1.

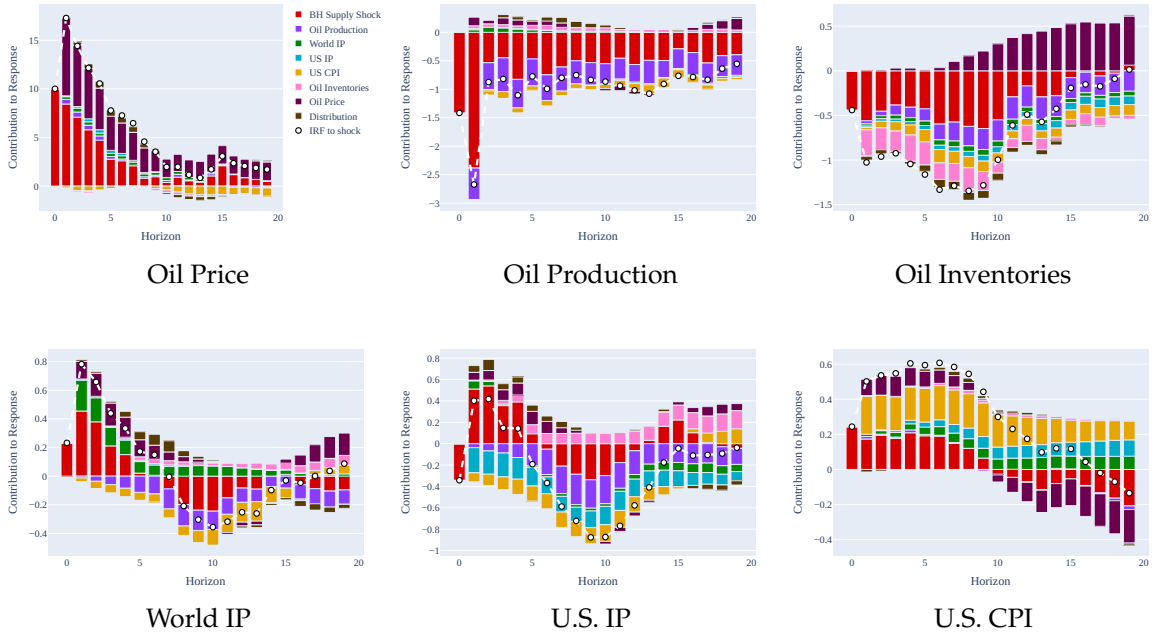
Before diving in on each specific panel, a bird's-eye view reveals which channels dominate. Under BH realized disruptions, red and purple — the shock itself and oil production — are the biggest contributors to the aggregate dynamics. Under KZ supply news, the picture shifts: dark purple and mostly yellow dominate, that is, oil price and CPI—similar to the FEVD picture. The decomposition makes the inflation *stickiness* of the news shock concrete: once the announcement has repriced both oil and consumer prices on impact, those two channels propagate themselves forward and account for the bulk of the aggregate response for the rest of the horizon, with the shock itself and oil production contributing essentially nothing beyond the first quarter. This self-propagating-price pattern is absent under BH, where the shock and production channels remain the dominant carriers throughout. I confirm Kilian (2009)'s conclusion that not all oil price shocks are alike: The two oil supply shocks differ in nature and thus motivate different macroeconomic models and different policy.

The oil inventories plot in Panel A and Panel B shows the distinguishing features of the KZ shock. On impact, the shock — whether BH or KZ — is responsible for the entire response, reflected by the full red bar at $h = 0$ with no other channels. This is mechanical (the shock is ordered first), so it can be seen as an implementation check. From $h = 1$, new colors appear: the shock at $h = 0$ activated channels at $h = 1$ that will now propagate until $h = 20$.

From this point, the key similarities and differences emerge. The shocks are similar in the

Figure 4: Channel Decomposition of Oil Supply Shocks

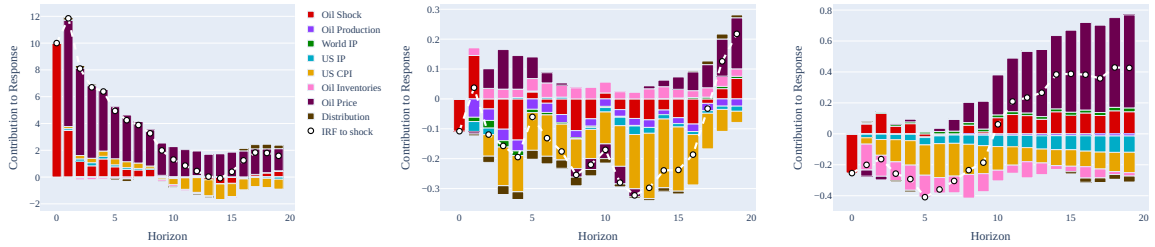
Panel A: Baumeister & Hamilton (2019) realized supply shock



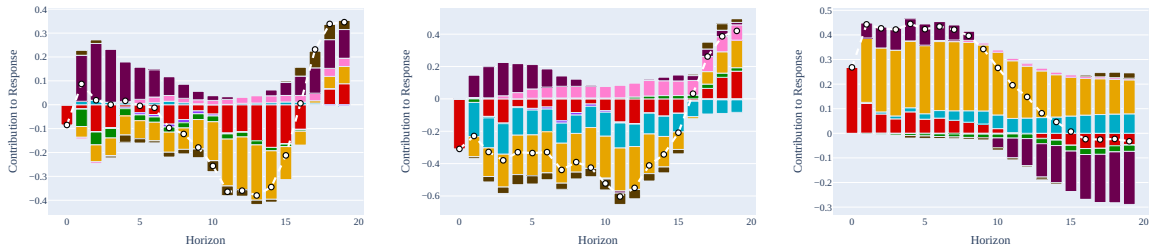
Oil Price | Oil Production | Oil Inventories

World IP | U.S. IP | U.S. CPI

Panel B: Känzig (2021) supply news shock



Oil Price | Oil Production | Oil Inventories



World IP | U.S. IP | U.S. CPI

Notes: Generalized impulse response decomposition (Dufour and Wang, 2024) of the macro impulse response into contributions from the oil supply shock, aggregate macro channels, and the distributional channel. Stacked bars: posterior median contribution of each channel at each horizon (color scheme identical across panels). Dashed white line and outlined white circles: median point estimates from Figure 1. Panel A: Baumeister and Hamilton (2019) realized supply shock; Panel B: Känzig (2021) oil supply news shock.

contribution of oil inventories and oil prices. In both cases, inventories exert downward pressure on themselves, which is in line with oil inventories being rigid, while the oil price channel becomes increasingly important after roughly two years and then propagates strongly. The magnitude of this price mechanism, however, differs across shocks. It is more muted under BH and more pronounced under KZ. Why? Under KZ, oil production (the IRF) is relatively unaffected, allowing supply to accommodate the increased demand associated with lower oil prices. Under BH, by contrast, production shortfalls are large and persistent, limiting this adjustment. This difference is visible in the decomposition: the production channel (purple) plays essentially no role in the inventories response under KZ, but is clearly present under BH.

The shocks' channel themselves (red) will differ in their implications for inventories, but also generally. The shocks contribution is larger and persistent under BH and much smaller under KZ, but operate in the same direction for all aggregates except one: oil inventories. The BH shock exerts a sizeable *downward* pressure on oil inventories throughout the horizon. In contrast, the KZ shock—beyond the impact response—generates *upward* pressure on inventories, particularly after two years. This pattern supports the interpretation of the KZ shock as directly driving inventory accumulation.

Examining the role of the U.S. distribution for the global oil market system reveals that it is a small one. The response of U.S. CPI is not driven by the distribution, and only a small fraction of U.S. IP is. With only two domestic variables — U.S. IP and U.S. CPI — there is also little scope for the U.S. distribution to matter, since the contribution of each channel will depend on the scope of relevant channels available. Appendix J extends the macro block to more domestic aggregates, where the distribution's contribution is larger (but still minor) and has the common feature of becoming active in the medium-run and as an amplifier.

6 The Distributional Effects of Oil Supply Shocks

This section provides insights on how the two different shocks affect inequality and household-level consumption responses. As before, I compare the Baumeister and Hamilton (2019) realized oil supply shocks with the Känzig (2021) oil supply news shock, to distinguish the two economic channels: sudden physical disruptions versus expectational shifts in oil supply.

6.1 Aggregate Measures of Inequality

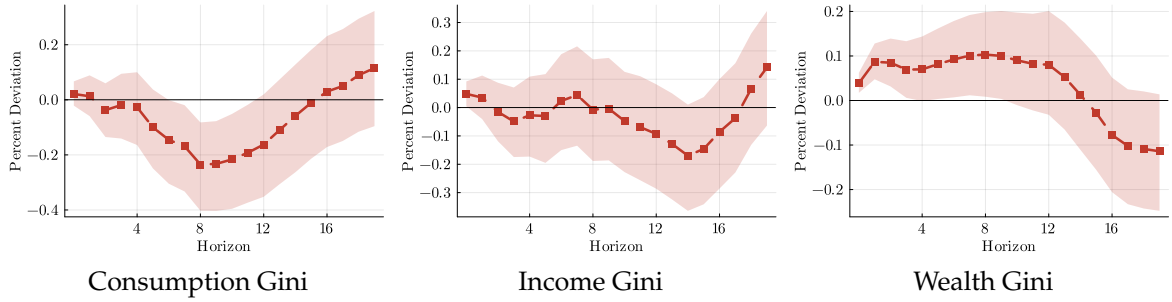
I examine how oil supply shocks affect inequality across the three dimensions of the joint distribution: consumption, income, and wealth. Figure 5 reports the accumulated stationary

impulse responses of the Gini coefficient.[22] Figures 6, 7, and 8 report shares held by different groups. Appendix K reports the log 90/10 and 90/50 quantile ratios, and the standard deviation of logs for each dimension.

Figure 5: Inequality responses to oil supply shocks

Panel A: Baumeister & Hamilton (2019) realized supply shock

Panel B: Känzig (2021) supply news shock



Notes: Impulse responses of the Gini coefficient for income, consumption, and wealth. Units: percentage point deviation from steady state. Solid lines: posterior median. Shaded areas: 68% credible sets.

Ginis. Under the BH shock, all three Gini coefficients increase, but with distinct short-run dynamics. Consumption and income inequality rise on impact, whereas wealth inequality responds with a delay. In the medium-run, after year two, all Ginis elevate and peak before waning off in the longer horizon. These patterns for income and wealth align with Lei, Ludwig, and Ma (2025), while the consumption response is novel. Under the KZ shock, a different story emerges: consumption and income inequality are rather flat on impact and more so for income. Whereas a physical disruption widens inequality in the medium-run, the anticipation channel appears to compress it at medium horizons before recovering toward the steady state by the end of the horizon.[23] Based on results I show later (Figure 10), I find that this is because

[22] For all three dimensions I use the bounded Gini normalization of Raffinetti, Siletti, and Vernizzi (2015), which replaces the classical denominator $2(N-1)\bar{x}$ with $2(N-1)(T^+ + T^-)$, where T^+ and T^- are the weighted totals of positive and absolute-negative values. This guarantees $G \in [0, 1]$ even when some observations are negative—relevant for wealth, where leveraged households can carry negative net worth—and collapses to the standard Gini when all values are non-negative.

[23] Although not explicitly documented, this decrease in income inequality can be found, to some extent, in Figure A.3 of Lei, Ludwig, and Ma (2025).

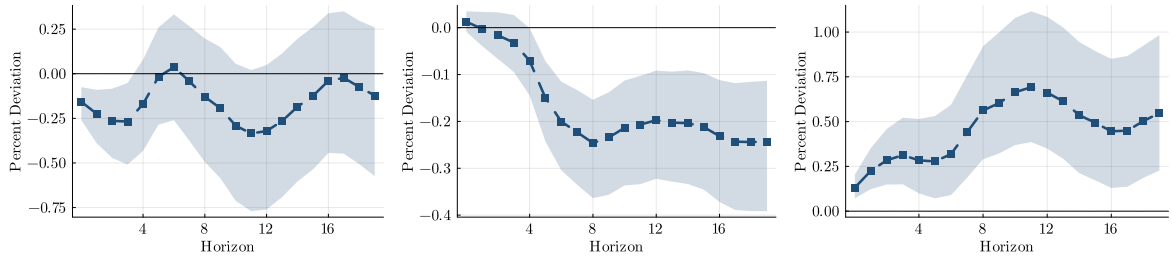
households in the upper half of the wealth distribution reduce consumption preemptively while transfer-dependent households at the bottom are initially shielded (Figure 43). This is consistent with households who participate in financial markets being more forward-looking and responsive to news. The wealth Gini is the only dimension where both shocks agree, increasing under both identifications; but still, their profiles are different. Under KZ, wealth inequality rises mildly and has a flat profile; under BH it is rather hump-shaped. Figures 6, 7, and 8 decompose the Gini dynamics into the underlying group shares and reveal what is moving them.

Consumption shares. Under the BH shock, the Bottom 50% share falls on impact and remains negative for three quarters, while the Top 10% rises by a comparable amount—consistent with the standard oil-redistribution channel (e.g., Lei, Ludwig, and Ma, 2025). After one year, however, the pattern shifts: the increase in the Gini is driven primarily by a persistent contraction in the Middle 50%-90% share, which bottoms out around quarter eight and dominates at medium horizons, regardless of the oil supply shock. Although the initial decline in the Bottom 50% is larger, its response is volatile, likely reflecting labor supply adjustments (intensive and extensive margins, Figure 37), transfer timing (Figure 43), and non-homothetic consumption (e.g., higher energy shares among low-income households; Edelstein and Kilian, 2009). Under the KZ shock, redistribution also occurs but with a delay and in the opposite direction: the Bottom 50% share rises relative to the Middle 50%-90%, rather than a contemporaneous increase in the Top 10% as under BH.

Income shares. Figure 7 shows the income-share decomposition. Under the realized BH shock, the Top 10% income share response is strictly monotonic, peaking at 1.5% at quarter 8, possibly coming from capital and corporate-profit gains (Figure 45). The shares of the Bottom 50% and Middle 50%-90% respond equally strongly, but in the opposite direction, with the Bottom 50% responding two times stronger than the Middle 50%-90%. Under the KZ shock, the pattern is different: changes in income inequality only materialize in the medium-run, remaining basically flat for most of the time horizon. An interesting finding however is that from quarter eight onward, the shape of the IRFs is similar—both shocks produce a hockey-shaped trajectory in each share, just at opposite levels (BH peaks positive, KZ troughs negative)—and it is mostly in the short-run that they meaningfully differ.

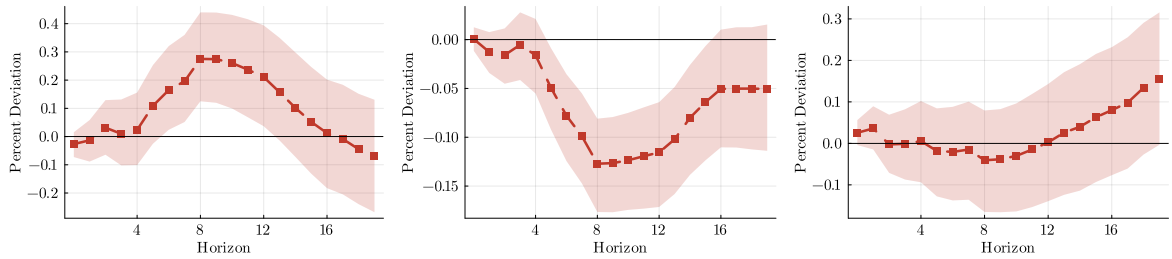
Figure 6: Consumption share responses to oil supply shocks

Panel A: Baumeister & Hamilton (2019) realized supply shock



Bottom 50% Middle 50%-90% Top 10%

Panel B: Känzig (2021) supply news shock

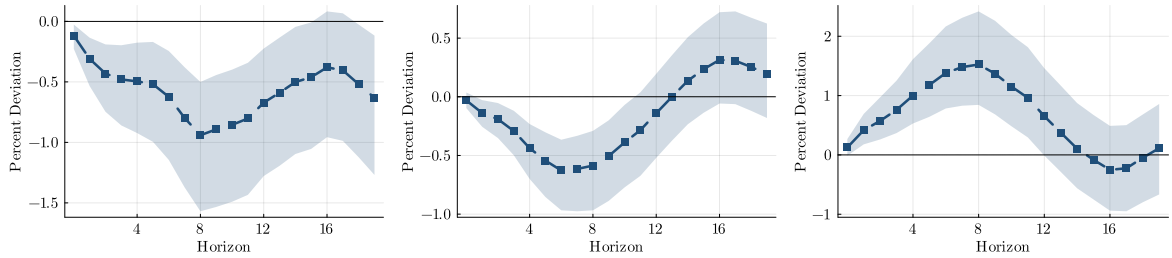


Bottom 50% Middle 50%-90% Top 10%

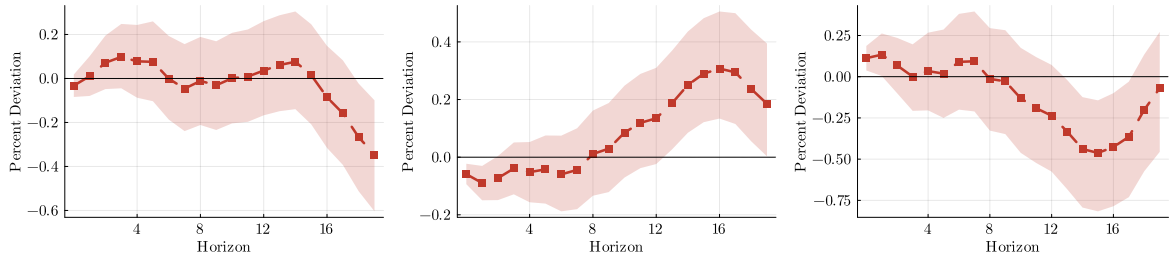
Notes: Impulse responses of the consumption share held by the Bottom 50% (0-50th percentile), Middle 50%-90% (50th-90th), and Top 10% (90th-100th) of the consumption distribution. Units: percentage point deviation from steady state. Solid lines: posterior median. Shaded areas: 68% credible sets.

Figure 7: Income share responses to oil supply shocks

Panel A: Baumeister & Hamilton (2019) realized supply shock



Panel B: Känzig (2021) supply news shock

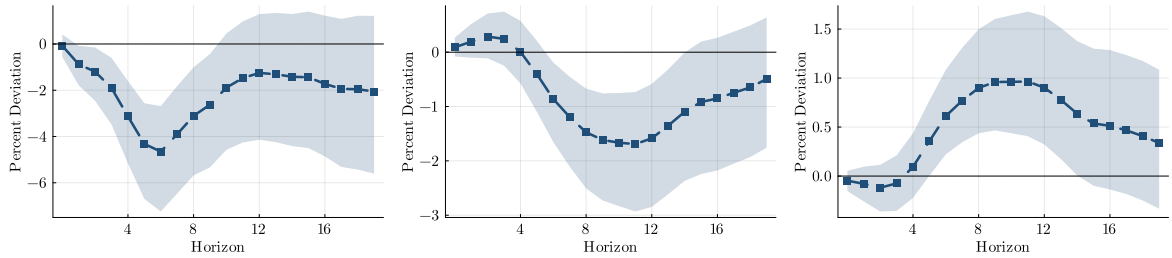


Notes: Impulse responses of the income share held by the Bottom 50%, Middle 50%-90%, and Top 10% of the income distribution. Units: percentage point deviation from steady state. Solid lines: posterior median. Shaded areas: 68% credible sets.

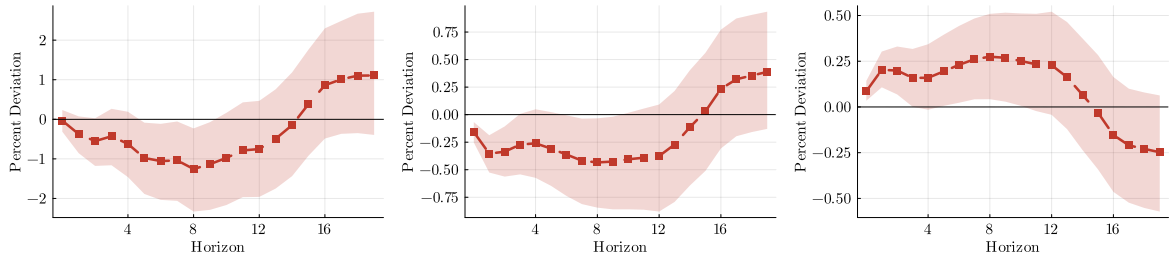
Wealth shares. Figure 8 shows the wealth-share decomposition. Under both identifications, I see a similar pattern: the Top 10% wealth share rises and the other groups fall. This is interesting because it suggests that the two shocks, in this dimension, possibly share the same underlying mechanism. Still, these responses will differ in their magnitudes and timing. Under BH, the Bottom 50% deplete their savings strongly within the first year, most likely attempting to curb the hike in energy prices. The Middle 50%-90% as well, but with a delay. The Top 10% also reacts with a delay, but ultimately peaks at 1%. All these effects are muted under a news shock and linear, suggesting that the news shock interacts with fewer channels, but these channels are persistent and correlated with the other variables in the system. Figure 4 shows precisely this.

Figure 8: Wealth share responses to oil supply shocks

Panel A: Baumeister & Hamilton (2019) realized supply shock



Panel B: Känzig (2021) supply news shock



Notes: Impulse responses of the wealth share held by the Bottom 50%, Middle 50%-90%, and Top 10% of the wealth distribution. Units: percentage point deviation from steady state. Solid lines: posterior median. Shaded areas: 68% credible sets.

6.2 Household-Level Responses

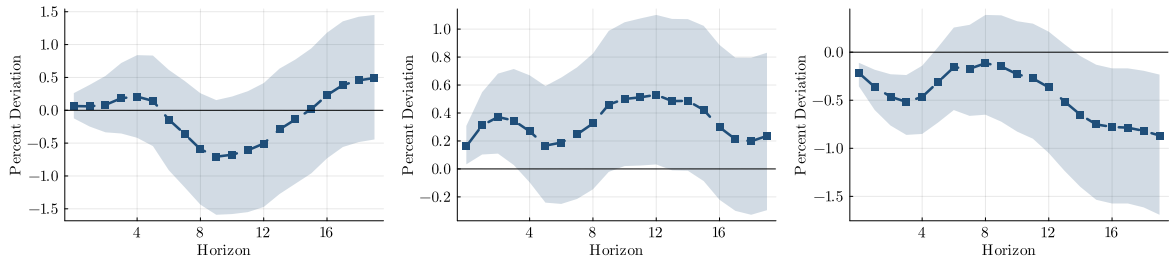
This section zooms in on consumption and provides insights on the sources of consumption inequality by showing household-level consumption impulse responses along the income and wealth distribution. This analysis uncovers heterogeneity that is hidden in the aggregate—offering a more nuanced and informative picture than the inequality measures alone. Figure 9 shows consumption by income and Figure 10 shows consumption by wealth, each for three

groups: Bottom 50%, Middle 50%-90%, and Top 10%. The section concludes with the full 5×5 income–wealth grid (Figures 11 and 12), which reports consumption responses for each (income, wealth) cell.

Consumption responses by income. Figure 9 reports the consumption impulse responses for households grouped into Bottom 50%, Middle 50%-90%, and Top 10% income groups.[24] Under the BH shock, the response of Bottom 50% income households is relatively flat the first five quarters before declining sharply to about -0.5% around quarter eight. Middle 50%-90% income households show a different, rather non-monotonic path, but generally increasing consumption over the time horizon. Top 10% income households display the opposite pattern of Middle 50%-90% income households: an immediate, but small decline that gradually zeros after the first year, but has a medium-run negative impact. Under the KZ news shock, there is again this flatness across all groups in the short-run, but responses then begin to mimic the BH shock.

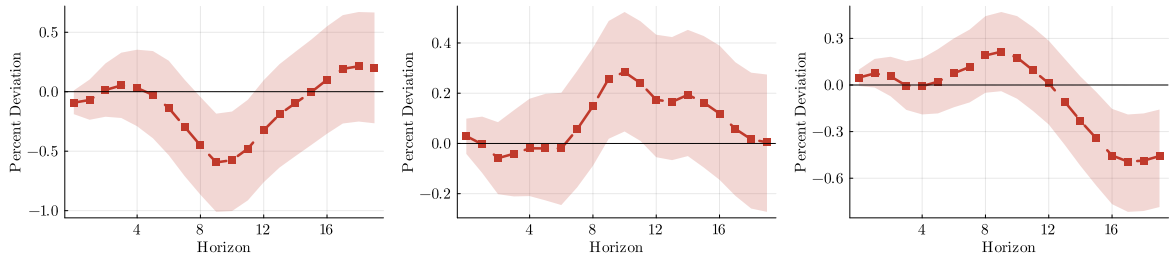
Figure 9: Consumption responses by position in the income distribution

Panel A: Baumeister & Hamilton (2019) realized supply shock



Bottom 50% Income | Middle 50%-90% Income | Top 10% Income

Panel B: Känzig (2021) supply news shock



Bottom 50% Income | Middle 50%-90% Income | Top 10% Income

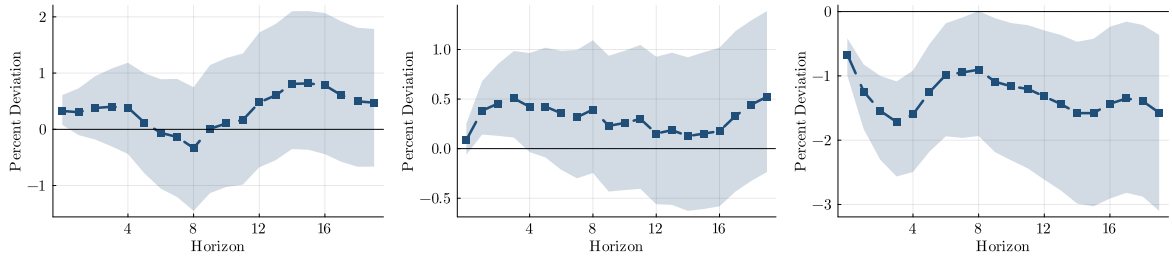
Notes: Impulse responses of mean consumption for households grouped into Bottom 50% (0–50%), Middle 50%-90% (50–90%), and Top 10% (90–100%) of the marginal income distribution. Solid lines: posterior median. Shaded areas: 68% credible sets. Horizon in quarters.

[24] Throughout this section I use five overlapping marginal cuts of the income (and analogously the wealth) distribution: Bottom 20% (0–20%), Bottom 50% (0–50%), Middle 50%-90% (50–90%), Top 20% (80–100%), and Top 10% (90–100%). The tail cuts (Bottom 20%, Top 10%) are nested inside the broader cuts (Bottom 50%, Top 20%) and capture the extremes of the marginal distribution; the remaining three span the bulk of the population. The body figures use the three broad cuts (Bottom 50%, Middle 50%-90%, Top 10%); the joint-distribution figures and the appendix tail cuts use all five.

Consumption responses by wealth. Figure 10 repeats the exercise along the wealth dimension. Under BH, Bottom 50% wealth households respond similar to Bottom 50% income households, but with larger uncertainty. The Middle 50%-90% group again increases consumption on impact, but uncertainty widens in the medium-run. The Top 10% wealth group decreases consumption by nearly 1% in the first year—a sharp result. These households do not return to steady state in the medium-run. Under KZ, Bottom 50% wealth households increase consumption in the short-run and in the medium-run. Middle 50%-90% wealth and Top 10% wealth households lose in the medium-run as in BH, but not in the short-run. This indicates the wealth channel is similar at the top, but different for the Bottom 90%. Appendix L extends both the income and wealth views to the tails of each margin (Bottom 20% and Top 10%), where the magnitudes are sharper.

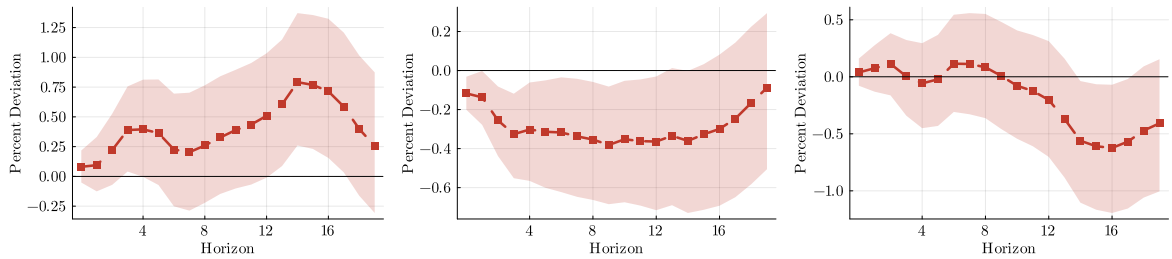
Figure 10: Consumption responses by position in the wealth distribution

Panel A: Baumeister & Hamilton (2019) realized supply shock



Bottom 50% Wealth | Middle 50%-90% Wealth | Top 10% Wealth

Panel B: Känzig (2021) supply news shock



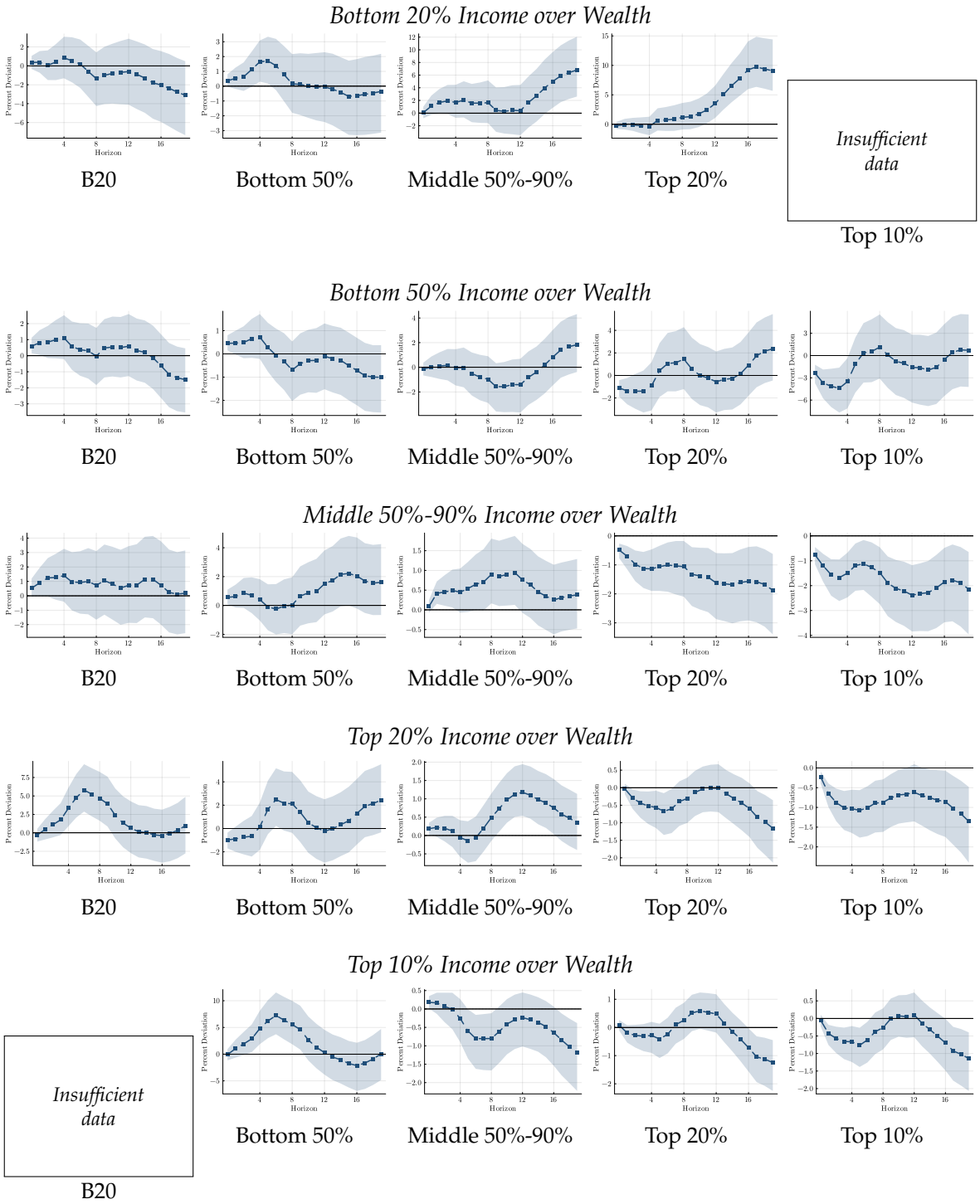
Bottom 50% Wealth | Middle 50%-90% Wealth | Top 10% Wealth

Notes: Impulse responses of mean consumption for households grouped into Bottom 50% (0–50%), Middle 50%-90% (50–90%), and Top 10% (90–100%) of the marginal wealth distribution. Solid lines: posterior median. Shaded areas: 68% credible sets. Horizon in quarters.

Consumption responses by income and wealth. Figures 11 and 12 report consumption impulse responses across the full 5×5 income–wealth grid for the BH and KZ shocks, respectively. Each row corresponds to an income group, similar to before, with two new groups: Bottom 20% and Top 20%; columns are the analogous wealth groups.[25]

[25] Two cells (Bottom 20% income/Top 10% wealth and Top 10% income/Bottom 20% wealth) are absent because the underlying micro data does not find a correlation between these groups.

Figure 11: Consumption responses by joint income–wealth cell — BH realized supply shock



Notes: Consumption impulse responses for households in each (Income, Wealth) cell of the joint distribution; Baumeister and Hamilton (2019) realized supply shock. Rows correspond to income groups (Bottom 20%, Bottom 50%, Middle 50%-90%, Top 20%, Top 10%; overlapping ranges, see footnote on first introduction); columns within each row correspond to the analogous wealth groups. Cells absent due to insufficient observations: Bottom 20% income/Top 10% wealth and Top 10% income/Bottom 20% wealth. Solid lines: posterior median. Shaded areas: 68% credible sets. Horizon in quarters.

Figure 12: Consumption responses by joint income–wealth cell — KZ oil supply news shock

Notes: Consumption impulse responses for households in each (Income, Wealth) cell of the joint distribution; Känzig (2021) oil supply news shock. Rows correspond to income groups (Bottom 20%, Bottom 50%, Middle 50%-90%, Top 20%, Top 10%; overlapping ranges, see footnote on first introduction); columns within each row correspond to the analogous wealth groups. Cells absent due to insufficient observations: Bottom 20% income/Top 10% wealth and Top 10% income/Bottom 20% wealth. Solid lines: posterior median. Shaded areas: 68% credible sets. Horizon in quarters.

Reading the BH grid (Figure 11) by blocks, four patterns stand out. (i) The *asset-rich, income-poor* cells in the upper-right 2×3 block (Bottom 20%, Bottom 50% income) $\times$ (Middle 50%-90%, Top 20%, Top 10% wealth) increase consumption in the medium run, often substantially: the wealth buffer at the bottom of the income distribution absorbs the recessionary impulse, and households in this block are the only ones that systematically *benefit* from the realized supply shock at intermediate horizons. (ii) The opposite block — the bottom right 3×3 (Middle 50%-90%, Top 20%, Top 10% income) $\times$ (Middle 50%-90%, Top 20%, Top 10% wealth) — contracts uniformly, with two exceptions: (Middle 50%-90% income, Middle 50%-90% wealth) and (Top 20% income, Middle 50%-90% wealth). (iii) The 2×2 of (Bottom 20%, Bottom 50% income) $\times$ (Bottom 20%, Bottom 50% wealth) display a small short-run rise followed by a medium-run decline — consistent with transfer-indexed nominal income protecting the bottom in the short run before the recessionary contraction binds—though the posterior contains zero. (iv) The (Middle 50%-90%, Top 20%, Top 10% income) $\times$ (Bottom 20%, Bottom 50% wealth) do not contract; they rise on impact and revert.

The KZ grid, Figure 12, preserves the upper-right "wealth-buffer wins" block (i) and block (iii), but block (ii) reads differently. Under KZ, Middle 50%-90% income households split: at Bottom 20% or Bottom 50% wealth they increase consumption gradually over the entire horizon, while at Middle 50%-90%, Top 20%, or Top 10% wealth they decrease in the short or medium run — a pattern consistent with asset-holders pricing the announcement and non-asset holders not. Top 20% income households at the bottom of the wealth distribution increase consumption in the medium run; at the top of the wealth distribution they rise marginally on impact before fading along zero. Top 10% income / Top 10% wealth households are the cleanest news responders: a stronger impact consumption increase consistent with capital-income gains as oil-related expectations reprice, slowly waning to zero. The contrast with BH is informative — realized supply contractions punish the asset-rich, income-rich; news shocks are priced in opposite directions on impact depending on wealth, with financial-market participants reacting first.

7 Conclusion

This paper has documented that oil supply shocks have first-order effects on the joint distribution of U.S. household income, consumption, and wealth, and that the distributional incidence depends qualitatively on the type of shock identified. Realized supply disruptions and supply news shocks move many of the same aggregate variables in similar directions but redistribute in opposite directions: the Baumeister and Hamilton (2019) realized shock raises consumption inequality through a cost-pull channel that punishes asset-rich, income-rich house-

holds after a one-year lag, while the Känzig (2021) news shock compresses consumption inequality on impact through an expectations channel that operates first on financial-market participants. Across robustness checks—alternative oil-price measures, sub-sample windows, lag lengths, financial-control sets, and posterior central tendencies—this wedge survives in sign and shape. The analysis thus shares the view of Kilian and Lewis (2011) that policy responses should depend on the underlying causes of oil price shocks.

Beyond the substantive findings, the paper's methodological contribution is the combination of the asymmetric conjugate Bayesian VAR of Chan (2022) with the within-model causal-mediation decomposition of Dufour and Wang (2024), applied to a system that embeds the joint income–consumption–wealth distribution of Bayer, Calderon, and Kuhn (2025)—a template that other applied work on shock transmission to heterogeneous households can adopt.

Three implications stand out. First, structural identification matters not only for the aggregate response, as Kilian (2009) and Känzig (2021) emphasize, but also for the redistributive incidence—different parts of the same OPEC decision (announcement versus realization) hit different parts of the distribution. Second, the joint distribution is not a passive add-on: dropping it from the VAR changes the aggregate IRFs themselves—most starkly, the monetary-policy response flips sign—so studying oil shocks at the aggregate level alone leads, in this dataset, to an omitted-variable problem. This is the empirical analogue of the demand-side-information argument of Kilian and Lewis (2011): heterogeneity in expenditure shares and household balance sheets carries demand-side variation that the small aggregate-only oil VARs of the literature systematically omit. Third, the wealth dimension is the one place where the two shocks agree directionally, suggesting that wealth inequality is shaped by oil shocks through balance-sheet revaluations that operate regardless of identification, while consumption and income inequality are shaped by the channel through which the shock reaches the household.

Several extensions come to mind that would complement the results presented here. First, the sufficient-statistics counterfactual framework of Caravello, McKay, and Wolf (2024) can be applied to the impulse responses to ask how the distributional incidence would look under alternative monetary regimes—a fixed-rate peg, strict inflation targeting, or a Taylor rule that internalizes distributional concerns—turning the present empirical wedge into a normative comparison of policy rules, building on Broer, Kramer, and Mitman (2025)'s application of the same framework to German labor-market microdata but with the joint income–consumption–wealth distribution as the object of incidence. Second, the mechanism documented here can perhaps be modeled in a heterogeneous-agent New Keynesian model with a realistic oil-market structure, replicating the flip inequality result I find, using also the distributional IRFs reported here as targets. Third, applying the framework to other supply-side

shocks should clarify whether the "physical-versus-informational" wedge is specific to oil or general to commodity markets where announcement effects coexist with realized supply changes.

A final caveat concerns non-linearities. My baseline is a linear Bayesian VAR, and the muted headline-inflation pass-through under the BH realized shock is a statement about the sample-average shock size. Baumeister et al. (2025) document substantial size and sign asymmetries in oil-shock transmission using Bayesian Additive Regression Trees, and earlier non-linear work (e.g., Hooker, 2002) shows that headline pass-through can re-emerge under unusually large or compounding episodes. A state-dependent or non-linear specification with a sufficiently large BH-type impulse—comparable to the 1973–74 or 1979–80 episodes—could plausibly find a more inflationary realized-supply response (and other responses) than the linear average I report.

References

[1] Juan Antolín-Díaz and Paolo Surico. "The Long-Run Effects of Government Spending". In: *American Economic Review* 115.7 (2025), pp. 2376–2413.

[2] Adrien Auclert. "Monetary policy and the redistribution channel". In: *American Economic Review* 109.6 (2019), pp. 2333–2367.

[3] Jushan Bai. "Estimating Cross-Section Common Stochastic Trends in Nonstationary Panel Data". In: *Journal of Econometrics* 122.1 (2004), pp. 137–183.

[4] Christiane Baumeister. "Measuring Market Expectations". In: *Handbook of Economic Expectations* (2023), pp. 413–442.

[5] Christiane Baumeister. *Discussion of" Local Projections or VARs? A Primer for Macroeconomists" by José Luis Montiel Olea, Mikkel Plagborg-Møller, Eric Qian, and Christian K*. Tech. rep. Wolf, Working paper, Prepared for the NBER Macroeconomics Annual, 2025.

[6] Christiane Baumeister and James D. Hamilton. "Structural Interpretation of Vector Autoregressions with Incomplete Identification: Revisiting the Role of Oil Supply and Demand Shocks". In: *American Economic Review* 109.5 (2019), pp. 1873–1910.

[7] Christiane Baumeister and Lutz Kilian. "Forty years of oil price fluctuations: Why the price of oil may still surprise us". In: *Journal of Economic Perspectives* 30.1 (2016), pp. 139–160.

[8] Christiane Baumeister et al. *Oil, Inflation Expectations, and Household Characteristics: A Nonlinear Heterogeneous Agent VAR Approach*. Tech. rep. University of Notre Dame, 2025.

[9] Christian Bayer, Benjamin Born, and Ralph Luetticke. "Shocks, frictions, and inequality in US business cycles". In: (2020).

[10] Christian Bayer, Luis Calderon, and Moritz Kuhn. *Distributional Dynamics*. Tech. rep. 2025.

[11] Paul Beaudry and Franck Portier. "Stock prices, news, and economic fluctuations". In: *American economic review* 96.4 (2006), pp. 1293–1307.

[12] Edmond Berisha et al. "Income inequality and oil resources: Panel evidence from the United States". In: *Energy Policy* 159 (2021), p. 112603.

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

[14] Tobias Broer, John Kramer, and Kurt Mitman. "The Distributional Effects of Oil Shocks". In: *IMF Economic Review* 73.3 (2025), pp. 851–889.

[15] Martin Bruns and Helmut Lütkepohl. "Comparing external and internal instruments for vector autoregressions". In: *Journal of Economic Dynamics and Control* 177 (2025), p. 105131.

[16] Dario Caldara and Edward Herbst. "Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs". In: *American Economic Journal: Macroeconomics* 11.1 (2019), 157–92.

[17] Guillermo A. Calvo. "Staggered prices in a utility-maximizing framework". In: *Journal of Monetary Economics* 12.3 (1983), pp. 383–398.

[18] Tomás E Caravello, Alisdair McKay, and Christian K Wolf. *Evaluating policy counterfactuals: A var-plus approach*. Tech. rep. National Bureau of Economic Research, 2024.

[19] Joshua C C Chan. "Asymmetric conjugate priors for large Bayesian VARs". In: *Quantitative Economics* 13.3 (2022), pp. 1145–1169.

[20] Minsu Chang, Xiaohong Chen, and Frank Schorfheide. "Heterogeneity and Aggregate Fluctuations". In: *Journal of Political Economy* 132.12 (2024), pp. 4021–4067.

[21] Minsu Chang and Frank Schorfheide. *On the Effects of Monetary Policy Shocks on Income and Consumption Heterogeneity*. Working Paper 32166. National Bureau of Economic Research, 2024.

[22] Gabriel Chodorow-Reich et al. "Homeownership and the Cost of Living". In: *Manuscript, Harvard University* (2025).

[23] Olivier Coibion and Yuriy Gorodnichenko. "Information rigidity and the expectations formation process: A simple framework and new facts". In: *American Economic Review* 105.8 (2015), pp. 2644–2678.

[24] Jeff D. Colgan. "The Emperor Has No Clothes: The Limits of OPEC in the Global Oil Market". In: *International Organization* 68.3 (2014), pp. 599–632.

[25] Ferre De Graeve and Andreas Westermark. *How long is long enough? Long-lag VARs*. Tech. rep. Sveriges Riksbank Working Paper Series, 2025.

[26] Riccardo Degasperi. *Identification of expectational shocks in the oil market using OPEC announcements*. University of Warwick, Department of Economics, 2025.

[27] Felipe N Del Canto et al. *Are inflationary shocks regressive? A feasible set approach*. Tech. rep. National Bureau of Economic Research, 2023.

[28] Thomas Doan, Robert Litterman, and Christopher Sims. "Forecasting and conditional projection using realistic prior distributions". In: *Econometric reviews* 3.1 (1984), pp. 1–100.

[29] Theo Drossidis, Haroon Mumtaz, and Angeliki Theophilopoulou. "The distributional effects of oil supply news shocks". In: *Economics Letters* 240.C (2024).

[30] Jean-Marie Dufour and Eric Renault. "Short run and long run causality in time series: theory". In: *Econometrica* (1998), pp. 1099–1125.

[31] Jean-Marie Dufour and Endong Wang. *Causal mechanism and mediation analysis for macroeconomics dynamics: a bridge of Granger and Sims causality*. Tech. rep. 2024.

[32] Paul Edelstein and Lutz Kilian. "How sensitive are consumer expenditures to retail energy prices?" In: *Journal of Monetary Economics* 56.6 (2009), pp. 766–779.

[33] Stephanie Ettmeier. *No Taxation Without Reallocation: The Distributional Effects of Tax Changes*. CRC TR 224 Discussion Paper Series crctr224_2023_436. University of Bonn and University of Mannheim, Germany, 2023.

[34] Simon Freyaldenhoven. "Factor models with local factors—determining the number of relevant factors". In: *Journal of Econometrics* 229.1 (2022), pp. 80–102.

[35] Yoshito Funashima. "Global Economic Activity Indexes Revisited". In: *Economics Letters* 193 (2020), p. 109269.

[36] Domenico Giannone, Michele Lenza, and Giorgio E Primiceri. "Prior selection for vector autoregressions". In: *Review of Economics and Statistics* 97.2 (2015), pp. 436–451.

[37] Domenico Giannone, Michele Lenza, and Giorgio E Primiceri. "Priors for the long run". In: *Journal of the American Statistical Association* 114.526 (2019), pp. 565–580.

[38] Simon Gilchrist and Egon Zakrajšek. "Credit spreads and business cycle fluctuations". In: *American economic review* 102.4 (2012), pp. 1692–1720.

[39] Oriol González-Casásus and Frank Schorfheide. *Misspecification-Robust Shrinkage and Selection for VAR Forecasts and IRFs*. Tech. rep. NBER Working Paper 33474, 2025.

[40] James D Hamilton. "Oil and the macroeconomy since World War II". In: *Journal of political economy* 91.2 (1983), pp. 228–248.

[41] James D Hamilton. *Causes and Consequences of the Oil Shock of 2007-08*. Tech. rep. National Bureau of Economic Research, 2009.

[42] James D Hamilton and Ana Maria Herrera. "Comment: oil shocks and aggregate macroeconomic behavior: the role of monetary policy". In: *Journal of Money, credit and Banking* (2004), pp. 265–286.

[43] Mark A Hooker. "Are oil shocks inflationary? Asymmetric and nonlinear specifications versus changes in regime". In: *Journal of money, credit and banking* (2002), pp. 540–561.

[44] Diego R Känzig. "The macroeconomic effects of oil supply news: Evidence from OPEC announcements". In: *American Economic Review* 111.4 (2021), pp. 1092–1125.

[45] Greg Kaplan, Benjamin Moll, and Giovanni L Violante. "Monetary policy according to HANK". In: *American Economic Review* 108.3 (2018), pp. 697–743.

[46] Robert K. Kaufmann. "The role of market fundamentals and speculation in recent price changes for crude oil". In: *Energy Policy* 39.1 (2011), pp. 105–115.

[47] Lutz Kilian. "Exogenous oil supply shocks: how big are they and how much do they matter for the US economy?" In: *The review of economics and statistics* 90.2 (2008), pp. 216–240.

[48] Lutz Kilian. "The economic effects of energy price shocks". In: *Journal of economic literature* 46.4 (2008), pp. 871–909.

[49] Lutz Kilian. "Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market". In: *American economic review* 99.3 (2009), pp. 1053–1069.

[50] Lutz Kilian. "Measuring global real economic activity: Do recent critiques hold up to scrutiny?" In: *Economics Letters* 178 (2019), pp. 106–110.

[51] Lutz Kilian and Logan T Lewis. "Does the Fed respond to oil price shocks?" In: *The Economic Journal* 121.555 (2011), pp. 1047–1072.

[52] Lutz Kilian and Daniel P. Murphy. "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil". In: *Journal of Applied Econometrics* 29.3 (2014), pp. 454–478.

[53] Lutz Kilian and Daniel P Murphy. "The role of inventories and speculative trading in the global market for crude oil". In: *Journal of Applied econometrics* 29.3 (2014), pp. 454–478.

[54] Lutz Kilian and Xiaoqing Zhou. "The econometrics of oil market VAR models". In: *Advances in Econometrics* 45 (2023), pp. 65–95.

[55] Marek Kolodzeij and Robert K. Kaufmann. "Oil demand shocks reconsidered: A cointegrated vector autoregression". In: *Energy Economics* 41 (2014), pp. 33–40.

[56] Xiaowen Lei, Julian F. Ludwig, and Xiaohan Ma. "Oil Prices and Inequality". In: *European Economic Review* (2025). Forthcoming.

[57] Michele Lenza and Ettore Savoia. *Do we need firm data to understand macroeconomic dynamics?* Working Paper Series 438. Sveriges Riksbank (Central Bank of Sweden), 2024.

[58] Robert B Litterman. "Techniques for forecasting with vector autoregressions". PhD thesis. Ph. D. thesis, University of Minnesota, 1980.

[59] Julian F. Ludwig. *Local Projections Are VAR Predictions of Increasing Order*. Tech. rep. SSRN Working Paper 4882149, 2024.

[60] James G. MacKinnon. "Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests". In: *Journal of Business & Economic Statistics* 12.2 (1994), pp. 167–176.

[61] Michael W. McCracken and Serena Ng. "FRED-QD: A Quarterly Database for Macroeconomic Research". In: *Review* 103.1 (2021), pp. 1–44.

[62] Alisdair McKay and Christian K Wolf. "What Can Time-Series Regressions Tell Us About Policy Counterfactuals?" In: *Econometrica* 91.5 (2023), pp. 1695–1725.

[63] Karel Mertens and Morten O Ravn. "The dynamic effects of personal and corporate income tax changes in the United States". In: *American economic review* 103.4 (2013), pp. 1212–1247.

[64] Silvia Miranda-Agrippino and Giovanni Ricco. "Identification with external instruments in structural VARs". In: *Journal of Monetary Economics* 135 (2023), pp. 1–19.

[65] Pavel Molchanov. "A statistical analysis of OPEC quota violations". In: *Economics* (2003), pp. 1–31.

[66] Lorenzo Mori and Gert Peersman. *Estimating the macroeconomic effects of oil supply news*. Tech. rep. CESifo Working Paper, 2024.

[67] Anton Nakov and Andrea Pescatori. "Oil and the great moderation". In: *The Economic Journal* 120.543 (2010), pp. 131–156.

[68] M Hashem Pesaran and Yongcheol Shin. "Generalized impulse response analysis in linear multivariate models". In: *Economics Letters* 58.1 (1998), pp. 17–29.

[69] Robert S. Pindyck. "The dynamics of commodity spot and futures markets: A primer". In: *The Energy Journal* 22.3 (2001), pp. 1–29.

[70] Mikkel Plagborg-Møller and Christian K Wolf. "Local projections and VARs estimate the same impulse responses". In: *Econometrica* 89.2 (2021), pp. 955–980.

[71] Emanuela Raffinetti, Elena Siletti, and Achille Vernizzi. "On the Gini coefficient normalization when attributes with negative values are considered". In: *Statistical Methods & Applications* 24.3 (2015), pp. 507–521.

[72] Ronald A Ratti and Joaquin L Vespignani. "OPEC and non-OPEC oil production and the global economy". In: *Energy Economics* 50 (2015), pp. 364–378.

[73] Christopher A Sims, James H Stock, and Mark W Watson. "Inference in linear time series models with some unit roots". In: *Econometrica: Journal of the Econometric Society* (1990), pp. 113–144.

[74] Christopher A Sims and Tao Zha. "Bayesian methods for dynamic multivariate models". In: *International Economic Review* (1998), pp. 949–968.

[75] Abe Sklar. "Vol. 8 of Fonctions de Répartition à n Dimensions et Leurs Marges, 229–231". In: *Paris: Publications de l'Institut de statistique de l'Université de Paris* (1959).

[76] James H Stock and Mark W Watson. *Disentangling the Channels of the 2007-2009 Recession*. Tech. rep. National Bureau of Economic Research, 2012.

[77] James H Stock and Mark W Watson. "Identification and estimation of dynamic causal effects in macroeconomics using external instruments". In: *The Economic Journal* 128.610 (2018), pp. 917–948.

[78] Rainer Storn and Kenneth Price. "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces". In: *Journal of global optimization* 11 (1997), pp. 341–359.

[79] Yan Tan and Utai Uprasen. "Asymmetric effects of oil price shocks on income inequality in ASEAN countries". In: *Energy Economics* 126.C (2023).

[80] Hiro Y Toda and Taku Yamamoto. "Statistical inference in vector autoregressions with possibly integrated processes". In: *Journal of econometrics* 66.1-2 (1995), pp. 225–250.

[81] Halbert White. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity". In: *Econometrica* 48.4 (1980), pp. 817–838.

[82] Holbrook Working. "The Theory of Price of Storage". In: *American Economic Review* 39.6 (1949), pp. 1254–1262.

A Granger-Causality Tests

Mori and Peersman (2024) document that structural oil-market shocks identified in standard monthly SVARs are predictable by financial variables, in violation of the lead-lag exogeneity required for valid identification. I replicate their predictability test at quarterly frequency for both shocks I use. (The companion "recoverability" test on the oil-price residual is mechanically satisfied in my specification, since the shock instrument enters the VAR endogenously and is therefore a linear combination of the system's innovations by construction; it is informative only for the external/proxy-SVAR reading, which I do not adopt.)

Granger-causality test. For each candidate predictor x_t (a financial variable or common factor), I run the auxiliary regression

$$z_t = \alpha + \gamma_1 x_{t-1} + \cdots + \gamma_L x_{t-L} + e_t, \tag{27}$$

where z_t is the shock instrument and $L \in \{2, 4\}$. I test $H_0 \colon \gamma_1 = \cdots = \gamma_L = 0$ with a heteroskedasticity-robust Wald statistic using the HC0 sandwich estimator of White (1980) and report the p-value from an $F(L, T - L - 1)$ reference distribution. The $L = 2$ choice corresponds to a six-month horizon, matching the lag structure used by Mori and Peersman (2024) at monthly frequency (six lags); $L = 4$ extends to one year as a robustness check.

Results. Table 2 reports the predictability test for both the Känzig (2021) oil supply news shock and the Baumeister and Hamilton (2019) supply shock. No individual financial variable robustly predicts either shock instrument across both lag specifications, supporting the exogeneity of the instruments at quarterly frequency. For the KZ shock, the one exception is the excess bond premium, which is correlated with Factor 5; for BH, the term spread and Factor 6 are the exceptions. The Granger-causality patterns documented by Mori and Peersman (2024) at monthly frequency are therefore substantially attenuated at quarterly frequency, with the

residual predictive content concentrated in EBP (KZ) and the term spread (BH)—the channels that the `mp_ebp` robustness specification controls for directly.

Information overlap between the QD factor block and financial variables. A natural question is whether the baseline robustness specification `paper_kanzig_kqd_smfdd`, which augments the baseline VAR with the six McCracken and Ng (2021) levels factors $\boldsymbol{F}_t^{\text{lev}} = (qdl_{f1}, \ldots, qdl_{f6})$, implicitly addresses the financial-variable predictability documented in Table 2. To assess this I compute, for each financial variable x_t, the in-sample R^2 of a linear projection on the six levels factors:

$$x_t \;=\; \alpha \;+\; \boldsymbol{\beta}' \boldsymbol{F}_t^{\text{lev}} \;+\; \eta_t. \tag{28}$$

Table 1 reports these R^2 values over the 1975Q1–2024Q3 sample.

Table 1: R^2 from regressing each financial variable on the six McCracken and Ng (2021) levels factors

Financial variable	R^2
Federal funds rate	0.986
3M T-bill	0.968
1Y interest rate	0.963
10Y interest rate	0.952
BAA corporate yield	0.949
S&P 500	0.924
Term spread	0.788
USD exchange rate	0.677
Excess bond premium	0.220
VIX	0.138

Note: OLS R^2 from regressing each financial variable on a constant and the six FRED-QD levels factors $qdl_{f1}, \ldots, qdl_{f6}$ used in `paper_kanzig_kqd_smfdd`. Sample 1975Q1–2024Q3, $n = 199$ for all rows except S&P 500 ($n = 193$). The factor block is constructed with the variable-overlap exclusions of Section 4, so the regressors do not contain the dependent variables themselves.

The six levels factors span 92–99 % of the variation in interest rates, the S&P 500, and the BAA corporate yield, and roughly 70–80 % of the term spread and USD exchange rate. The two financial variables that the factor block does *not* span are the excess bond premium ($R^2 = 0.22$) and the VIX ($R^2 = 0.14$): both reflect higher-frequency credit-stress and volatility innovations that are orthogonal to the slow-moving levels factor space. Two implications follow.

First, including $\boldsymbol{F}_t^{\text{lev}}$ in the VAR implicitly controls for the financial-predictor block of Table 2 along the rate-and-equity dimension. The Granger-causality concerns of Mori and Peersman (2024) for monthly SVARs—which manifest in my quarterly results primarily for short and long Treasury yields, the federal funds rate, the term spread, and corporate yields—are addressed by the kqd specification through factor-block conditioning, without the need to add each financial variable as a separate VAR observable.

Second, the kqd and `mp_ebp` specifications are complementary rather than substitutes. The QD factor block absorbs the slow-moving level and term-structure information; the Mori and Peersman (2024) financial block adds the credit-stress (EBP) and volatility (VIX) dimensions that levels factors do not span. I report both specifications as robustness checks: `paper_kanzig_kqd_smfdd` guards against omission of the rate/equity portion of the financial information set, while `paper_kanzig_mp_ebp` guards against omission of the residual credit and volatility components. Impulse responses are stable across both specifications, indicating that the baseline conclusions are not driven by either component of the financial information set in isolation.

Table 2: Granger-causality tests — Shock instruments

	Känzig (2021)		Baumeister & Hamilton (2019)	
	VAR(2)	VAR(4)	VAR(2)	VAR(4)
Common factors (all)	**0.00*****	**0.00*****	**0.00*****	**0.00*****
F_1	1.00	0.67	0.86	0.96
F_2	0.58	0.88	0.97	0.89
F_3	0.92	0.94	**0.03****	**0.01*****
F_4	0.32	0.34	0.31	0.40
F_5	0.13	0.31	0.59	0.08*
F_6	0.31	0.52	0.10	0.12
F_7	0.53	0.61	0.35	0.38
Dist. factors (all)	0.06*	**0.04****	**0.02****	**0.00*****
D_1	0.44	0.84	0.41	0.16
D_2	0.39	**0.00*****	0.58	0.16
D_3	0.16	0.40	**0.02****	**0.02****
D_4	0.45	0.78	0.52	0.90
D_5	0.07*	**0.02****	0.06*	**0.02****
D_6	0.49	0.54	0.73	0.82
D_7	**0.05****	0.29	0.07*	**0.01****
D_8	0.10	0.41	0.46	0.33
S&P 500	0.95	0.68	0.12	0.34
VIX	0.45	0.77	0.83	0.82
1Y interest rate	0.83	0.70	0.75	0.07*
10Y interest rate	0.88	0.33	0.35	0.16
3M T-bill	0.64	0.71	0.81	**0.03****
BAA corporate yield	0.39	0.47	0.26	0.10
Federal funds rate	0.57	0.36	0.80	**0.02****
Excess bond premium	0.26	0.09*	0.25	0.33
Term spread	0.17	0.22	**0.01****	**0.01****
USD exchange rate	0.65	0.70	0.48	0.76

Note: p-values from robust (HC0) F-tests of the null that lagged variables do not Granger-cause the shock instrument. Since the instrument is external to the VAR, results are invariant to the VAR information set; only the baseline specification is reported. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. Sample: 1975Q1–2023Q1. Common factors are the first seven principal components of the McCracken and Ng (2021) quarterly dataset. Distributional factors are the eight smoothed states of the Bayer–Calderon–Kuhn (2025) state-space model on PSID income, consumption, and copula coefficients. Term spread is the 10-year minus 3-month Treasury yield.

B Data

This appendix documents the macroeconomic, financial, and oil-market variables used in the baseline VAR and across the robustness specifications. All series are at quarterly frequency; monthly source data is aggregated appropriately. The full set is bundled in the canonical macro file `oil_macro_all.csv`, which is constructed from three sources: the Känzig (2021) replication archive (oil-market block), McCracken and Ng (2021) FRED-QD (macroeconomic and financial), and FRED monthly downloads for series not in FRED-QD (selected CPI components and the Brent futures price). Distributional variables and their reconstruction are documented in Appendix C.

Table 3: Macroeconomic, financial, and oil-market variables

Variable	Description	Source	Transform	Used in
Oil market block — baseline				
`poil`	Real WTI spot price	Känzig (2021)	log levels	all baseline configs
`oilprod`	Global crude oil production	Känzig (2021)	log levels	all baseline configs
`worldip`	OECD + 6 NME industrial production	Känzig (2021)	log levels	all baseline configs
`usip`	U.S. industrial production	FRED-QD (`INDPRO`)	log levels	all baseline configs
`log_cpi`	U.S. CPI, all items	FRED-QD (`CPIAUCSL`)	log levels	all baseline configs
`oilstocksM_BH`	OECD crude inventories, BH cumulated	Baumeister and Hamilton (2019) replication archive (extended via EIA inputs through 2024Q2; anchored at 1975Q1)	log levels	baseline VAR
Oil-market robustness				
`poil_rac`	Refiners' acquisition cost (imported), real	EIA	log levels	`paper_*_rac`
`poil_rac_comp`	Refiners' acquisition cost (composite), real	EIA	log levels	sensitivity (`poil_rac*`)
`poil_rac_dom`	Refiners' acquisition cost (domestic), real	EIA	log levels	sensitivity (`poil_rac*`)
`poil_brent`	Real Brent (deflated by CPI)	EIA `MCOILBRENTEU`	log levels (1987Q2+)	`paper_*_brent`
`oilstocksM`	OECD crude oil inventories, noisy series	Känzig (2021) (Kilian–Murphy 2014 construction)	log levels	`paper_*_oldinv`

Continued on next page

Table 3 – continued from previous page

Variable	Description	Source	Transform	Used in
`oilstocksL`	OECD crude inventories, level (no log)	derived from `oilstocksM`	raw level	sensitivity
`kilian_rea`	Kilian Index of Global Real Economic Activity (dry-cargo shipping rate)	Kilian (2009), Dallas Fed update	quarterly mean of monthly (already % deviation from trend)	`paper_*_rea`
`oilprod_opec`	OPEC crude (incl. lease condensate) production, Kanzig-anchored via EIA share	EIA INTL (`productId=57`, X-12-ARIMA SA) × Känzig (2021) world	log levels (1973Q1+, sample limited by `oilprod`)	`paper_*_opec`
`oilprod_nopec`	Non-OPEC crude production (residual: World − OPEC, identity preserved)	same as above	log levels	`paper_*_opec`
Real activity (FRED-QD)				
`gdp`	Real GDP	`GDPC1`	log levels	medium block
`disp_income`	Real disposable income	`DPIC96`	log levels	FAVAR
`consumption`	Real PCE	`PCECC96`	log levels	FAVAR
`investment`	Real gross private domestic invest.	`GPDIC1`	log levels	FAVAR
`pce_durables`	Real PCE durables	`PCDGx`	log levels	FAVAR
`pce_nondurables`	Real PCE nondurables	`PCNDx`	log levels	FAVAR
`pce_services`	Real PCE services	`PCESVx`	log levels	FAVAR
`unrate`	Civilian unemployment rate	`UNRATE`	level (%)	medium block, FAVAR
`payems`	Total nonfarm payrolls	`PAYEMS`	log levels	FAVAR
`manemp`	Manufacturing employment	`MANEMP`	log levels	FAVAR
`ahetpi`	Avg. hourly earnings, prod. workers	`AHETPIx`	log levels	FAVAR
`weekly_hours`	Avg. weekly hours, prod. workers	`CES0600000007`	level	FAVAR
`sentiment`	Michigan consumer sentiment	`UMCSENTx`	log levels	FAVAR
Prices (FRED-QD + FRED monthly)				
`cpi_core`	Core PCE price index	`PCEPILFE`	log levels	medium block
`cpi_oer`	CPI owners' equivalent rent	`CUSR0000SEHC`	log levels	FAVAR
`cpi_food`	CPI food	FRED `CPIUFDSL`	log levels	FAVAR

Continued on next page

Table 3 – continued from previous page

Variable	Description	Source	Transform	Used in
cpi_energy	CPI energy	FRED CPIENGSL	log levels	FAVAR
cpi_rent	CPI rent of primary residence	FRED CUSR0000SEHA	log levels	FAVAR
cpi_transp	CPI transportation	CPITRNSL	log levels	FAVAR
ppi	PPI all commodities	PPIACO	log levels	FAVAR
imp_prices	Import price index, BLS IPP	FRED IR, spliced pre-1982Q3 with BEA B021RG3Q086SBEA	log levels	medium block, FAVAR
Financial (FRED-QD)				
sp500	S&P 500 index	WolfAggs (CRSP)	log levels	MP+EBP, FAVAR
vix	CBOE VIX	VIXCLSx	log levels	MP+EBP, FAVAR
gs1	1-year Treasury yield	GS1	level (%)	MP+EBP, FAVAR
gs10	10-year Treasury yield	GS10	level (%)	FAVAR
tb3ms	3-month T-bill yield	TB3MS	level (%)	FAVAR
baa	BAA corporate yield	BAA	level (%)	FAVAR
mortgage30	30-year mortgage rate	MORTGAGE30US	level (%)	FAVAR
fedfunds	Federal funds rate	FEDFUNDS	level (%)	FAVAR
term_spread	10Y minus 3M Treasury spread	GS10TB3Mx	level (%)	FAVAR
ebp	Excess bond premium	Gilchrist and Zakrajšek (2012)	level	MP+EBP
exrate	Trade-weighted USD index	TWEXAFEGSMTHx	log levels	FAVAR
corp_profits	Corporate profits before tax	B020RE1Q156NBEA	level	FAVAR
house_prices	Case-Shiller 20-city HPI	SPCS20RSA	log levels	FAVAR
Macro aggregates (WolfAggs)				
inv	Real gross private domestic investment	WolfAggs (replication archive of McKay and Wolf, 2023, based on FRED GPDIC1)	log levels	extended-macro VAR
cons	Real personal consumption expenditures	WolfAggs (FRED PCECC96)	log levels	extended-macro VAR
wages	Real wages, production / non-supervisory workers	WolfAggs (FRED AHETPIx)	log levels	extended-macro VAR
unemp	Civilian unemployment rate	WolfAggs (FRED UNRATE)	level (%)	extended-macro VAR
infl	PCE-deflator inflation, q-on-q	WolfAggs (FRED PCECTPI)	level (%)	extended-macro VAR

Continued on next page

Table 3 – continued from previous page

Variable	Description	Source	Transform	Used in
`ffr`	Federal funds rate	WolfAggs (FRED `FEDFUNDS`)	level (%)	extended-macro VAR
Household credit (FRED-QD + FRED monthly)				
`consumer_credit`	Total consumer credit outstanding, real	FRED-QD `TOTALSLx`	log levels	extended-macro VAR, FAVAR
`consumer_loans`	Consumer loans at all commercial banks, real	FRED-QD `CONSUMERx`	log levels	FAVAR
`cc_debt`	Revolving consumer credit, real (≈ credit-card debt)	FRED-QD `REVOLSLx`	log levels	FAVAR
`nonrev_credit`	Non-revolving consumer credit, real (auto + student)	FRED-QD `NONREVSLx`	log levels	FAVAR
`dlq_consumer`	Delinquency rate on consumer loans	FRED `DRCLACBS`	level (%)	FAVAR
`dlq_cc`	Delinquency rate on credit-card loans	FRED `DRCCLACBS`	level (%)	FAVAR
`dlq_mortgage`	Delinquency rate on single-family residential mortgages	FRED `DRSFRMACBS`	level (%)	FAVAR
`chgoff_consumer`	Charge-off rate on consumer loans	FRED `CORCACBS`	level (%)	FAVAR
`chgoff_cc`	Charge-off rate on credit-card loans	FRED `CORCCACBS`	level (%)	FAVAR
Income composition (NIPA)				
`transfers`	Personal current transfer receipts (nominal $bn)	FRED `A063RC1`	log levels	FAVAR
`income_ex_transfers`	Real personal income excluding transfer receipts	FRED `W875RX1`	log levels	FAVAR
Inflation expectations				
`infl_exp_mich`	U. Michigan 1-year expected inflation (consumer survey)	FRED `MICH`	level (%)	FAVAR
`infl_exp_spf_cpi1y`	SPF median 1-year CPI inflation forecast (forecasters)	Philadelphia Fed SPF (`INFCPI1YR`)	level (%)	FAVAR
Common factors (PCA on FRED-QD)				
$F^{\text{lev}}_{1..6}$	Levels common factors (Bai 2004)	PCA on log-levels of FRED-QD ($n = 113 - 12$, post-exclusion); 6 factors	—	`paper_*_kqd_smfdd`
Shock instruments				

Continued on next page

Table 3 – continued from previous page

Variable	Description	Source	Transform	Used in
`oil_supply_news`	OPEC announcement futures surprise	Känzig (2021)	quarterly sum	baseline (KZ)
`bh_supply_shock`	Realized structural supply shock	Baumeister and Hamilton (2019)	—	baseline (B&H)
`paveldje_shock`	Demand-side news shock	Degasperi (2025)	—	inactive (commented)

Notes: All series at quarterly frequency, 1975Q1–2024Q3 unless otherwise noted. Monthly source data is aggregated by quarterly mean. "Log levels" means $\log(x_t)$ used directly without further differencing or detrending; the Bayesian framework handles mixed orders of integration without pre-testing (Section 4). The FRED-QD common factors are extracted with the variable-overlap exclusion list (`OILPRICEx`, `INDPRO`, `CPIAUCSL`, `CPILFESL`, `FEDFUNDS`, `GS1`, `GS10`, `VIXCLSx`, `S&P 500`, `UNRATE`, `GDPC1`, `GPDIC1`, `PCECC96`, `AHETPIx`, `PAYEMS`) so that any series used elsewhere in the VAR as an explicit observable is removed from the PCA input panel before extraction. Brent (`poil_brent`) and the BH-cumulated inventory series (`oilstocksM_BH`) are constructed in the build pipeline; details on the latter are in Appendix E. The Kilian REA (`kilian_rea`) is already in percent-deviation-from-trend units and enters the VAR without further log-transform.

C Reconstruction of Distributional Objects

This appendix details how impulse responses estimated in factor space are mapped back to distributional objects—quantile functions, copula densities, and household-level moments—following the methodology of Bayer, Calderon, and Kuhn (2025).

C.1 From Factor IRFs to Coefficient IRFs

The VAR is estimated on a small number of distributional factors $\boldsymbol{F}_{D,t} \in \mathbb{R}^K$, with $K = 8$ in the baseline specification. Let $\hat{\boldsymbol{\theta}}_t \in \mathbb{R}^N$ denote the raw Legendre polynomial coefficient vector, with $N = (O+1)^d + d \cdot (O+1)$ collecting copula and marginal quantile-function coefficients (cf. Section 3). The factors are obtained via principal components analysis on the *standardized* coefficients $\boldsymbol{\Sigma}_\theta^{-1/2}\big(\hat{\boldsymbol{\theta}}_t - \bar{\boldsymbol{\theta}}\big)$, so the implied factor model for the raw coefficients is

$$\hat{\boldsymbol{\theta}}_t \;=\; \bar{\boldsymbol{\theta}} \;+\; \boldsymbol{\Sigma}_\theta^{1/2}\big(\boldsymbol{\Lambda}\,\boldsymbol{F}_{D,t} + \boldsymbol{\gamma}\,\boldsymbol{f}_t\big), \tag{29}$$

where $\bar{\boldsymbol{\theta}}$ is the sample mean, $\boldsymbol{\Sigma}_\theta^{1/2}$ is the diagonal matrix of per-coefficient standard deviations, $\boldsymbol{\Lambda} \in \mathbb{R}^{N \times K}$ is the factor loading matrix, and $\boldsymbol{f}_t$ captures idiosyncratic variation orthogonal to the common factors.

Mixed-frequency measurement and the four-quarter projection. A subtle point governs the loading used for reconstruction. Bayer, Calderon, and Kuhn (2025) estimate the factor model in a mixed-frequency state-space framework: the latent factors $\boldsymbol{F}_t$ live at quarterly frequency, but the PSID coefficient observations are annual averages. The measurement equation for the PSID dataset takes the form

$$\tilde{\boldsymbol{\theta}}_t^{\text{PSID}} \;=\; \frac{1}{4}\,\boldsymbol{\Lambda}\left(\boldsymbol{F}_t + \boldsymbol{F}_{t-1} + \boldsymbol{F}_{t-2} + \boldsymbol{F}_{t-3}\right) \;+\; \boldsymbol{\nu}_t^{\text{PSID}}, \tag{30}$$

where $\boldsymbol{\nu}_t^{\text{PSID}}$ captures sampling and operationalisation noise. Equation (30) is the source of a reconstruction issue when the smoothed factors enter the VAR as observables. Applying $\boldsymbol{\Lambda}$ alone to a quarterly factor IRF produces an over-attenuated coefficient response, because each quarter of the IRF activates a single factor rather than the four-quarter sum the loadings $\boldsymbol{\Lambda}$ were estimated to fit. To recover the quarterly-frequency loading that maps a single quarter of factor variation into its associated coefficient impact, I fit by ordinary least squares

$$\hat{\boldsymbol{\Lambda}}_{4q} \;=\; \arg\min_{\boldsymbol{\Lambda}} \sum_t \left\| \hat{\boldsymbol{\theta}}_{t|T}^{\text{PSID}} \;-\; \boldsymbol{\Lambda}\,\tfrac{1}{4}\big(\hat{\boldsymbol{F}}_{t|T} + \hat{\boldsymbol{F}}_{t-1|T} + \hat{\boldsymbol{F}}_{t-2|T} + \hat{\boldsymbol{F}}_{t-3|T}\big) \right\|^2, \tag{31}$$

where $\hat{\boldsymbol{\theta}}_{t|T}^{\text{PSID}}$ and $\hat{\boldsymbol{F}}_{t|T}$ are the BCK smoother's posterior-mean reconstructions of the PSID coefficient and the latent quarterly factor at quarter t, respectively. The four-quarter rolling sum on the right-hand side enforces the BCK measurement equation (30); OLS thus recovers the loading $\hat{\boldsymbol{\Lambda}}_{4q}$ that closes the measurement-equation residual.

Mapping factor IRFs to coefficient IRFs. Quarterly factor responses $\boldsymbol{\Psi}_h^F$ are mapped to raw-coefficient responses by

$$\Delta\hat{\boldsymbol{\theta}}_h \;=\; \boldsymbol{\Sigma}_\theta^{1/2}\,\hat{\boldsymbol{\Lambda}}_{4q}\,\boldsymbol{\Psi}_h^F, \tag{32}$$

which undoes the per-coefficient standardization $\boldsymbol{\Sigma}_\theta^{1/2}$ from the PCA. The result $\Delta\hat{\boldsymbol{\theta}}_h$ is the IRF of the polynomial coefficients in raw units; the steady-state mean is not reintroduced.

Reconstructing nonlinear functionals. To evaluate distributional functionals ϕ (quantile values, copula densities, household-group means, inequality indices) at horizon h, the reconstruction is non-linear in the coefficient vector and therefore requires the level coefficient

$$\hat{\boldsymbol{\theta}}_h \;=\; \bar{\boldsymbol{\theta}} + \Delta\hat{\boldsymbol{\theta}}_h, \tag{33}$$

where $\bar{\boldsymbol{\theta}}$ is the time-invariant steady-state coefficient (including any trend component removed prior to estimation). The reconstructed objects are evaluated via (35)–(36) below at the level

$\hat{\boldsymbol{\theta}}_h$, and the IRF of any functional ϕ is reported as

$$\Psi_h^{\phi} \;=\; \phi(\hat{\boldsymbol{\theta}}_h) \;-\; \phi(\bar{\boldsymbol{\theta}}). \tag{34}$$

For functionals that are linear in $\boldsymbol{\theta}$ (the marginal quantile values evaluated pointwise via (35)), this reduces to direct evaluation at $\Delta\hat{\boldsymbol{\theta}}_h$. For copula-based functionals (household-group means, conditional moments, Ginis), the level coefficient must be used because monotonicity and non-negativity constraints couple the reconstruction nonlinearly across coefficients.

C.2 Reconstructing Quantile Functions and Copula Densities

The coefficient vector $\hat{\theta}_h$ partitions into copula coefficients $\kappa_h \in \mathbb{R}^{(O+1)^d}$ and marginal quantile-function coefficients $\xi_h^m \in \mathbb{R}^{O+1}$ for each dimension $m \in \{\text{consumption, income, wealth}\}$. These reconstruct the distributional objects via the Legendre polynomial basis $\{Q_o\}_{o=0}^{O}$:

Marginal quantile functions. For dimension m evaluated on a grid $u \in [0,1]$:

$$\hat{\Xi}_{m,h}^{-1}(u) = \sum_{o=0}^{O} \xi_{o,h}^m \, Q_o(u). \tag{35}$$

In practice, I evaluate this on a grid of $G = 20$ equally spaced quantile points (vigintiles), giving the income, consumption, or wealth level at each vigintile of the distribution.

Copula density. The joint dependence structure is reconstructed as:

$$d\hat{C}_h(u_1,u_2,u_3) = \sum_{o_1,o_2,o_3=0}^{O} \kappa_{(o_1,o_2,o_3),h} \prod_{m=1}^{3} Q_{o_m}(u_m). \tag{36}$$

This is evaluated on the $G^3 = 8{,}000$ grid points of the three-dimensional unit cube, giving the joint density at each combination of consumption, income, and wealth vigintiles.

C.3 Household Groups and Conditional Moments

A household group is defined by a range of quantiles in each dimension. For a group occupying quantile range $U_i^c \times U_i^y \times U_i^w$ (e.g., the bottom two income vigintiles crossed with all wealth vigintiles), I compute:

Mass (population share). The fraction of households in the cell:

$$\omega_{i,h} = \iiint_{(u_c,u_y,u_w)\in U_i^c\times U_i^y\times U_i^w} d\hat{C}_h(u_c, u_y, u_w). \tag{37}$$

Conditional mean consumption. The average consumption of households in the cell:

$$\bar{c}_{i,h} = \frac{1}{\omega_{i,h}} \iiint_{U_i^c\times U_i^y\times U_i^w} \hat{\Xi}_{c,h}^{-1}(u_c)\, d\hat{C}_h(u_c, u_y, u_w). \tag{38}$$

Income and wealth conditional means are defined analogously by replacing $\hat{\Xi}_{c,h}^{-1}$ with the corresponding marginal quantile function.

Numerical implementation. The continuous integrals (37)–(38) are discretized on the $G \times G \times G$ vigintile grid. Let $\mathcal{G}_i \subseteq \{1,\dots,G\}^3$ denote the grid indices belonging to household group i, and write $u_g = (g-1/2)/G$ for the midpoint of vigintile g. The discrete analogues of (37) and (38) are

$$\hat{\omega}_{i,h} = \frac{1}{G^3} \sum_{\mathbf{g}\in\mathcal{G}_i} d\hat{C}_h(u_{g_c}, u_{g_y}, u_{g_w}), \qquad \hat{\bar{c}}_{i,h} = \frac{1}{G^3\,\hat{\omega}_{i,h}} \sum_{\mathbf{g}\in\mathcal{G}_i} \hat{\Xi}_{c,h}^{-1}(u_{g_c})\, d\hat{C}_h(u_{g_c}, u_{g_y}, u_{g_w}), \tag{39}$$

with $\mathbf{g} = (g_c, g_y, g_w)$. The copula density is rescaled across the full grid so that $\sum_{\mathbf{g}} d\hat{C}_h(u_{g_c}, u_{g_y}, u_{g_w}) = G^3$, ensuring $\sum_i \hat{\omega}_{i,h} = 1$ when groups partition the unit cube.

C.4 Computation Summary

Table 4: Reconstruction parameters

Parameter	Value
Polynomial order (O)	11
Distribution dimensions (d)	3 (consumption, income, wealth)
Total coefficients ($N = (O+1)^d + d\cdot(O+1)$)	1,764
Copula coefficients ($(O+1)^d$)	1,728
Marginal coefficients ($d\cdot(O+1)$)	36
Distributional factors (K)	8
Variance explained by K factors	$> 95\%$
Evaluation grid (G)	20 (vigintiles)
Grid points for copula (G^3)	8,000
Gini evaluation grid (n)	200
Household groups	5 income $\times$ 5 wealth $=$ 25 cells
Posterior draws	5,000

For each of the 5,000 posterior draws, the full reconstruction pipeline (factor IRF $\rightarrow$ coefficient IRF $\rightarrow$ quantile functions + copula density $\rightarrow$ household moments + inequality measures)

is executed at each horizon $h = 1, \ldots, 20$. The reported impulse responses are posterior medians with 68% credible sets computed pointwise across draws. The tensor contraction over the G^d grid is the computational bottleneck; the implementation exploits the separable structure of the Legendre basis—each factor $Q_{o_m}(u_m)$ depends on only one dimension—to reduce the $O\big(G^d(O+1)^d\big)$ naïve evaluation to $O\big(G^d(O+1)\big)$ operations per draw via successive 1D contractions along each dimension.

Formally, the reconstruction proceeds as follows. Given the estimated factors $\hat{\boldsymbol{F}}_t$, I first recover the standardized coefficients, then undo normalization and trend adjustments:

$$\hat{\tilde{\theta}}_t = \hat{\Lambda}\,\hat{\boldsymbol{F}} \qquad \text{(Project factors)} \tag{40}$$

$$\hat{\theta}_{nt} = \sigma_{\zeta(n)}\hat{\tilde{\theta}}_{\zeta(n)t} + \mu_n + g(t)_n \qquad \text{(Unstandardize and add trend)} \tag{41}$$

$$\hat{\theta}_t = (\hat{\xi}_{m,o_1,t}, \ldots, \hat{\kappa}_{(o_1,\ldots,o_d),t}) \qquad \text{(Decomposition)} \tag{42}$$

$$\tag{43}$$

See Table 5 for details on notation. To generate economically interpretable objects—quantile functions, copula densities, Gini coefficients, percentile shares, and other moments—these reconstructed objects are evaluated on $u \in [0,1]^d$, with boundary points excluded in practice (the grid uses $[10^{-6}, 0.9995]$ for numerical stability) and then integrated to obtain the distributional statistics reported in the body.

Table 5: Reconstruction of Distributional Objects: Notation

Symbol	Description
$\hat{\boldsymbol{F}}_t$	Estimated high-frequency factors
$\hat{\Lambda}$	Factor loadings mapping factors to coefficients
$\hat{\tilde{\theta}}_t$	Standardized coefficient representation
$\hat{\theta}_{nt}$	Final coefficient vector for coefficient n
$\zeta(n)$	denotes a distributional object—either a quantile function or a copula, which consists of n coefficients
$\sigma_{\zeta(n)}$	Standard deviation *per* dist. object
μ_n	Mean adjustment *per* coefficient
$g(t)_n$	Trend component per coefficient
$\hat{\xi}_{m,o,t}$	Marginal polynomial coefficients
$\hat{\kappa}_{(o_1,\ldots,o_d),t}$	Copula coefficients
$Q_o(u_m)$	Legendre polynomial basis function

Notes: The table summarizes the mapping from factor estimates to reconstructed distributional objects.

D Marginal-Only Specifications

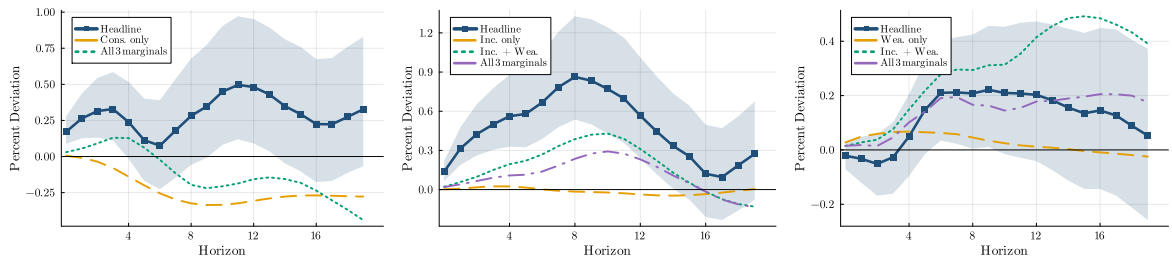
This appendix complements the baseline specification — which models the joint distribution of consumption, income, and wealth via the copula together with all three marginals — with five *marginal-only* specifications that selectively drop distributional information from the VAR. In each case, the macroeconomic block is unchanged; only the distributional factor block varies. The purpose is twofold. First, it benchmarks the baseline against the empirical strategy that dominates the literature: estimation on a single marginal dimension or a stack of marginals *without* the copula. Second, it isolates how much of the baseline inequality response is driven by the joint distribution rather than by any individual marginal.

Concretely, the five specifications are: (i) consumption marginal only (12 Legendre coefficients), (ii) income marginal only (12), (iii) wealth marginal only (12), (iv) income and wealth marginals stacked but no copula (24), and (v) all three marginals stacked but no copula (36). All five share the same baseline macro block, lag length, prior, and posterior draws; they differ only in which distributional coefficients enter the VAR and (mechanically) in the number of distributional principal components carried into the system, which is capped at three for marginal-only specifications.

Figures 13 and 14 report the Gini IRF for each dimension under the baseline (solid line, 68% credible band) and under each marginal-only specification that includes that dimension (dashed lines, posterior median only). The takeaway is direct: when the copula is dropped, the inequality response shifts in magnitude and sometimes in shape, and the shift is more pronounced under the realized supply shock than under the news shock. Modeling the joint matters most for the dimensions where wealth-income comovement carries the cyclical signal — consumption and wealth — and matters least for income, whose marginal response is largely picked up by the income coefficients alone.

Figure 13: Gini IRFs across marginal-only specifications

Panel A: Baumeister & Hamilton (2019) realized supply shock

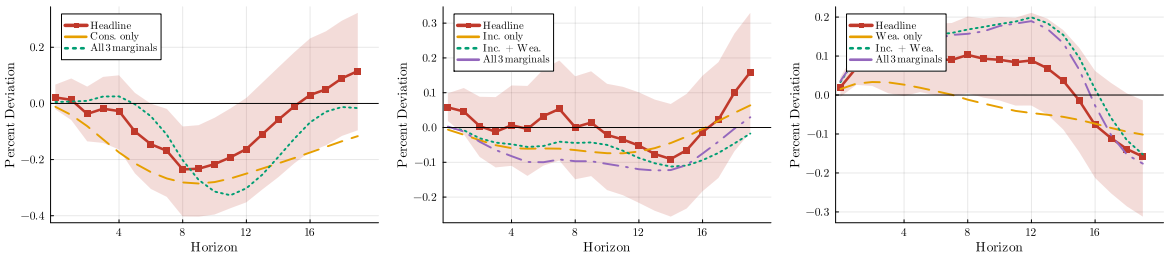


Consumption Gini | Income Gini | Wealth Gini

Notes: IRF of the Gini, percent deviation from steady state. Solid line + 68% credible band: baseline VAR (full distribution: copula + all three marginals). Dashed lines: marginal-only specifications that include the named dimension (consumption-only, income-only, wealth-only, income+wealth, all three marginals; see Section D for definitions). Macro block, lag length, prior, and posterior draws are identical across specifications.

Figure 14: Gini IRFs across marginal-only specifications

Panel B: Känzig (2021) supply news shock



Consumption Gini | Income Gini | Wealth Gini

Notes: IRF of the Gini, percent deviation from steady state. Solid line + 68% credible band: baseline VAR (full distribution: copula + all three marginals). Dashed lines: marginal-only specifications that include the named dimension. Macro block, lag length, prior, and posterior draws are identical across specifications.

E Inventory Measurement Error and Robustness

Baumeister and Hamilton (2019), Section III, document substantial measurement error in the OECD crude-oil inventory series of Kilian and Murphy (2014a) and propose a measurement-error-corrected reconstruction; details are in their paper. I adopt their construction for the baseline VAR: cumulating their monthly Δi_t series, anchored at `oilstocksM` in 1975Q1, yields the quarterly log-level series `oilstocksM_BH`, which I extend through 2024Q2 by replicating their construction on current EIA inputs and splicing at 2019M12. Figure 15 compares the result to the legacy `oilstocksM` of Känzig (2021): the level correlation is 0.997, but the $\Delta \log$ correlation is only 0.27 — the empirical content of BH's measurement-error equation, since a VAR identifies dynamics from changes.

Appendix F.1 reverts to the noisier legacy `oilstocksM` measure, holding the structural setup, prior, identification, and lag length fixed. The qualitative pattern is intact, ruling out a measurement-error explanation.

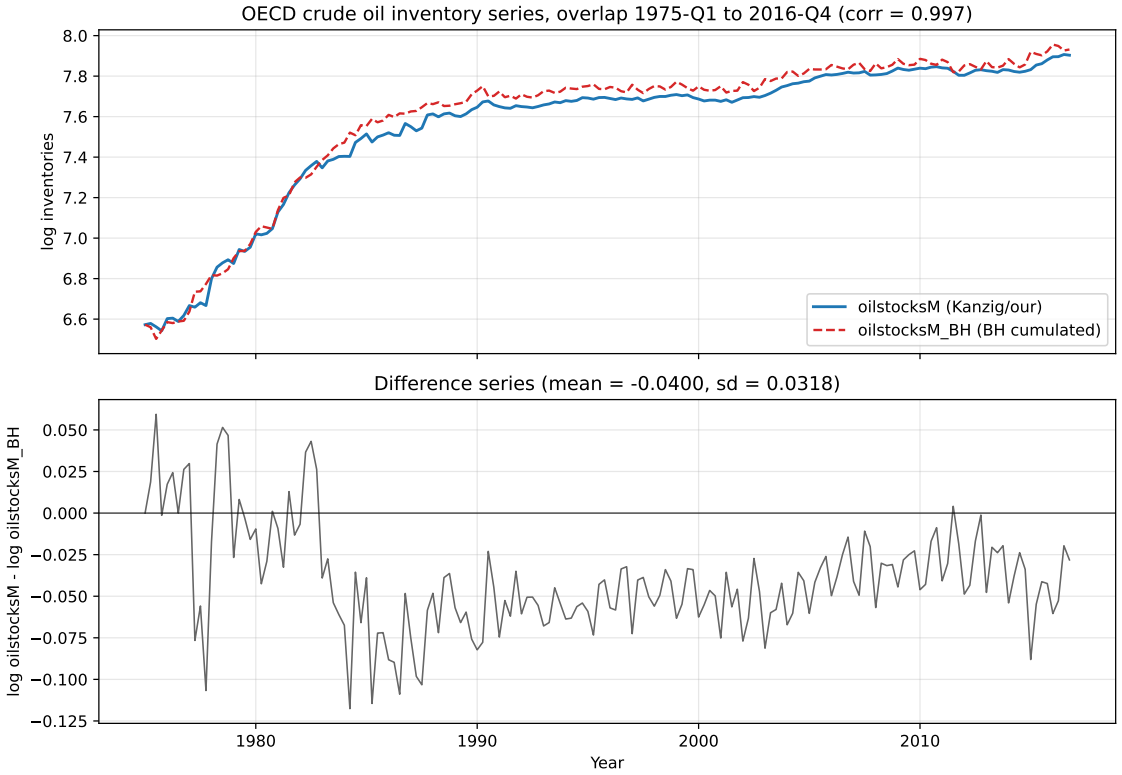


Figure 15: Comparison of two OECD crude-oil inventory series, 1975Q1–2016Q4

Notes: `oilstocksM`: legacy series from Känzig (2021). `oilstocksM_BH`: reconstructed by cumulating Baumeister and Hamilton (2019)'s monthly Δi_t data, anchored at 1975Q1. Top: log-level overlay. Bottom: difference series.

F Robustness

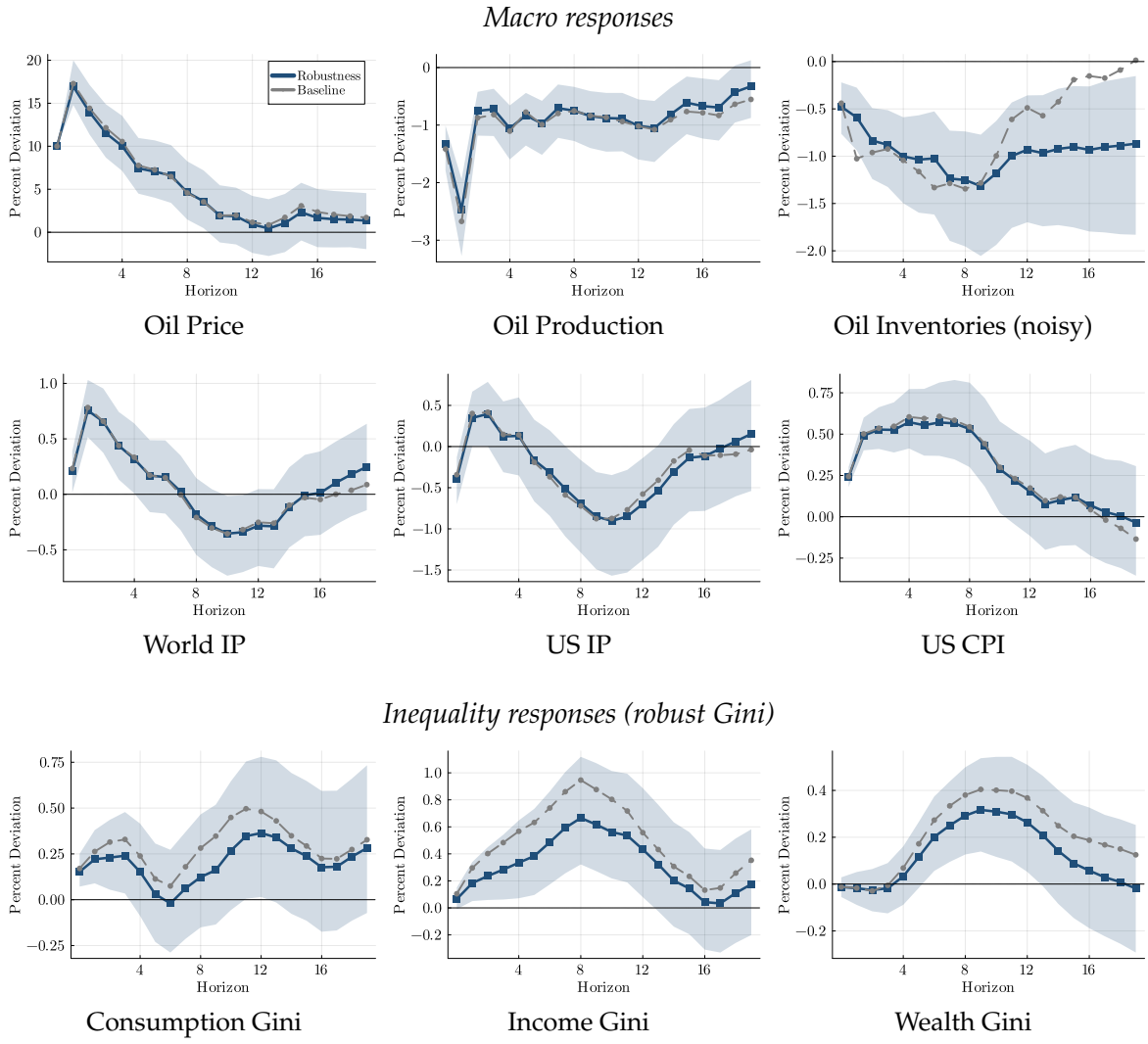
This appendix collects baseline-vs-robustness IRF comparisons for every robustness specification estimated in the paper. Each subsection corresponds to one robustness theme; within a subsection, each figure overlays the baseline IRF (solid + 68% credible band, shock-aware color) with the robustness IRF (dashed gray) for six macro variables and three robust-Gini measures of inequality.[26]

F.1 Legacy Inventory Series

I revert to the legacy `oilstocksM` inventory measure (Kanzig 2021's deseasonalised OECD series) instead of the baseline `oilstocksM_BH` (Baumeister & Hamilton 2019's measurement-error-corrected reconstruction); reverting to `oilstocksM` restores the full 1975Q1–2024Q3 sample (vs. 1975Q1–2016Q4 under the baseline).

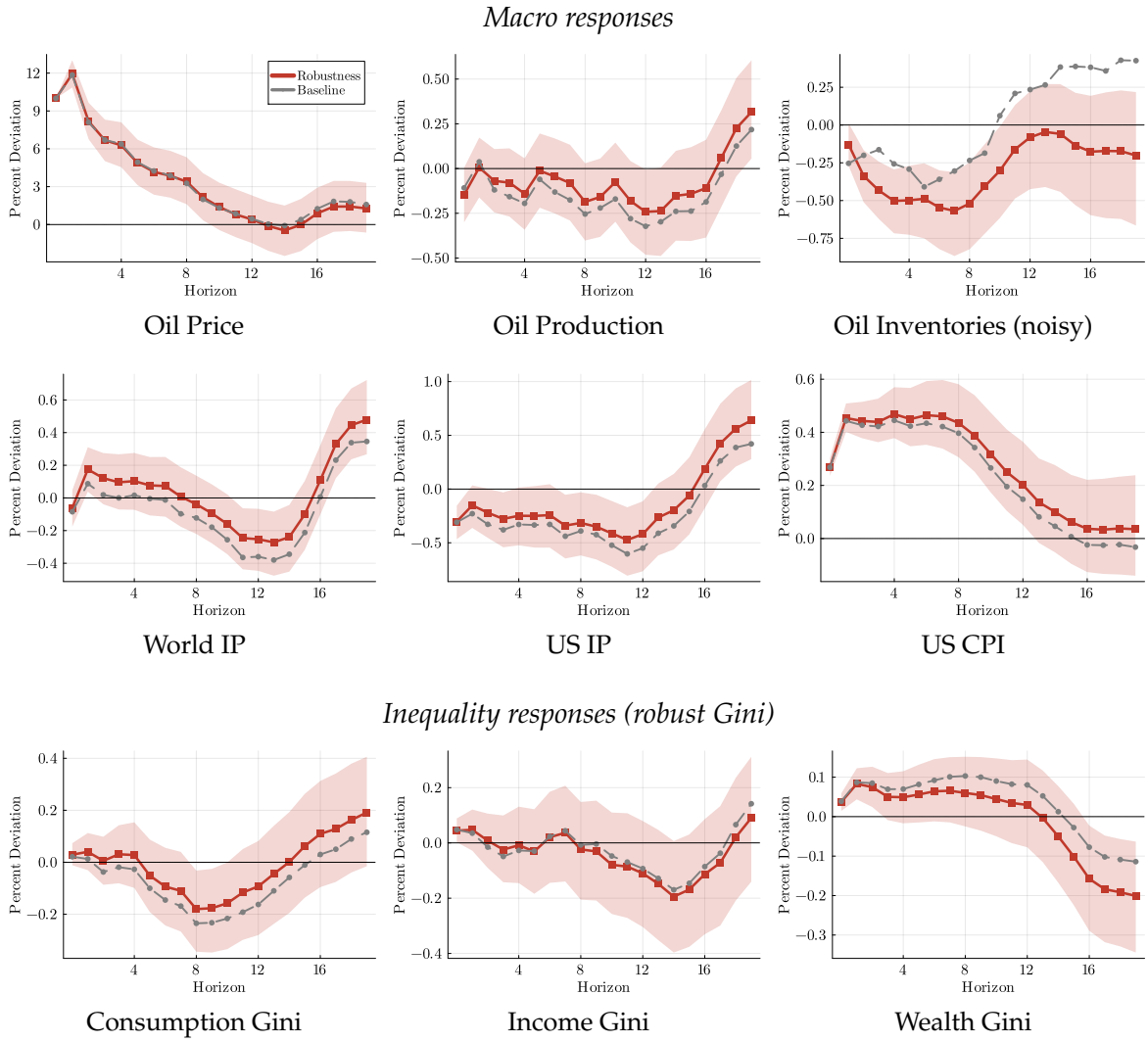
[26] Results are also unchanged under several additional checks not reported here for space: alternative oil-price measures (the U.S. refiners' acquisition cost of imported crude and the real-log Brent front-month price, addressing Kilian and Zhou 2023); the QD-levels factor specification with the distributional block dropped; posterior mean and posterior mode in place of the median (addressing the asymmetry concerns of Kilian and Zhou 2023); and 90% pointwise credible bands in place of the 68% bands. Figures available on request.

Figure 16: Legacy oilstocksM — Baumeister and Hamilton (2019) supply shock



Notes: Solid + 68% band: robustness (Legacy oilstocksM). Dashed gray: baseline. Macro responses are level IRFs.

Figure 17: Legacy oilstocksM — Känzig (2021) supply news shock

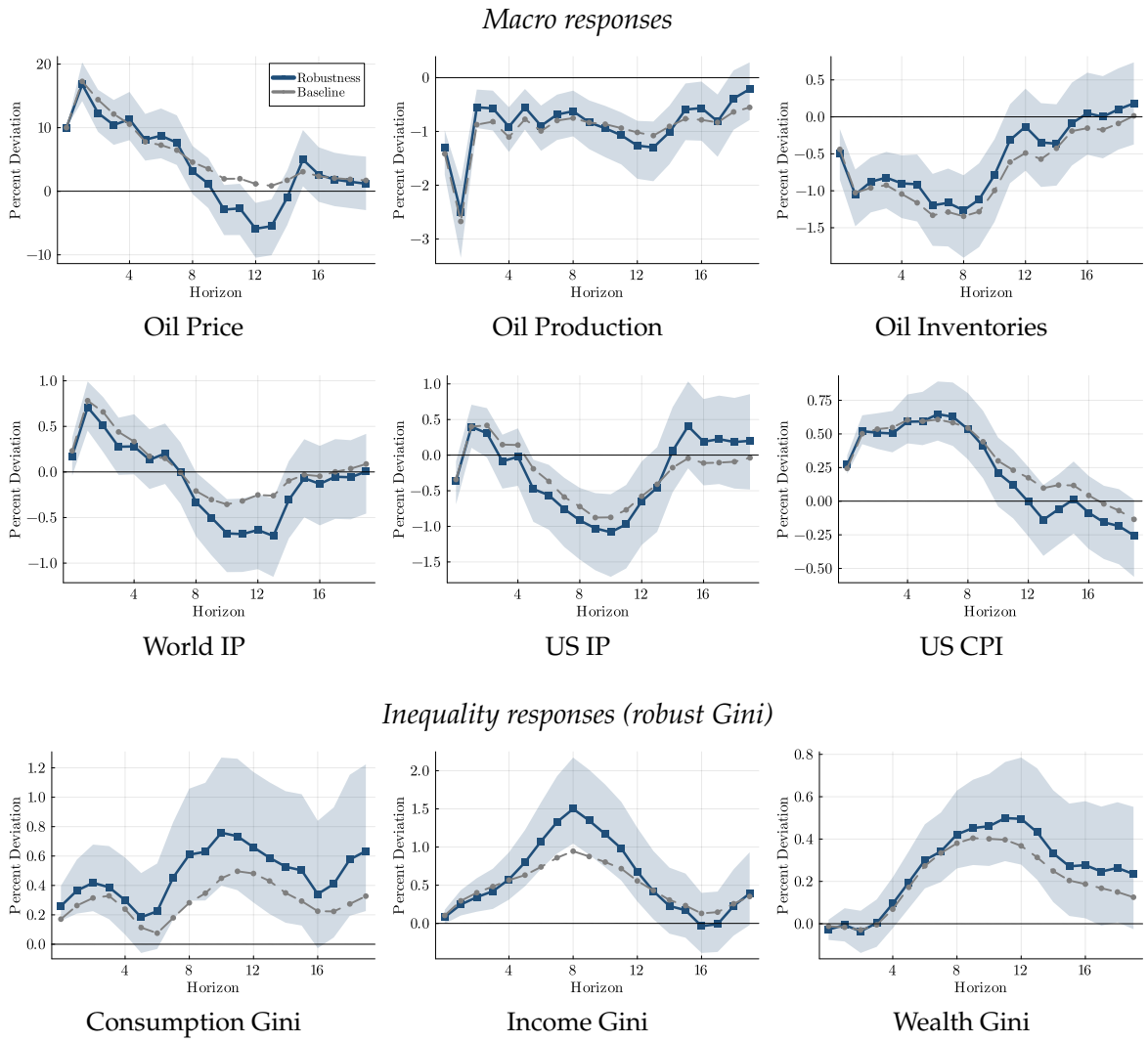


Notes: Solid + 68% band: robustness (Legacy oilstocksM). Dashed gray: baseline. Macro responses are level IRFs.

F.2 Mori-Peersman Financial Block

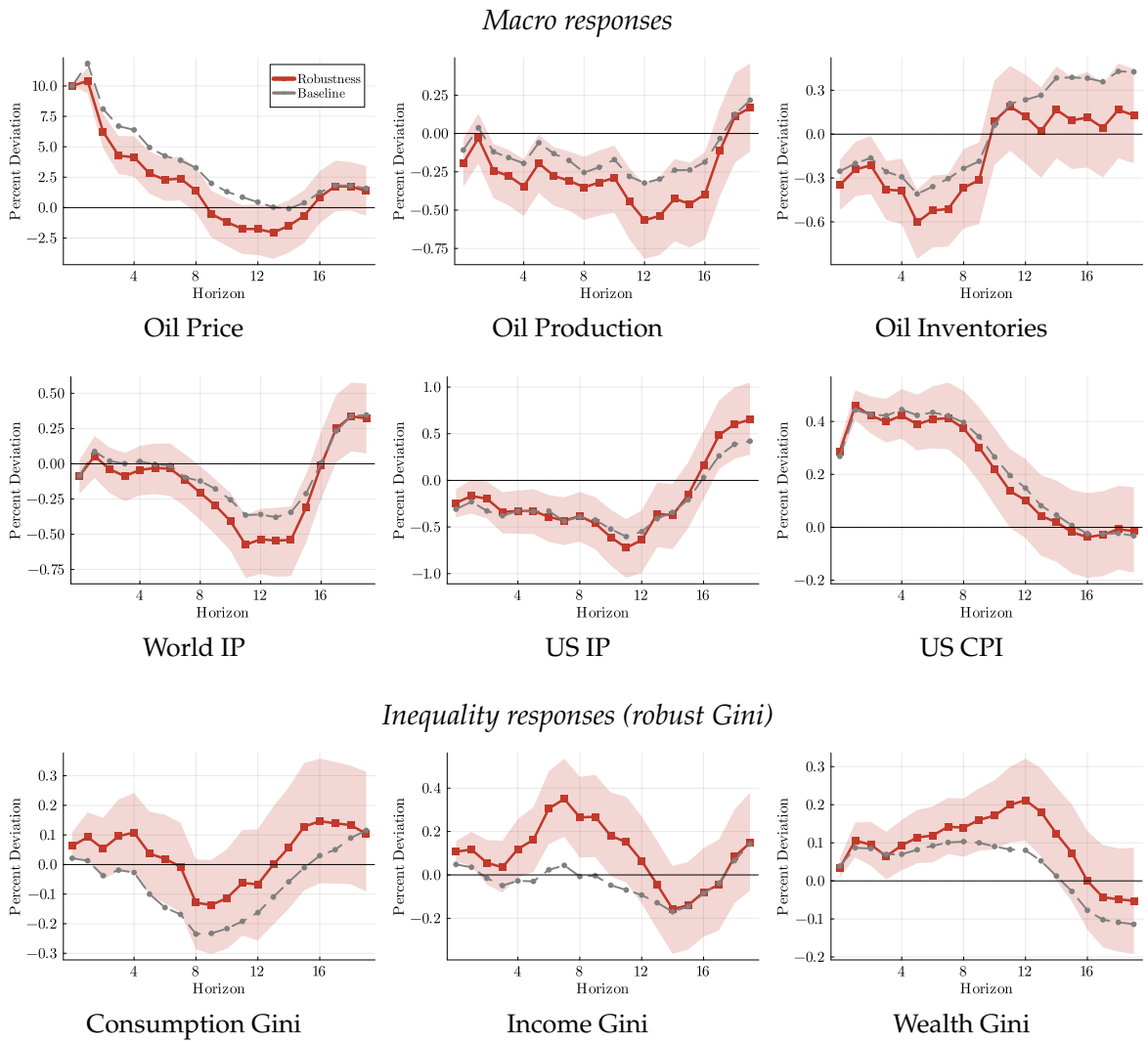
I augment the baseline VAR with the Mori-Peersman financial block in two variants. The bare *MP* block adds the 1-year Treasury yield, S&P 500, and VIX. The *MP+EBP* block additionally appends the Gilchrist-Zakrajšek excess bond premium (`ebp`). The motivation is the Mori and Peersman (2024) predictability concern: at quarterly frequency, my own Granger tests (Appendix A) point to the EBP as the predictor for the KZ shock, while interest rates predict the BH shock. The MP block is the literature-matched specification; MP+EBP additionally controls for the EBP channel that binds at quarterly frequency.

Figure 18: MP financial block (GS1, S&P 500, VIX) — Baumeister and Hamilton (2019) supply shock



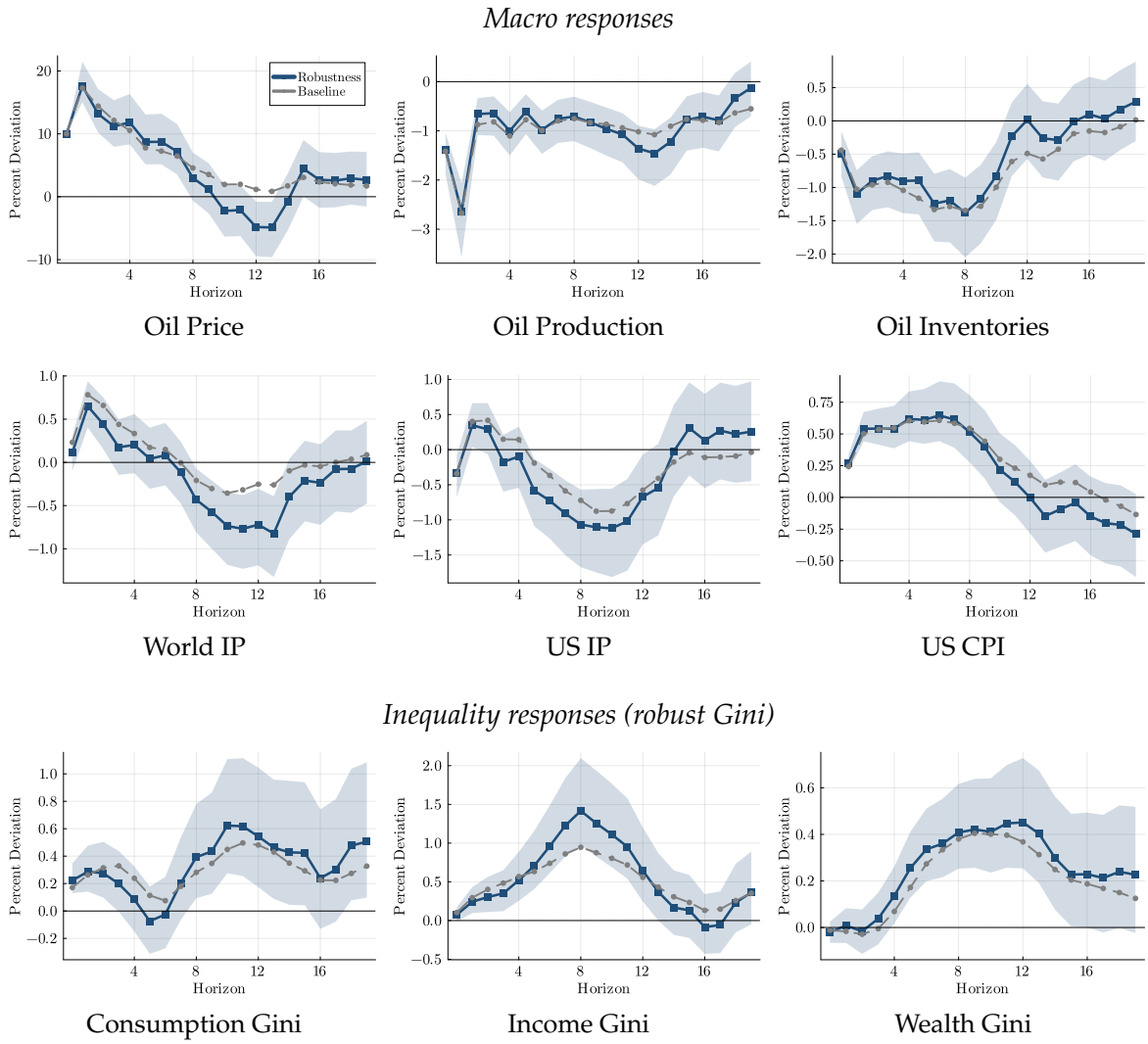
Notes: Solid + 68% band: robustness (MP financial block (GS1, S&P 500, VIX)). Dashed gray: baseline. Macro responses are level IRFs.

Figure 19: MP financial block (GS1, S&P 500, VIX) — Känzig (2021) supply news shock



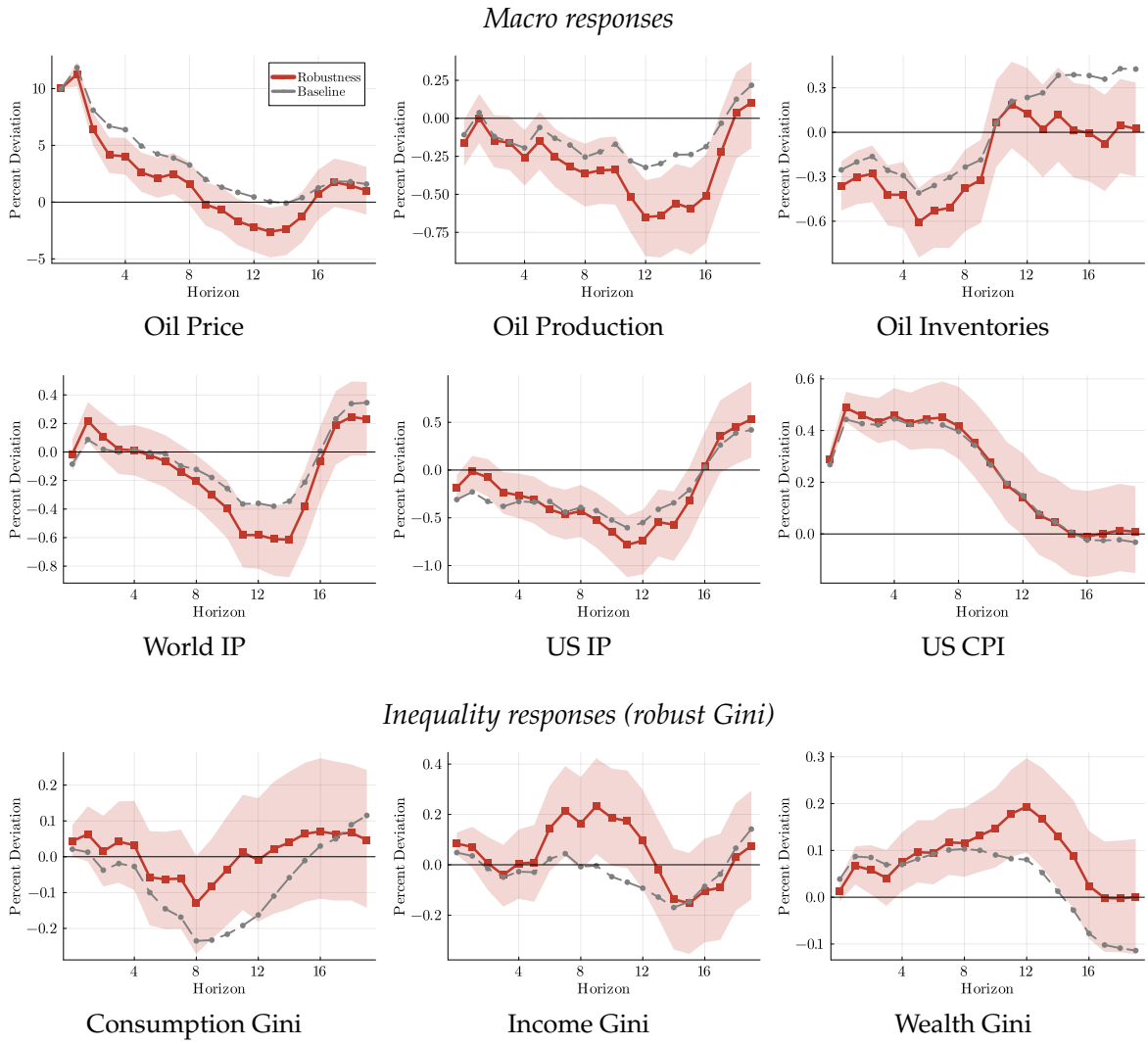
Notes: Solid + 68% band: robustness (MP financial block (GS1, S&P 500, VIX)). Dashed gray: baseline. Macro responses are level IRFs.

Figure 20: MP + EBP financial block — Baumeister and Hamilton (2019) supply shock



Notes: Solid + 68% band: robustness (MP + EBP financial block). Dashed gray: baseline. Macro responses are level IRFs.

Figure 21: MP + EBP financial block — Känzig (2021) supply news shock

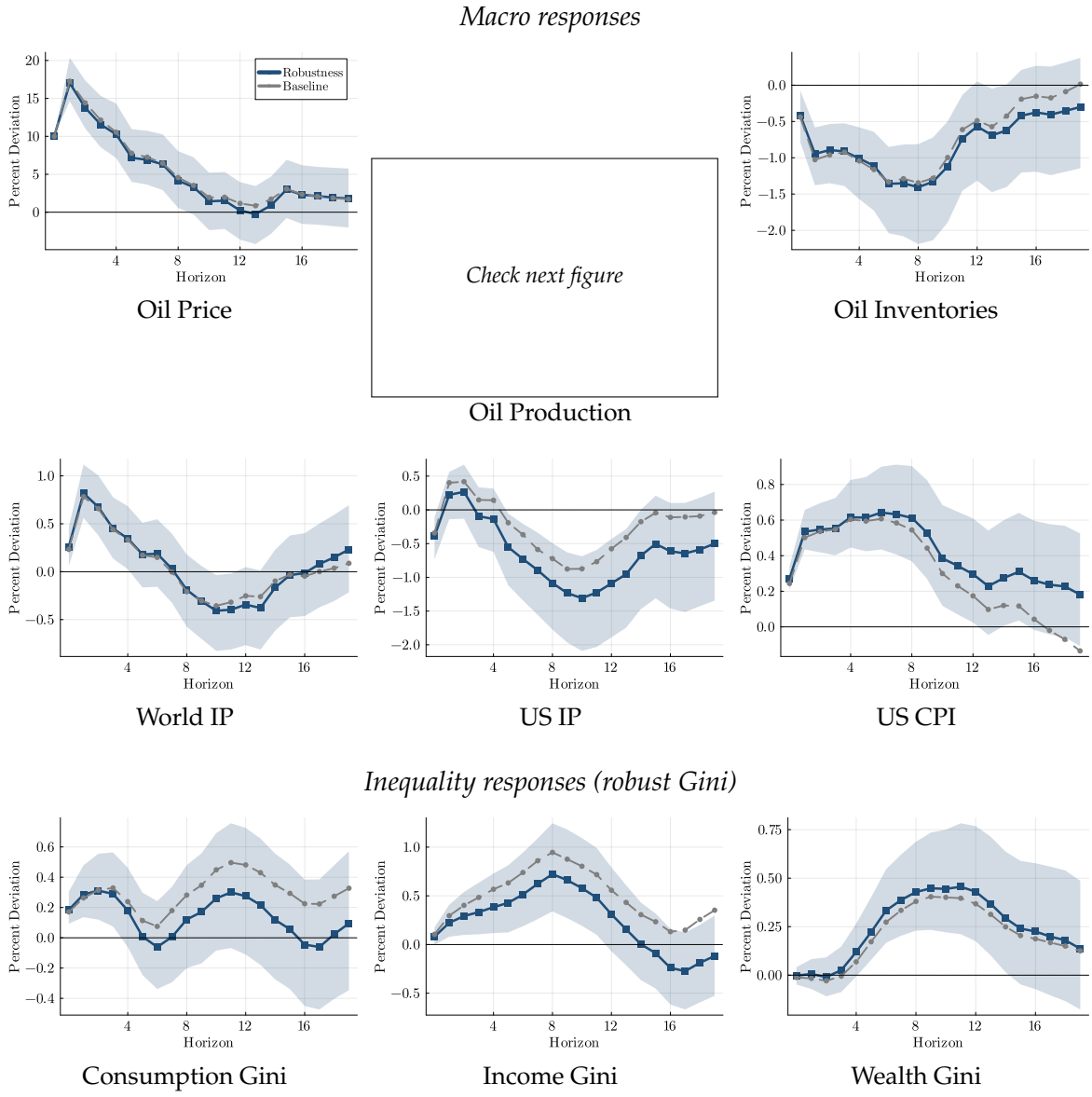


Notes: Solid + 68% band: robustness (MP + EBP financial block). Dashed gray: baseline. Macro responses are level IRFs.

F.3 OPEC / Non-OPEC Production Split

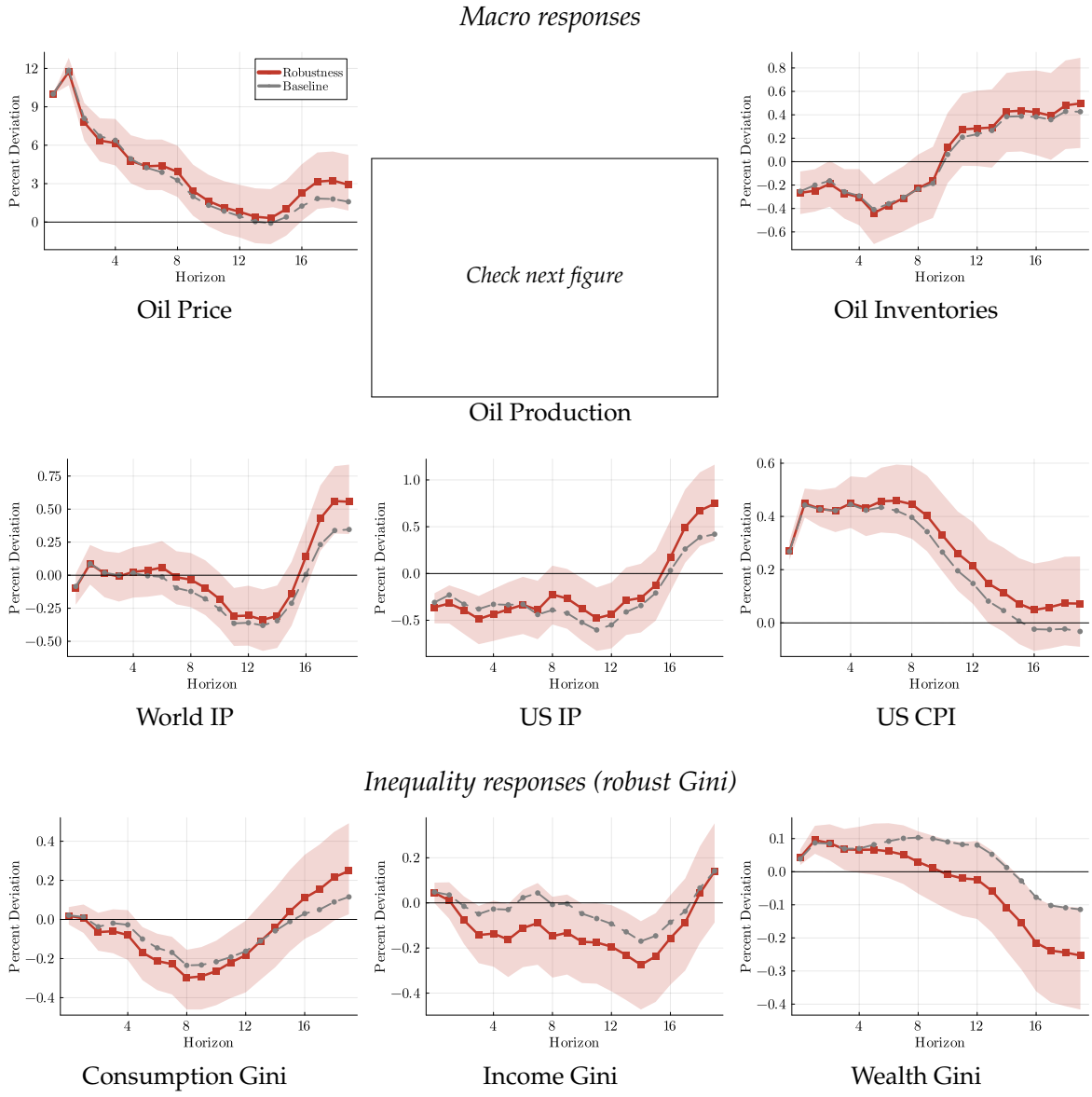
Following Kolodzeij and Kaufmann (2014), who argue that OPEC and non-OPEC producers operate under different output-setting criteria so that reductions in OPEC production tend to raise oil prices while non-OPEC movements act differently, I replace the aggregate `oilprod` variable with two components: `oilprod_opec` (OPEC production) and `oilprod_nopec` (Non-OPEC production). Series are constructed by applying EIA INTL OPEC and Non-OPEC shares (X-12-ARIMA seasonally adjusted) to the baseline `oilprod` level, so that $\exp(\texttt{oilprod\_opec}) + \exp(\texttt{oilprod\_nopec}) = \exp(\texttt{oilprod})$ by construction. This tests whether baseline supply-shock dynamics are sensitive to treating OPEC discipline separately from non-OPEC supply (US shale, North Sea, etc.). The `oilprod` panel below is absent because the OPEC specification does not contain an aggregate production variable; the OPEC and non-OPEC component IRFs are reported separately in Figure 24.

Figure 22: OPEC + Non-OPEC split — Baumeister and Hamilton (2019) supply shock



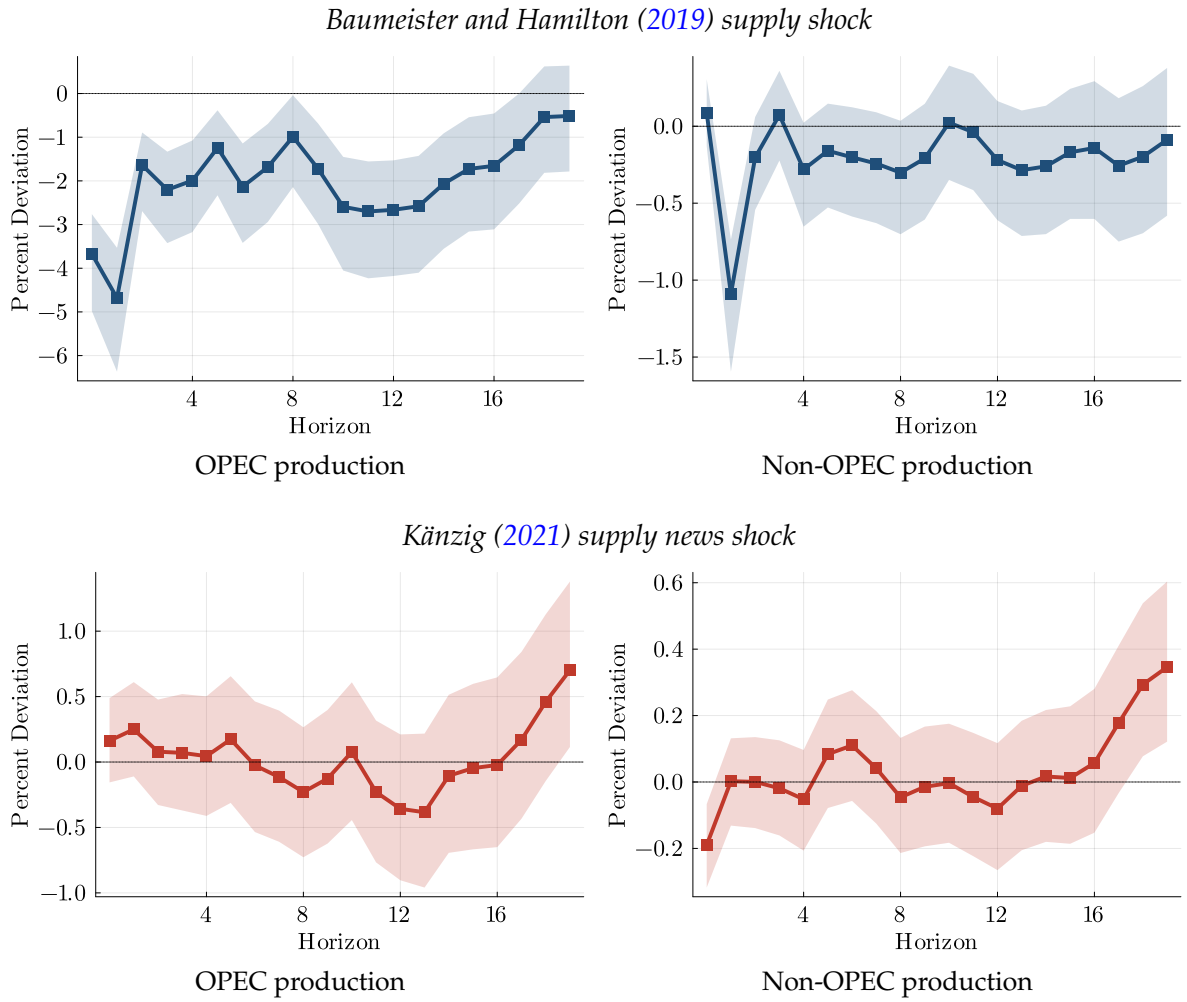
Notes: Solid + 68% band: robustness (OPEC + Non-OPEC split). Dashed gray: baseline. Macro responses are level IRFs.

Figure 23: OPEC + Non-OPEC split — Känzig (2021) supply news shock



Notes: Solid + 68% band: robustness (OPEC + Non-OPEC split). Dashed gray: baseline. Macro responses are level IRFs.

Figure 24: OPEC and Non-OPEC production responses (OPEC-split specification)

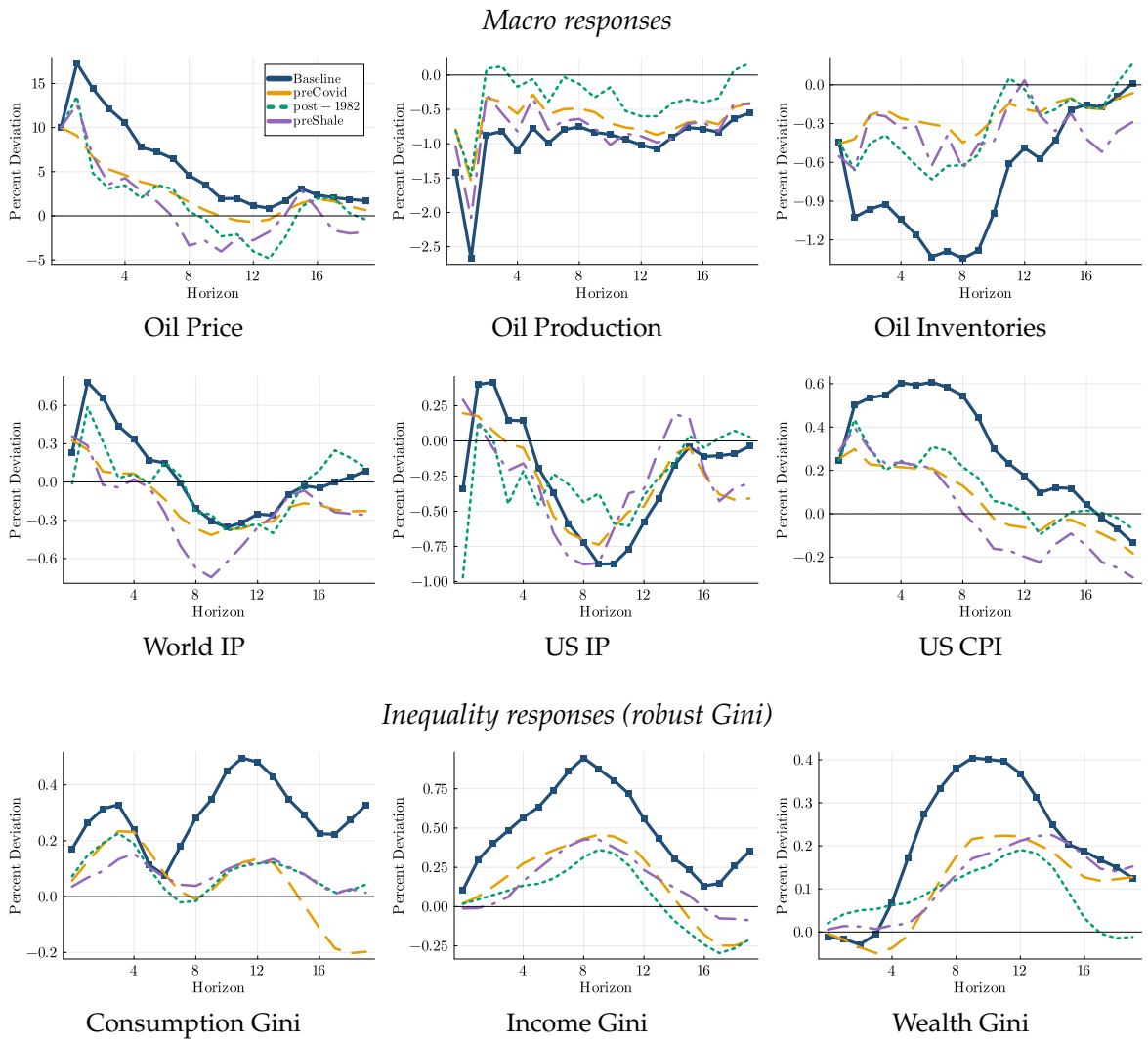


Notes: Solid line: posterior median; shaded band: 68% credible set. The OPEC-split specification replaces `oilprod` with `oilprod_opec` and `oilprod_nopec`; these IRFs have no baseline counterpart in the baseline VAR (where only the aggregate `oilprod` is identified) and are reported on their own.

F.4 Subsamples

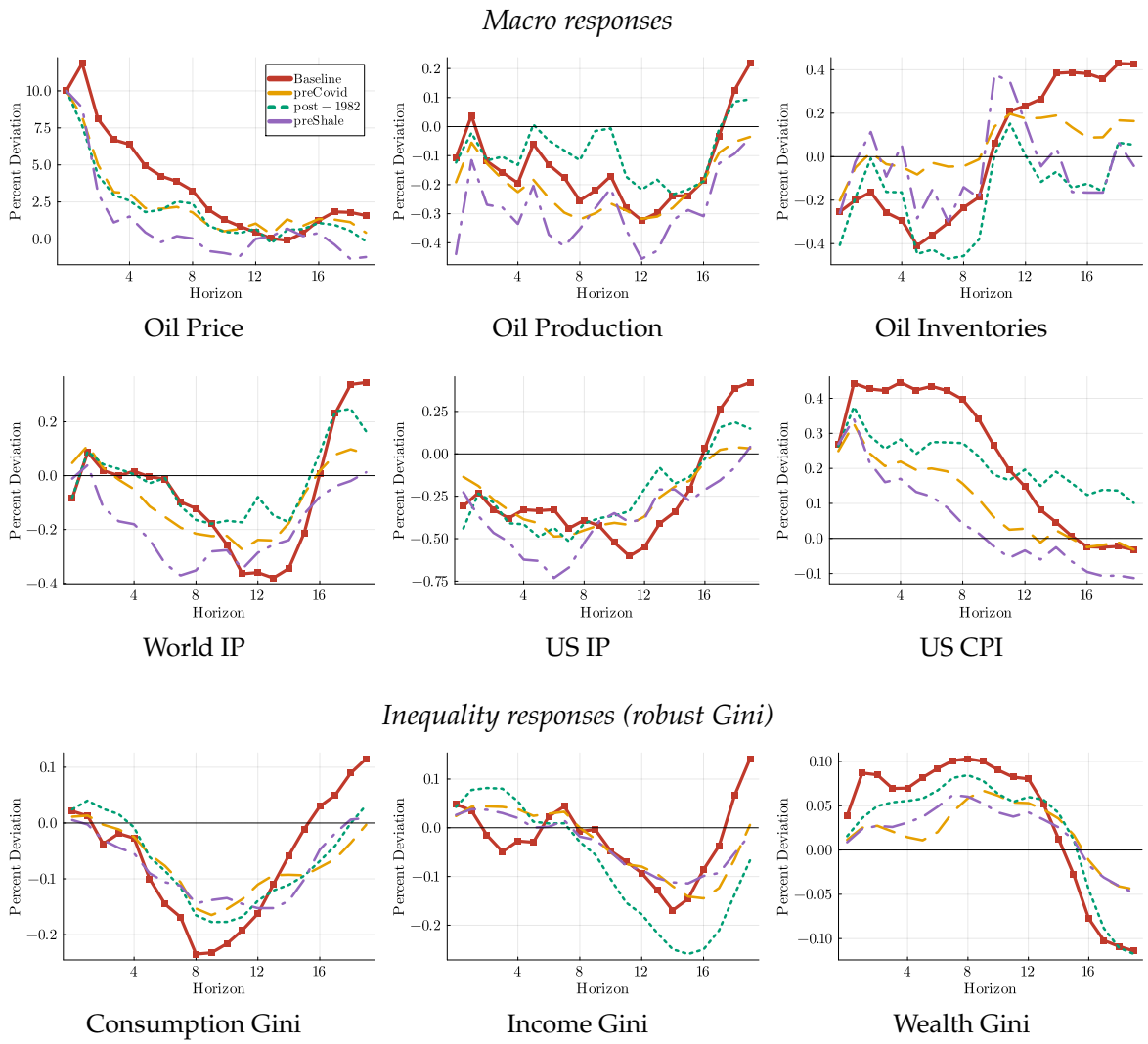
I re-estimate the baseline VAR on three subsamples: *preCovid* (truncated at 2019Q4 to exclude the pandemic and its aftermath), *post1982* (starting in 1982Q1 to drop the early Volcker disinflation period), and *preShale* (truncated at 2010Q4, before the U.S. shale-oil revolution materially raised the medium-run elasticity of supply). The post-1982 cut is also motivated by Nakov and Pescatori (2010), who note that the period around 1981 marks a regime break in world oil markets, in U.S. energy production and consumption, and in monetary policy and inflation; Hooker (2002) further documents that the pass-through from oil prices to inflation weakens substantially after this break, so the subsample doubles as a check that the inflation responses are not artifacts of pre-Volcker passthrough.

Figure 25: Subsample sweep — Baumeister and Hamilton (2019) supply shock



Notes: Each panel: baseline (solid, shock-aware color, full sample) + three dashed lines for *preCovid* ($\leq$ 2019Q4), *post-1982* ($\geq$ 1982Q1), and *preShale* ($\leq$ 2010Q4). No credible bands. Macro responses are level IRFs.

Figure 26: Subsample sweep — Känzig (2021) supply news shock

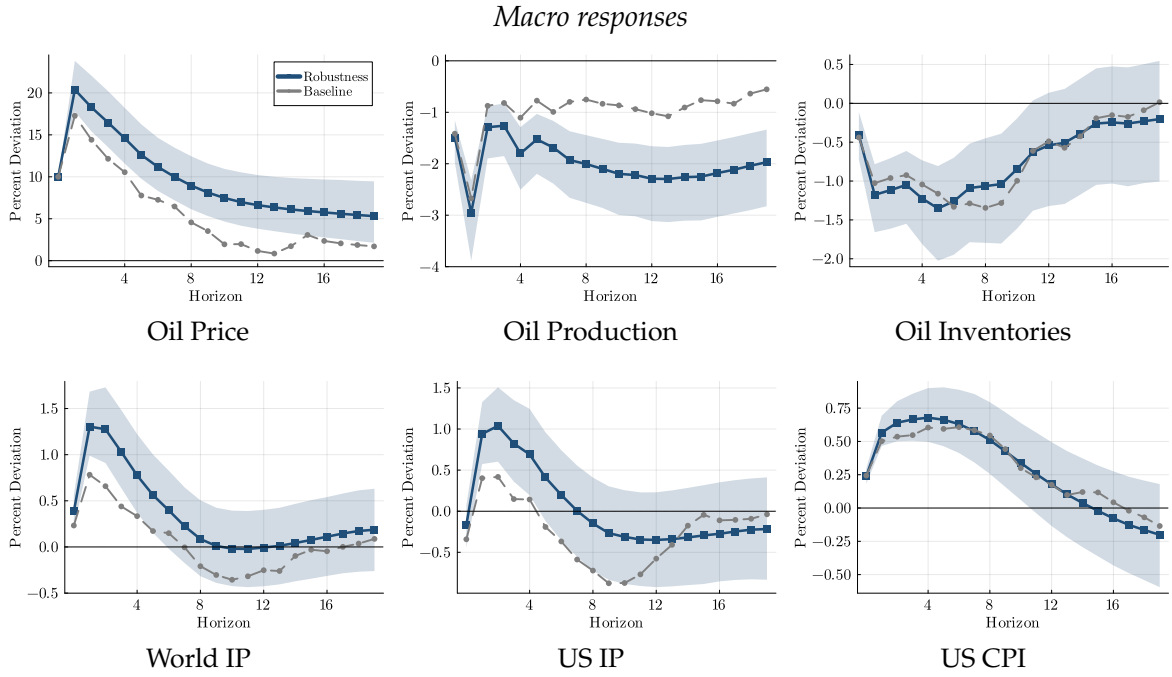


Notes: Each panel: baseline (solid, shock-aware color, full sample) + three dashed lines for *preCovid* ($\leq$ 2019Q4), *post-1982* ($\geq$ 1982Q1), and *preShale* ($\leq$ 2010Q4). No credible bands. Macro responses are level IRFs.

F.5 Aggregates-Only Specification

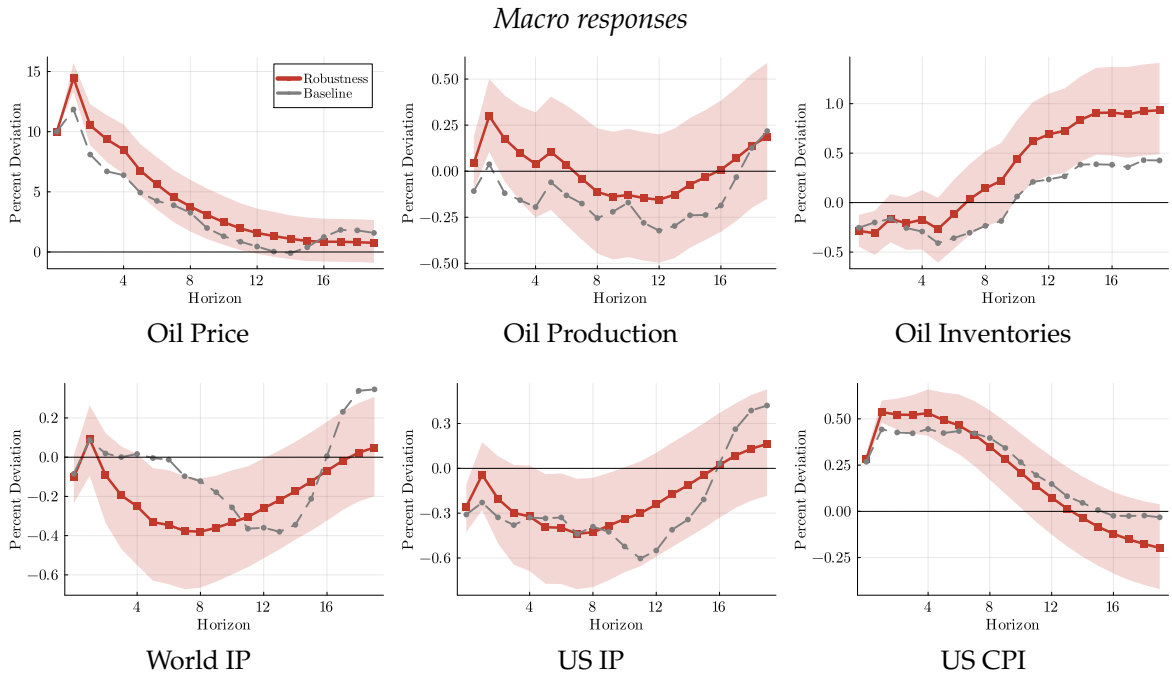
I drop the distributional block from the VAR. This is the standard small SVAR; the comparison reveals whether including the distribution changes the aggregate IRFs. (Inequality panels are absent because the model has no distributional measures.)

Figure 27: Aggregates only — Baumeister and Hamilton (2019) supply shock



Notes: Solid + 68% band: robustness (Aggregates only). Dashed gray: baseline. Macro responses are level IRFs.

Figure 28: Aggregates only — Känzig (2021) supply news shock

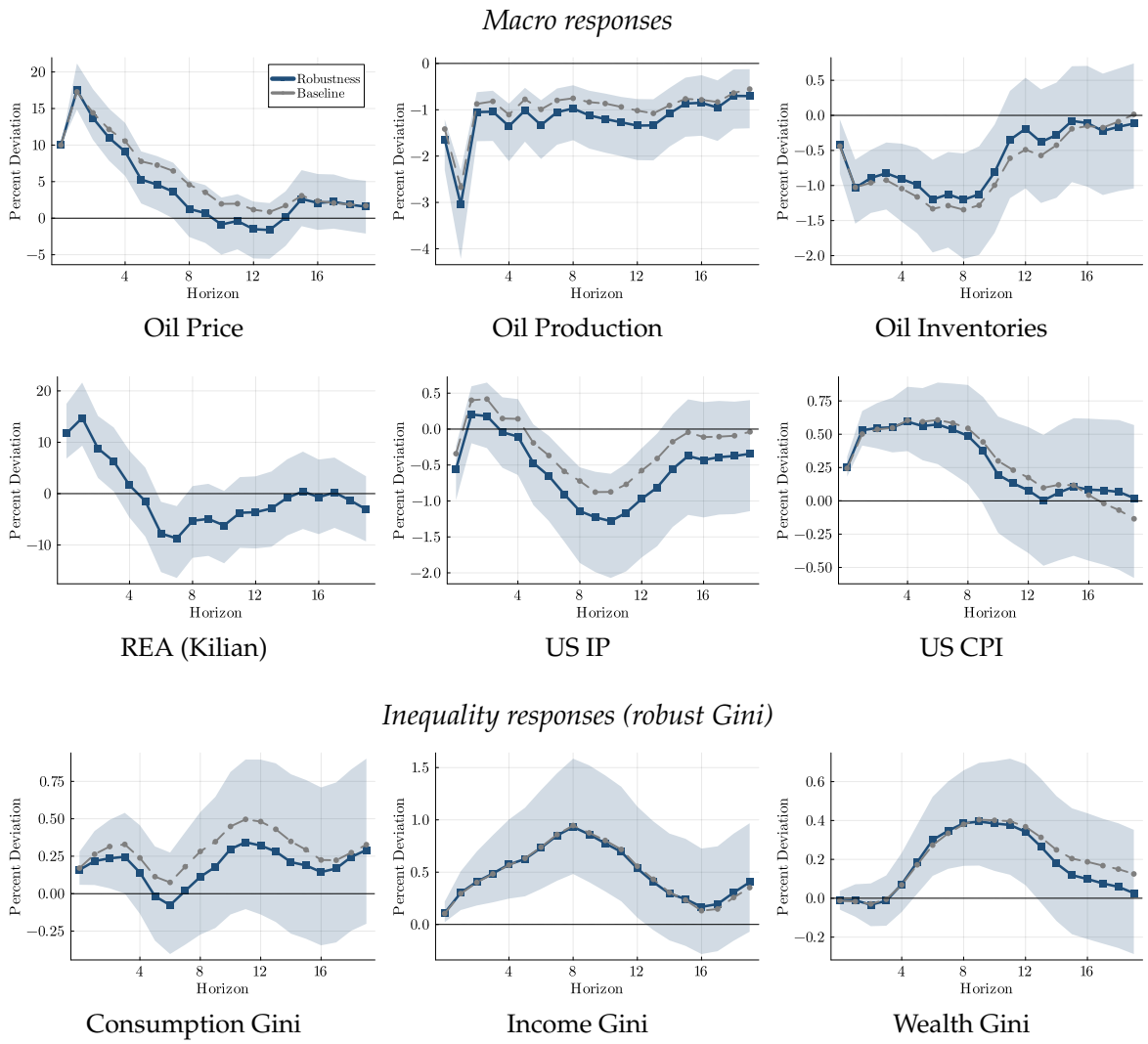


Notes: Solid + 68% band: robustness (Aggregates only). Dashed gray: baseline. Macro responses are level IRFs.

F.6 Kilian REA in Place of World IP

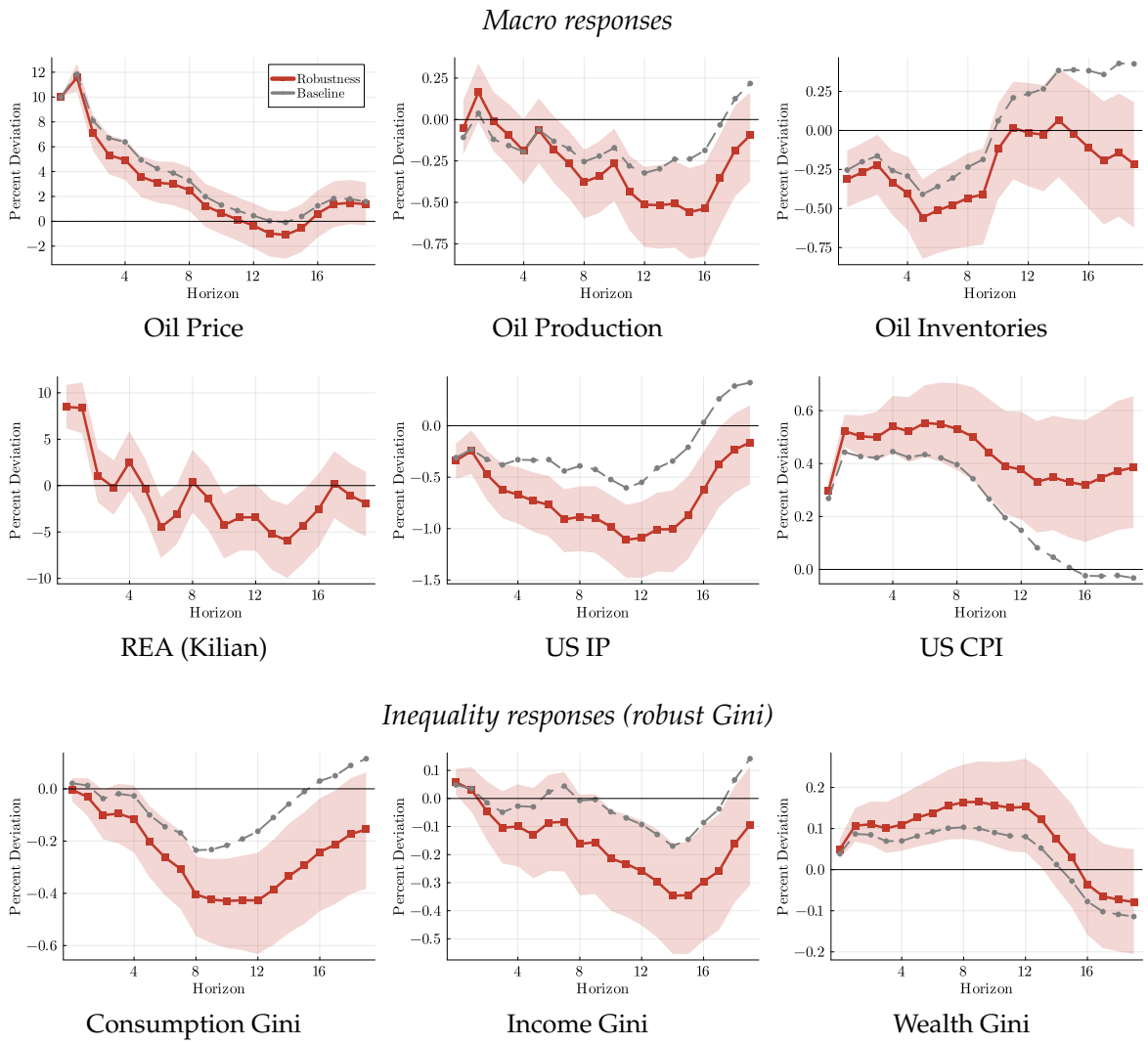
I replace world industrial production with Kilian's (2009a) Real Economic Activity index (deflated dry-cargo shipping rate, detrended). The data window is determined by the REA series.

Figure 29: Kilian REA — Baumeister and Hamilton (2019) supply shock



Notes: Solid + 68% band: robustness (Kilian REA). Dashed gray: baseline. *World IP / REA panel*: only the robustness REA series is shown because the two measures are on different scales. Macro responses are level IRFs.

Figure 30: Kilian REA — Känzig (2021) supply news shock

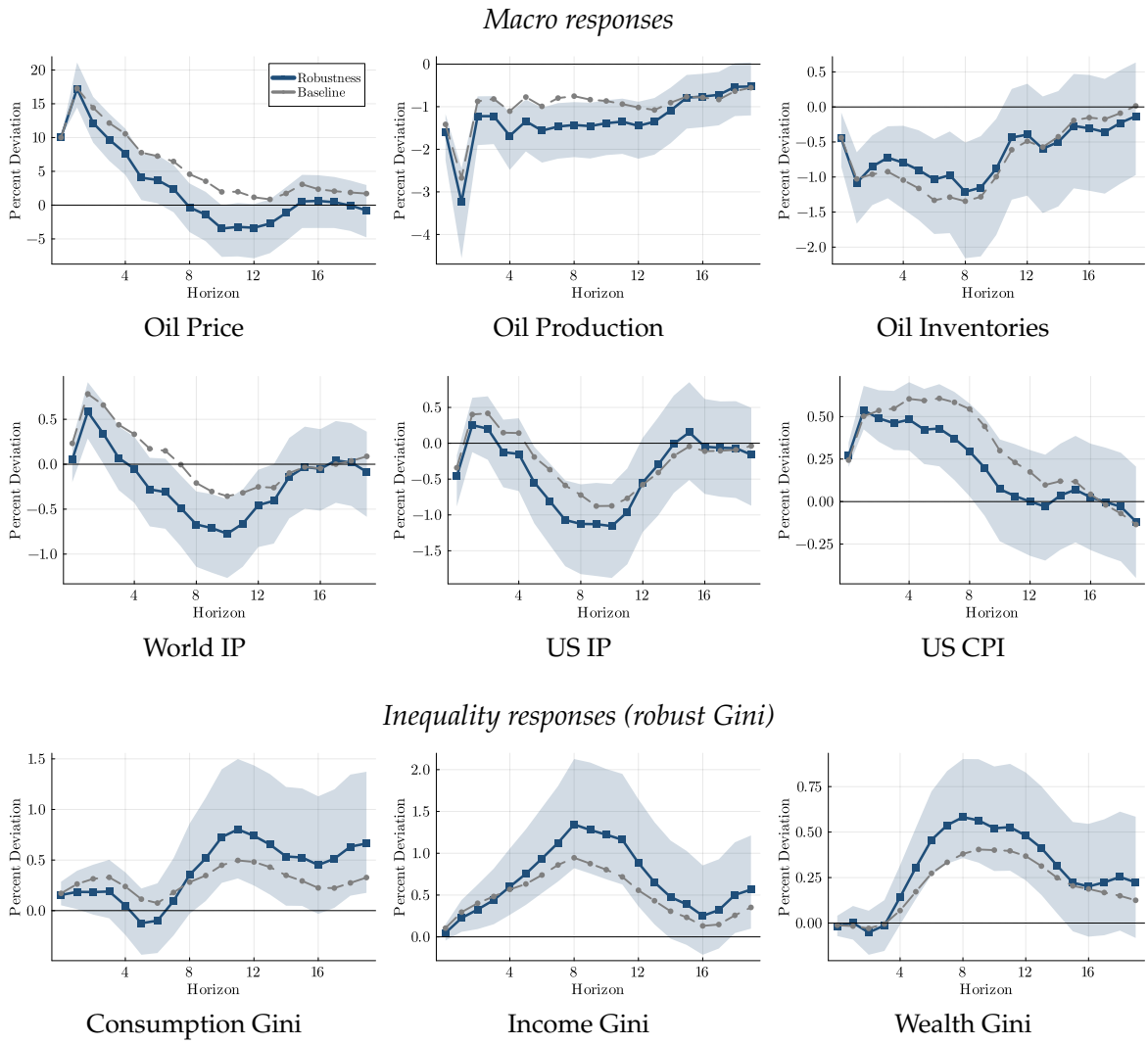


Notes: Solid + 68% band: robustness (Kilian REA). Dashed gray: baseline. *World IP / REA panel*: only the robustness REA series is shown because the two measures are on different scales. Macro responses are level IRFs.

F.7 QD-Levels Macro Factors

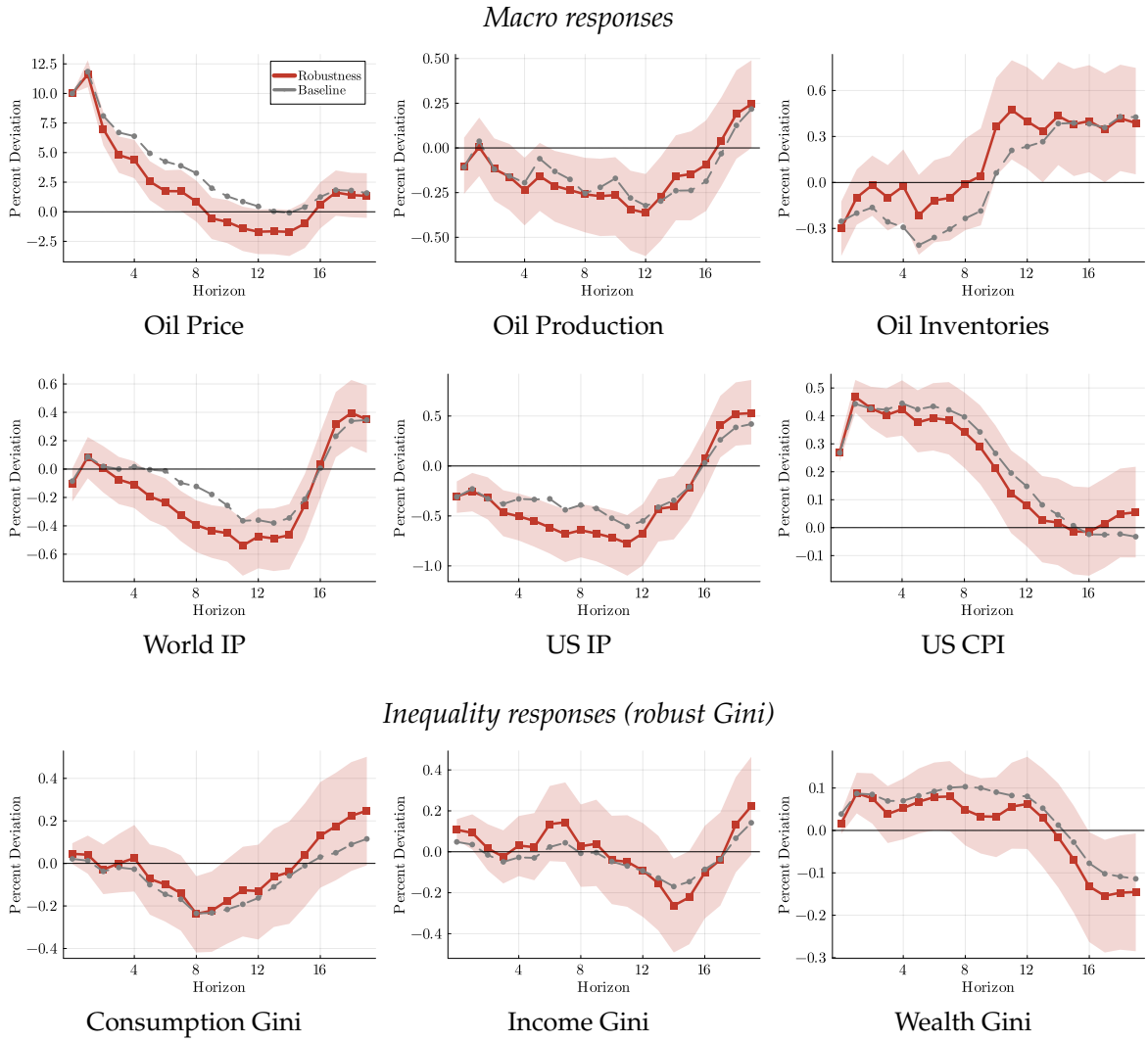
I re-estimate the baseline VAR augmenting the macro block with the first principal components of the FRED-QD log-level macro panel (Bai 2004 non-stationary PCA). This guards against the baseline result being an artefact of a six-variable macro block missing common factors.

Figure 31: QD-levels factors — Baumeister and Hamilton (2019) supply shock



Notes: Solid + 68% band: robustness (QD-levels factors). Dashed gray: baseline. Macro responses are level IRFs.

Figure 32: QD-levels factors — Känzig (2021) supply news shock



Notes: Solid + 68% band: robustness (QD-levels factors). Dashed gray: baseline. Macro responses are level IRFs.

G Prior for the Long Run: Construction and Diagnostic

This appendix details the construction of the prior for the long run (PLR) used in estimation and reports the auto-detection output for the baseline VARs (`paper_bh_smfdd` and `paper_kanzig_smfdd`: shock instrument + 6 macroeconomic variables + 8 smoothed distributional factors = 15 endogenous variables, 16 lags, 1975–2024 quarterly).

Dummy-observation construction. The PLR of Giannone, Lenza, and Primiceri (2019) regularizes the long-run multiplier matrix $\boldsymbol{A}(1) = \boldsymbol{I} - \boldsymbol{A}_1 - \cdots - \boldsymbol{A}_p$ by augmenting the likelihood with n artificial observations $(\boldsymbol{Y}_d, \boldsymbol{X}_d)$ prepended to the data. The dummy in row i encodes the prior belief that the long-run combination $\boldsymbol{H}_{i\cdot}\boldsymbol{Y}_t$ has mean $\bar{y}_{0,i}$, with strength inversely proportional to a tightness parameter ϕ_i:

$$\boldsymbol{Y}_{d,i,\cdot} \;=\; \frac{\boldsymbol{H}_{i\cdot}\,\bar{\boldsymbol{y}}_0}{\phi_i}\,[\boldsymbol{H}^{-1}]_{\cdot i}, \tag{44}$$

with $\phi_i \to 0$ collapsing the prior toward exact cointegration in direction $\boldsymbol{H}_{i\cdot}$ and $\phi_i \to \infty$ removing it.

Adaptation to the asymmetric conjugate prior. The PLR was originally derived for the symmetric Normal–Inverse-Wishart prior of (9). I extend the dummy-observation construction to the per-equation NIG marginal likelihood of (13) so that all priors operate on the same augmented data, and the joint marginal-likelihood optimization remains internally consistent. This resolves the implementation issue that had previously made PLR incompatible with the Chan (2022) per-equation prior structure I use in estimation.

Joint optimization. The PLR tightness vector $\boldsymbol{\phi}$ is optimized *jointly* with the Minnesota hyperparameters $(\kappa_1, \kappa_2, \lambda_3)$ at the system level by maximizing the closed-form marginal likelihood under (13) on the dummy-augmented data; the per-equation persistence parameters δ_i that would otherwise be system-level (and discarded once the per-equation grid in (15) fires) are held at loose defaults during this stage. The resulting structure—a Chan asymmetric conjugate base augmented with PLR dummies, with both stages selected jointly via marginal likelihood—accommodates the mixed persistence of my system (an $I(0)$ shock instrument, near-$I(1)$ macroeconomic aggregates, and $I(0)$ distributional factors) without requiring the researcher to take a stand on individual unit root tests, which tend to have low power (Sims and Zha, 1998).

Auto-detection of $\boldsymbol{H}$. I construct $\boldsymbol{H}$ data-adaptively: each variable is classified $I(0)$ or $I(1)$ via an augmented Dickey–Fuller test *with deterministic trend* (MacKinnon (1994) critical value -3.41 at 5%), and pairs of $I(1)$ variables whose first difference is itself stationary are encoded as cointegrating rows of $\boldsymbol{H}$ in the form $\boldsymbol{e}_i - \boldsymbol{e}_j$. The trend-included specification is essential: the no-trend ADF can spuriously reject the unit root for series with strong deterministic components by absorbing the trend into the intercept.

ADF classification with deterministic trend. Table 6 reports the trend-included ADF t-statistic for each of the 14 endogenous variables. The 5 % MacKinnon (1994) critical value is -3.41. Variables with $t < -3.41$ are classified $I(0)$; the rest are $I(1)$ candidates for the cointegration step.

Table 6: ADF unit-root tests (with deterministic trend), baseline VAR

Variable	t-stat	Class	Row interpretation in $\boldsymbol{H}$
`poil`	-2.03	$I(1)$	identity
`oilprod`	-3.08	$I(1)$	cointegrating: `oilprod` $-$ `oilstocksM_BH`
`worldip`	-2.20	$I(1)$	identity
`usip`	-1.44	$I(1)$	identity
`log_cpi`	-4.06	$I(0)$	identity (trend-stationary)
`oilstocksM_BH`	-2.57	$I(1)$	identity
`dist_f1`	-7.40	$I(0)$	identity (stationary)
`dist_f2`	-6.91	$I(0)$	identity (stationary)
`dist_f3`	-7.28	$I(0)$	identity (stationary)
`dist_f4`	-5.58	$I(0)$	identity (stationary)
`dist_f5`	-5.32	$I(0)$	identity (stationary)
`dist_f6`	-5.13	$I(0)$	identity (stationary)
`dist_f7`	-5.45	$I(0)$	identity (stationary)
`dist_f8`	-5.13	$I(0)$	identity (stationary)

Note: ADF specification: $\Delta y_t = \alpha + \delta\, t + \gamma\, y_{t-1} + \sum_{j=1}^{p^*} \beta_j\, \Delta y_{t-j} + \varepsilon_t$, with $p^* = \lfloor (T-1)^{1/3} \rfloor$ capped at 12. MacKinnon (1994) 5 % critical value -3.41. The trend term is essential for log-level macroeconomic series with non-zero drift; without it, the regression absorbs the deterministic trend into the intercept and over-rejects the unit root for variables with persistent upward levels (e.g., `oilstocksM_BH` returns $t = -4.11$ under the no-trend specification).

$\boldsymbol{H}$ matrix. The auto-built $\boldsymbol{H}$ is 15×15 with two cointegrating rows; the rest are identity. The leading detected pair is

$$\boldsymbol{e}_{\texttt{oilprod}} - \boldsymbol{e}_{\texttt{oilstocksM\_BH}},$$

which combines into $\log(\texttt{oilprod}) - \log(\texttt{oilstocksM\_BH}) = -\log(\text{days of cover})$, the negative log of the OECD inventory-to-production ratio (bounded by storage capacity and operational requirements).

Optimal tightness vector ϕ. Joint marginal-likelihood optimization over $(\kappa_1, \kappa_2, \lambda_3, \phi_1, \ldots, \phi_{15})$ at the system level is run separately for each shock identification, since the shock equation dif-

fers across the BH and KZ runs and the joint marginal likelihood depends on all 15 equations:

Table 7: Optimal PLR tightness ϕ_i^* under BH and KZ identifications, baseline VARs

Row of $\boldsymbol{H}$	BH ϕ_i^*	KZ ϕ_i^*
shock instrument	62.06	**14.11**
`poil`	90.75	79.65
`oilprod`	99.55	97.53
`worldip`	99.87	99.35
`usip`	98.08	99.94
`log_cpi`	99.93	99.17
`oilstocksM_BH`	98.88	99.37
`dist_f1`	75.14	49.56
`dist_f2`	54.37	25.38
`dist_f3`	32.94	89.05
`dist_f4`	42.05	83.98
`dist_f5`	**0.37**	43.86
`dist_f6`	12.89	47.17
`dist_f7`	27.72	31.57
`dist_f8`	93.05	75.03

Note: Search range $\phi_i \in [0.01, 100]$ on a log grid; optimization via differential evolution (Storn and Price, 1997) on the closed-form Chan marginal likelihood, dummy-augmented per (44). Larger ϕ_i corresponds to a looser long-run prior on row i of $\boldsymbol{H}$. Optimal Minnesota hyperparameters: $(\kappa_1, \kappa_2, \lambda_3) = (0.156, 0.022, 0.430)$ for BH and $(0.116, 0.017, 0.361)$ for KZ. Bold entries flag the rows where the PLR is most binding under each shock.

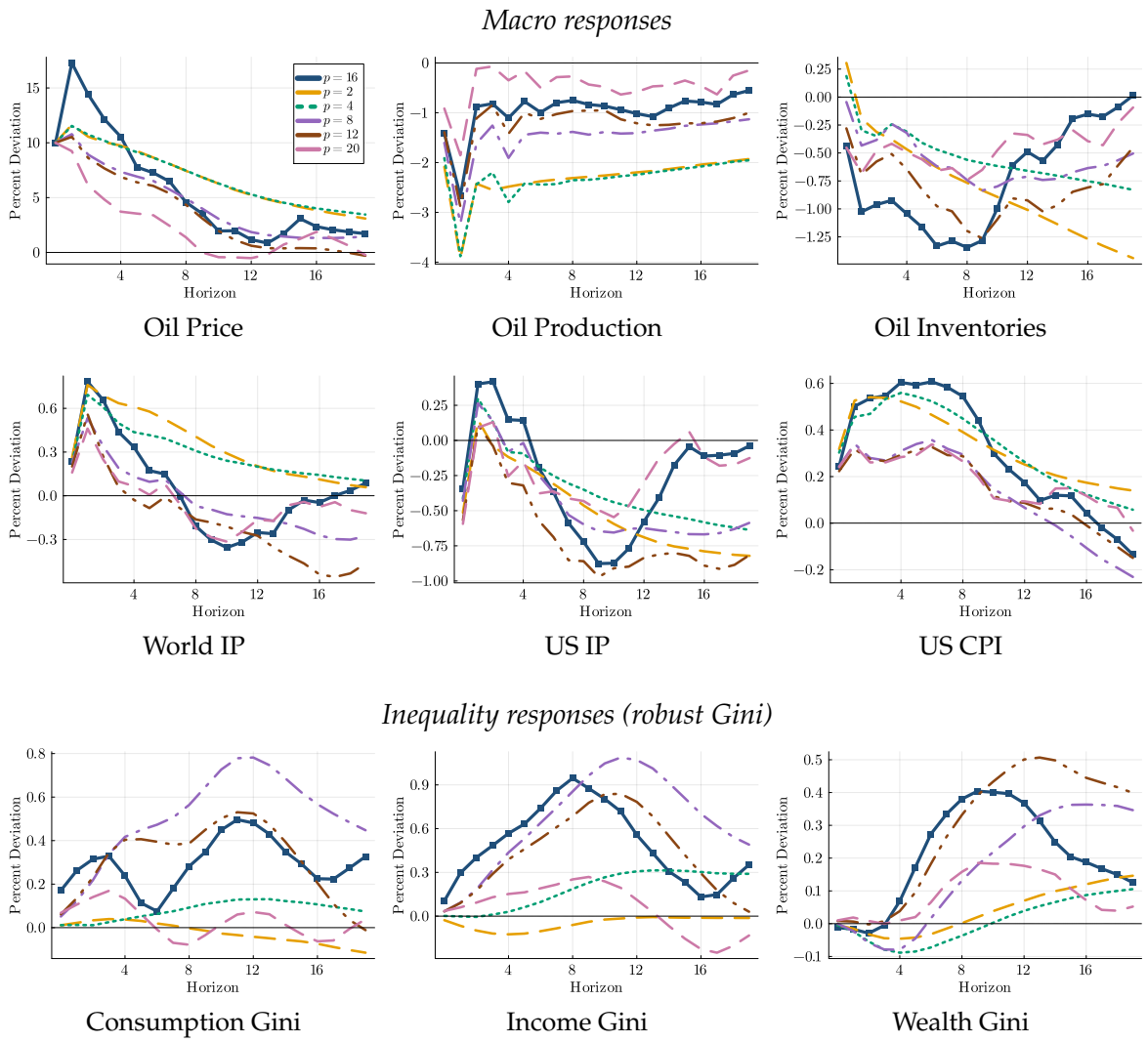
Interpretation. Three patterns stand out. First, the macro block is essentially non-binding under both shocks: `poil` sits in the loose-to-upper-bound range and the remaining five macro rows hover at the upper bound, indicating that the long-run prior contributes negligible additional shrinkage once the per-equation Chan persistence $\delta_i \approx 1$ is in force. The cointegrating row (`oilprod` $-$ `oilstocksM_BH`) receives a corner ϕ under both identifications: the auto-detected pair is interpretable as a stationary log-days-of-cover ratio (Working, 1949; Pindyck, 2001) but not statistically binding in this sample.[27] Second, the shock-instrument row is shock-specific by construction and the data prefer different long-run shrinkage on it: $\phi^*_{\text{KZ}} = 14.11$ is meaningfully tight, $\phi^*_{\text{BH}} = 62.06$ is moderate. Third, the distributional block is where the PLR does real work, and the binding row depends on the shock: $\phi^*_{D_5} = 0.37$ under BH (very tight); under KZ the dist factors take moderate values throughout, with $\phi^*_{D_2} = 25$ the tightest. For smooth, near-stationary latent factors with small steady-state means, the cointegration-style prior provides informative shrinkage that complements the per-equation Minnesota structure—exactly the asymmetry the Chan (2022) design philosophy is meant to deliver, and one that the data sharpen differently across shocks.

[27] Kilian and Murphy (2014a) report no cointegration on a pre-2009 monthly sample. I leave the auto-detected pair in the prior because the post-2009 quarterly regime makes the long-run ratio empirically more stable, and the PLR is non-binding in this row regardless.

H Robustness: Lag Order

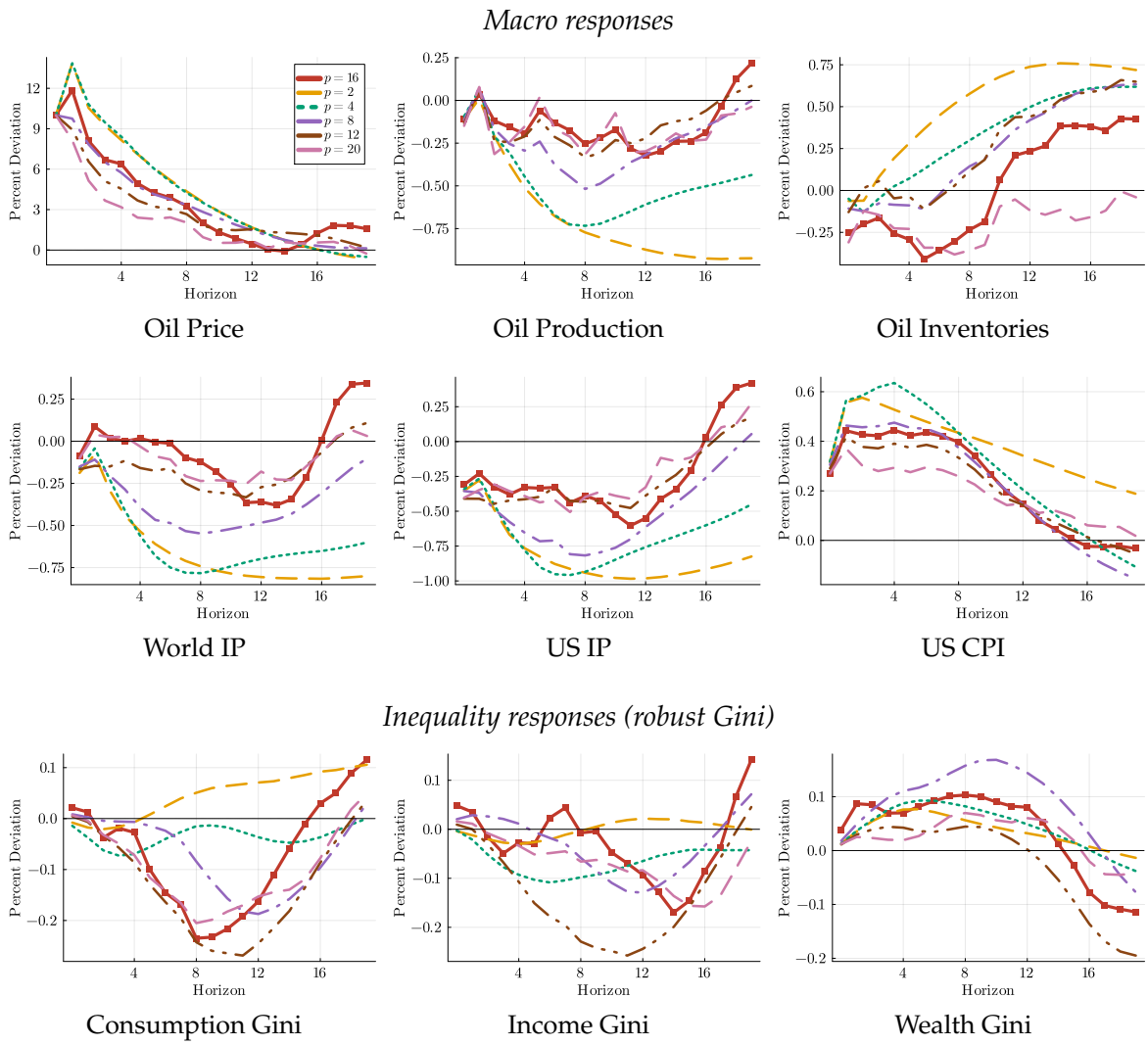
We vary the VAR lag length $p \in \{2, 4, 8, 12, 20\}$ around the baseline ($p = 16$). Shorter lags risk omitted dynamics; longer lags strain the data.

Figure 33: Lag length sweep — Baumeister and Hamilton (2019) supply shock



Notes: Each panel: baseline (solid, shock-aware color, $p = 16$) + five dashed lines for $p \in \{2, 4, 8, 12, 20\}$. No credible bands. Macro responses are level IRFs.

Figure 34: Lag length sweep — Känzig (2021) supply news shock



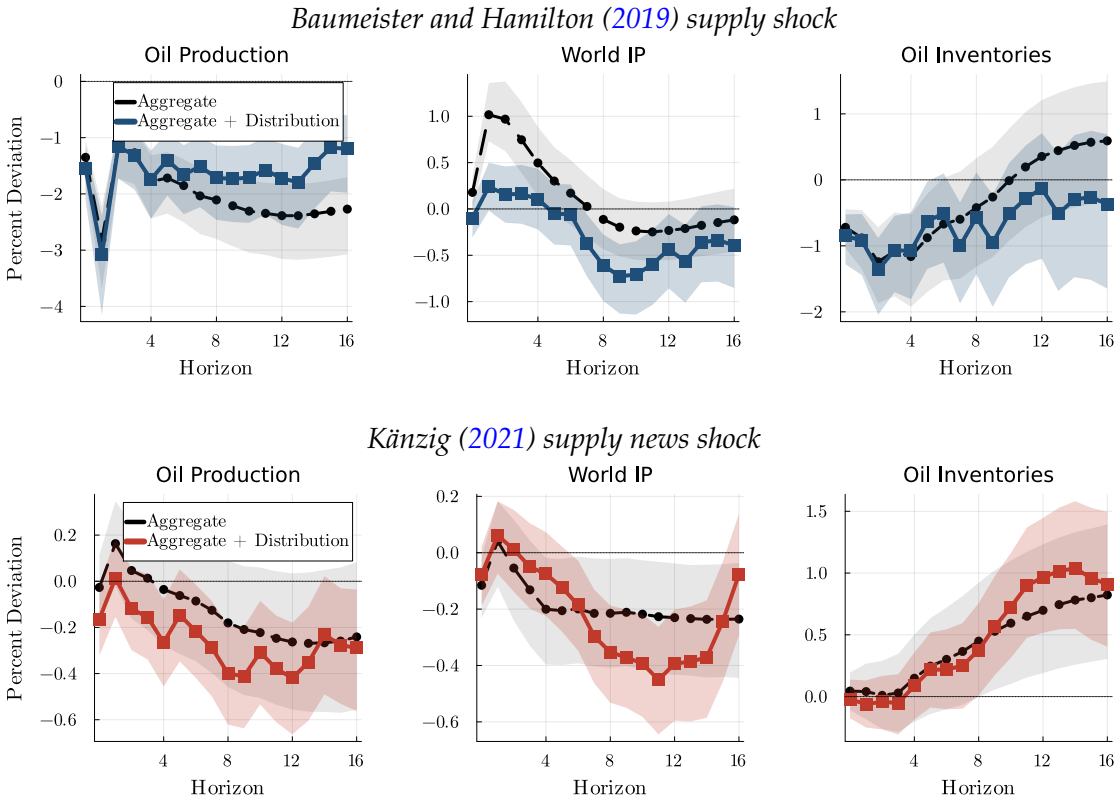
Notes: Each panel: baseline (solid, shock-aware color, $p = 16$) + five dashed lines for $p \in \{2, 4, 8, 12, 20\}$. No credible bands. Macro responses are level IRFs.

I IRFs from Rotating FAVAR

This appendix reports impulse responses from the rotating FAVAR exercise described in Section 4. Each FAVAR has the form $[\varepsilon^{\text{oil}}, \log p^{\text{oil}}, F_1, \ldots, F_7, x_t]$, where $F_1, \ldots, F_7$ are the seven stationary FRED-QD common factors (McCracken and Ng, 2021) and x_t is one of $|\mathcal{X}|$ macro and financial variables that enters the VAR ordered last. The shock is identified via the same internal-instrument scheme as the baseline VAR, and the lag length, prior, and posterior draws all match the baseline specification (Chan asymmetric conjugate prior, $p = 16$, 5,000 posterior draws, PLR enabled).

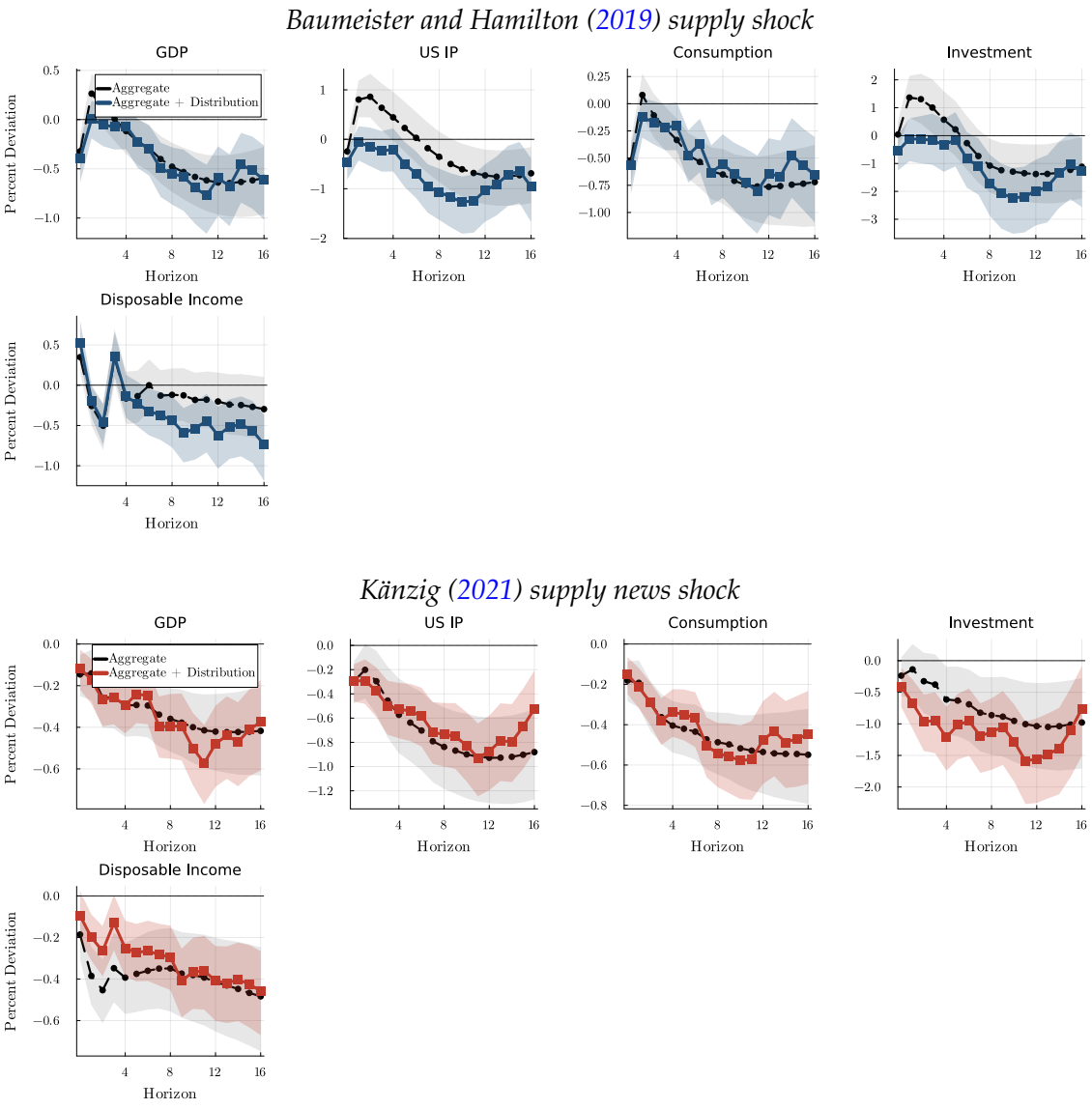
To gauge whether including the joint distribution alters the implied macro responses, each panel below overlays two IRFs for the same rotating variable: a solid line from the *aggregate-only* FAVAR (specification above), and a dashed line from a *distribution-augmented* FAVAR that appends the eight smoothed distributional factors $\boldsymbol{\xi}_t$ to the same aggregate block, $[\varepsilon^{\text{oil}}, \log p^{\text{oil}}, F_1, \ldots, F_7, \boldsymbol{\xi}_t, x$ Both lines come with 68% credible bands. The two shocks are color-coded: blue for the Baumeister and Hamilton (2019) realised supply shock and red for the Känzig (2021) supply news shock; within a figure, the dashed (distribution-augmented) line uses a darker shade of the same family. Variables are grouped into ten economic channels.

Figure 35: Rotating FAVAR — oil market



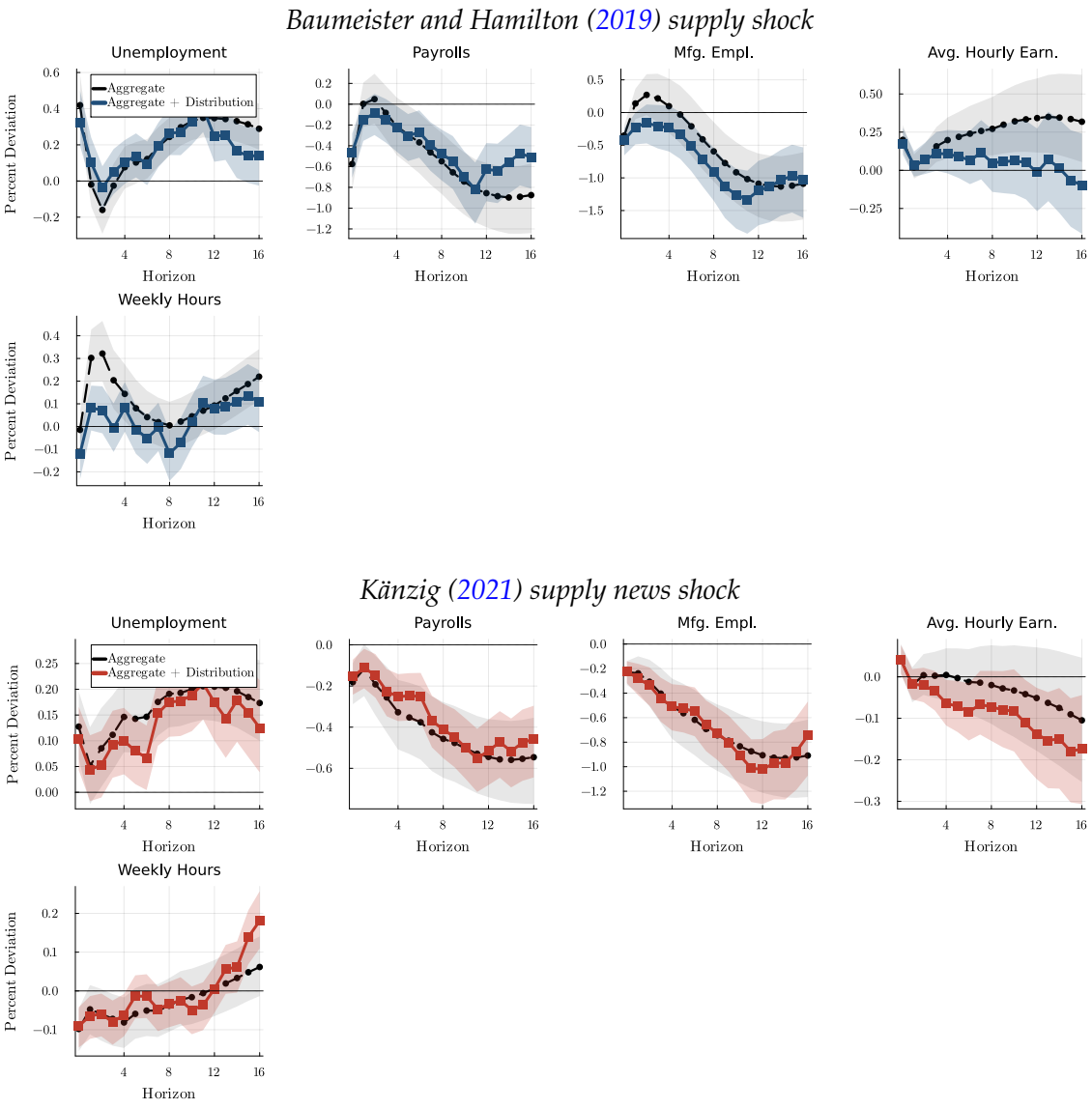
Notes: Each panel shows two IRFs for the named variable: solid line — aggregate-only FAVAR; dashed line — FAVAR augmented with the eight smoothed distributional factors $\boldsymbol{\xi}_t$. Shaded regions are 68% credible sets.

Figure 36: Rotating FAVAR — real economic activity



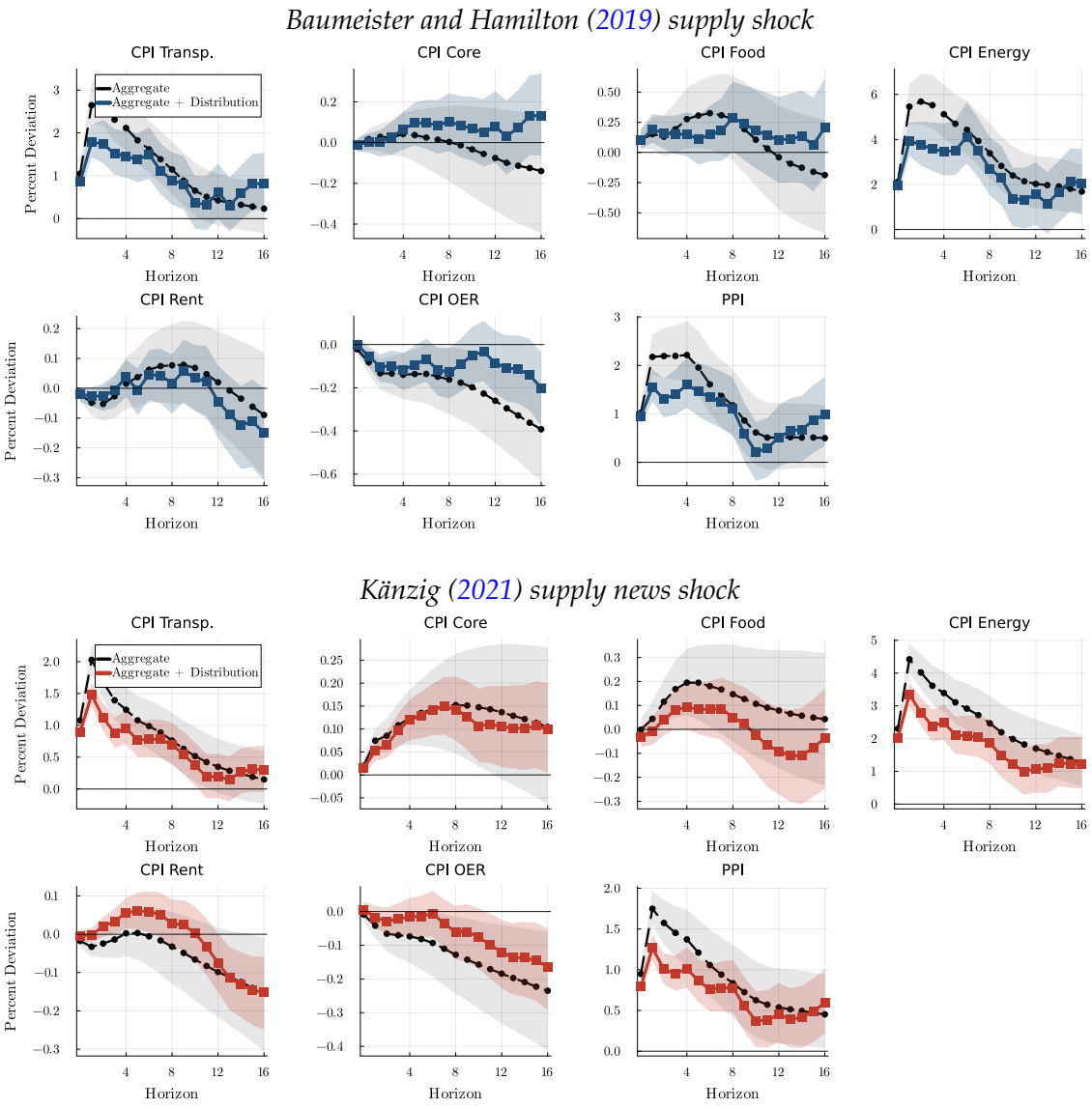
Notes: Solid: aggregate-only FAVAR. Dashed: FAVAR augmented with the eight smoothed distributional factors. Shaded regions are 68% credible sets.

Figure 37: Rotating FAVAR — labor market



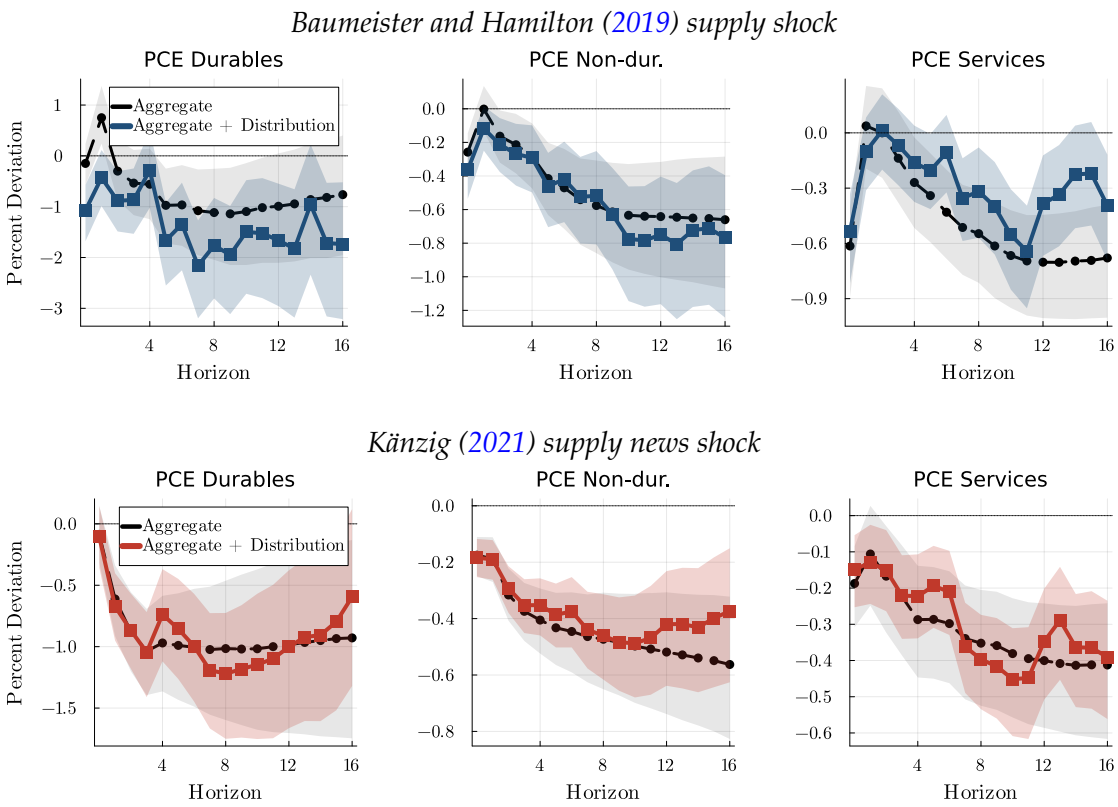
Notes: Solid: aggregate-only FAVAR. Dashed: FAVAR augmented with the eight smoothed distributional factors. Shaded regions are 68% credible sets.

Figure 38: Rotating FAVAR — consumer prices



Notes: Solid: aggregate-only FAVAR. Dashed: FAVAR augmented with the eight smoothed distributional factors. Disaggregated CPI components (food, energy, rent, transportation, OER). Shaded regions are 68% credible sets.

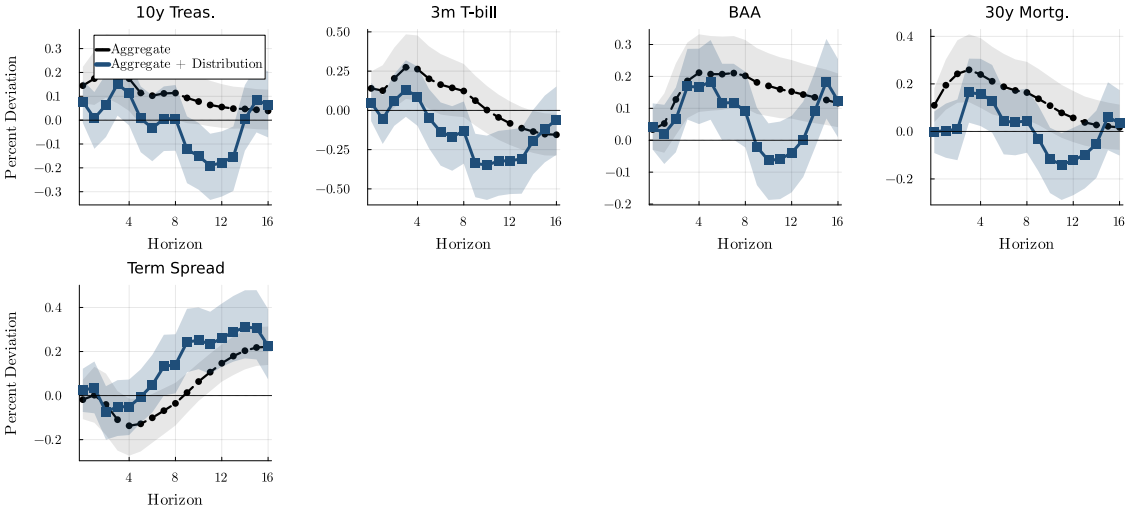
Figure 39: Rotating FAVAR — consumption



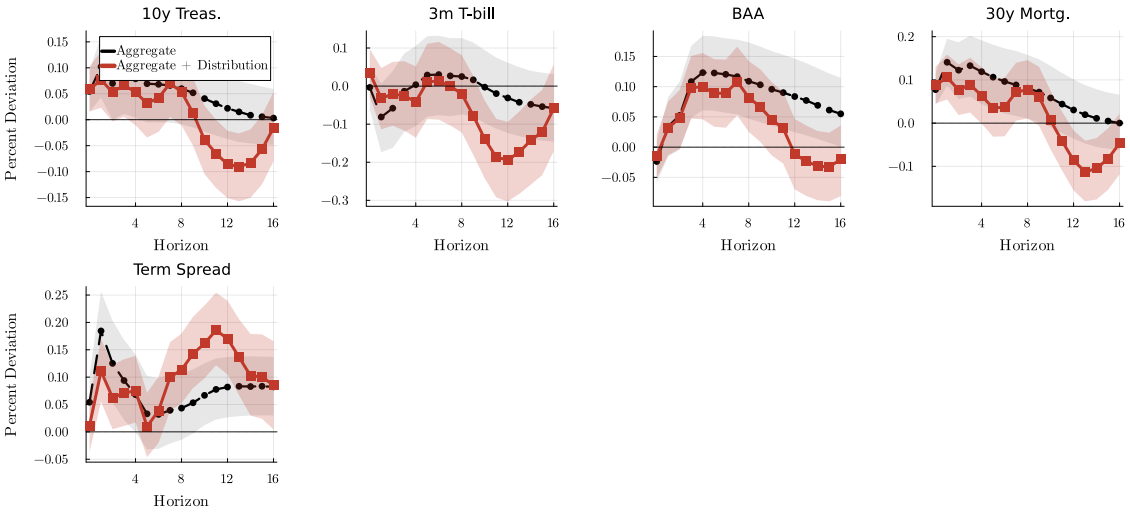
Notes: Solid: aggregate-only FAVAR. Dashed: FAVAR augmented with the eight smoothed distributional factors. PCE components (durables, non-durables, services). Shaded regions are 68% credible sets.

Figure 40: Rotating FAVAR — interest rates

Baumeister and Hamilton (2019) supply shock

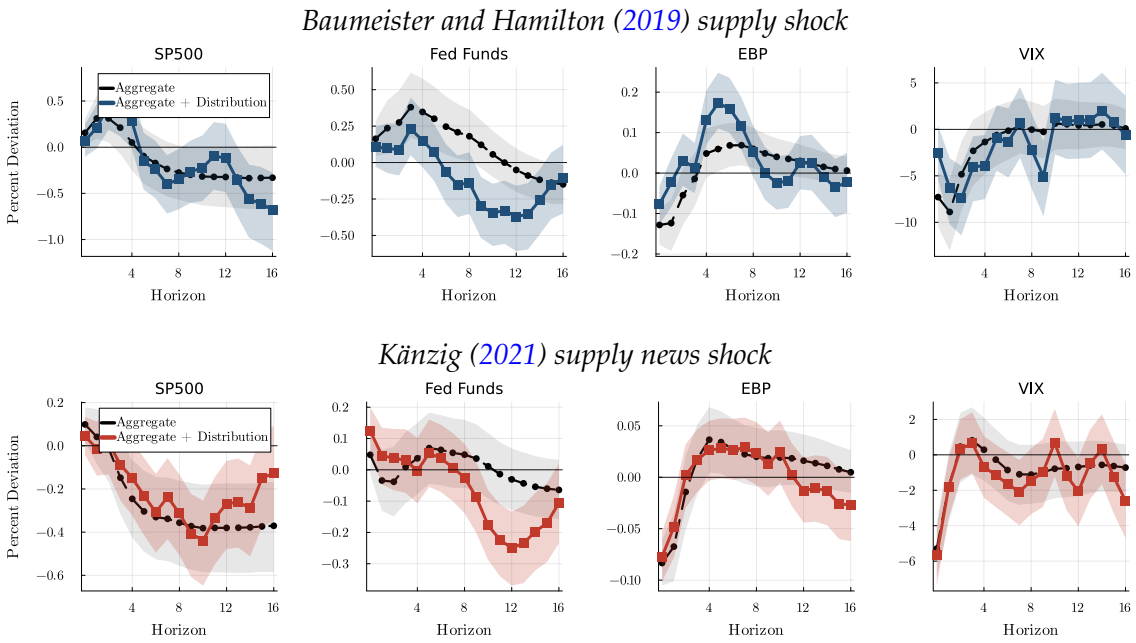


Känzig (2021) supply news shock



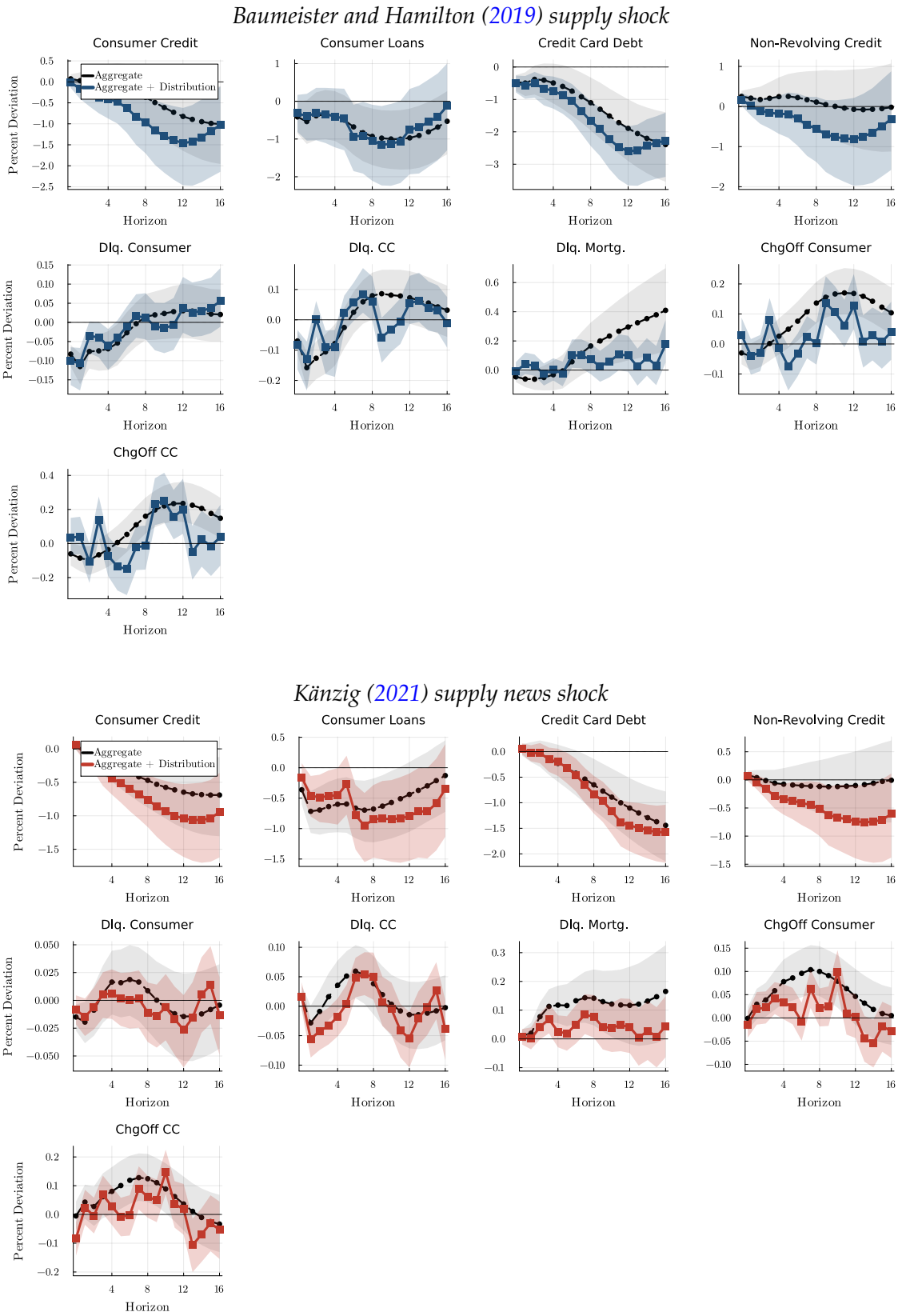
Notes: Solid: aggregate-only FAVAR. Dashed: FAVAR augmented with the eight smoothed distributional factors. Treasury, federal funds, and corporate yields. Shaded regions are 68% credible sets.

Figure 41: Rotating FAVAR — financial conditions



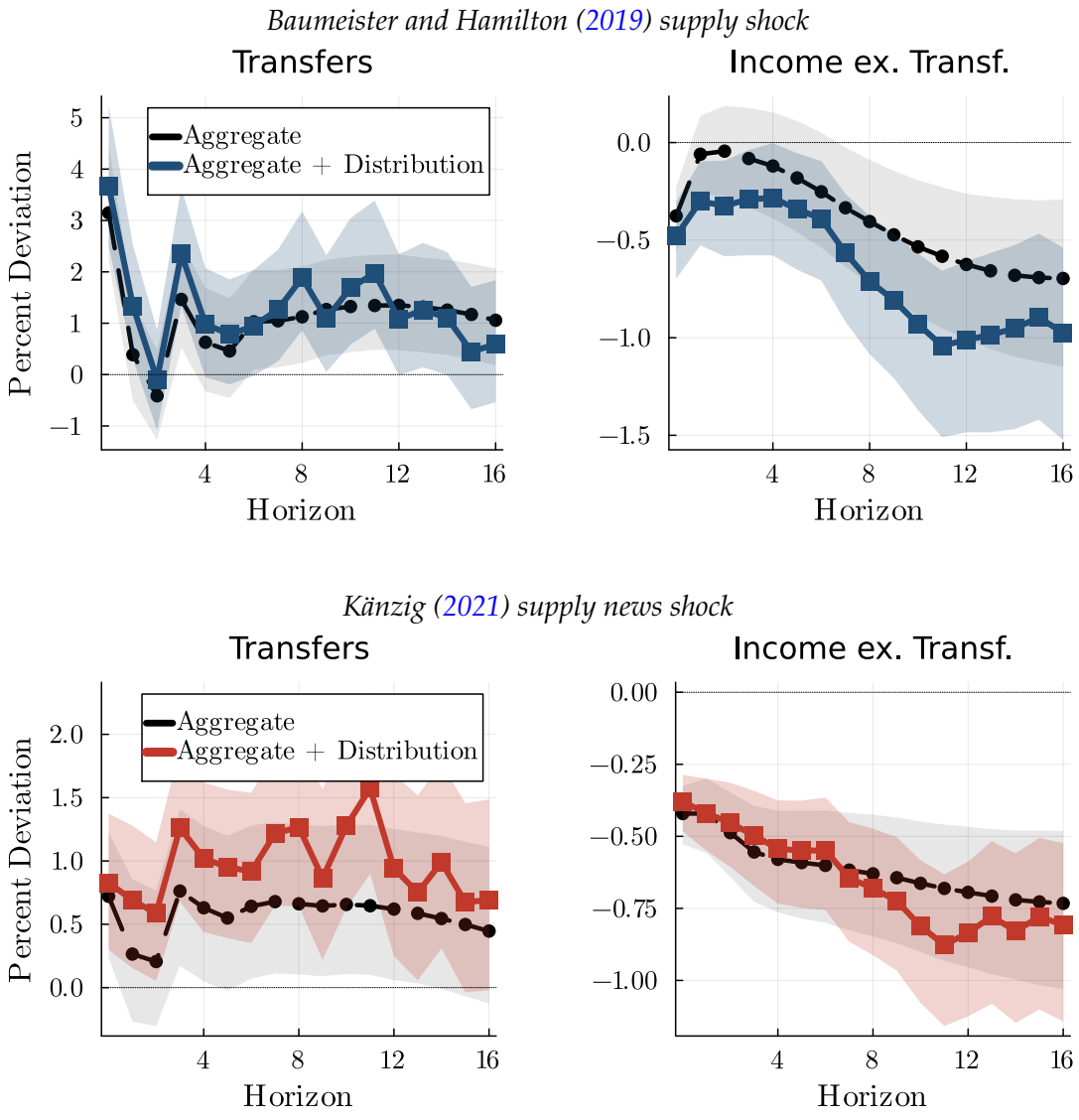
Notes: Solid: aggregate-only FAVAR. Dashed: FAVAR augmented with the eight smoothed distributional factors. S&P 500, VIX, Excess Bond Premium, federal funds rate. Shaded regions are 68% credible sets.

Figure 42: Rotating FAVAR — household credit supply and delinquency



Notes: Solid: aggregate-only FAVAR. Dashed: FAVAR augmented with the eight smoothed distributional factors. Consumer credit and consumer-loan stocks; delinquency and charge-off rates on consumer loans, credit-card balances, and mortgages. Shaded regions are 68% credible sets.

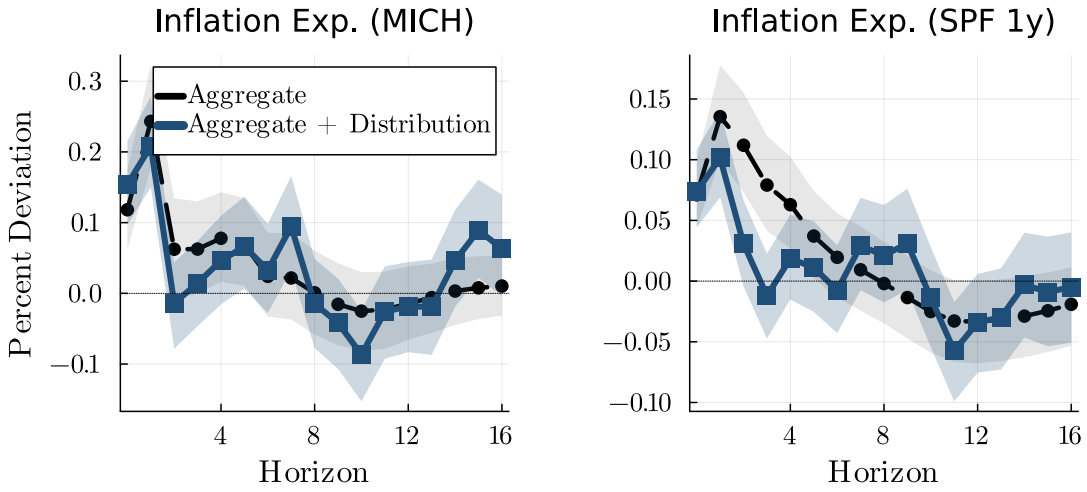
Figure 43: Rotating FAVAR — income composition



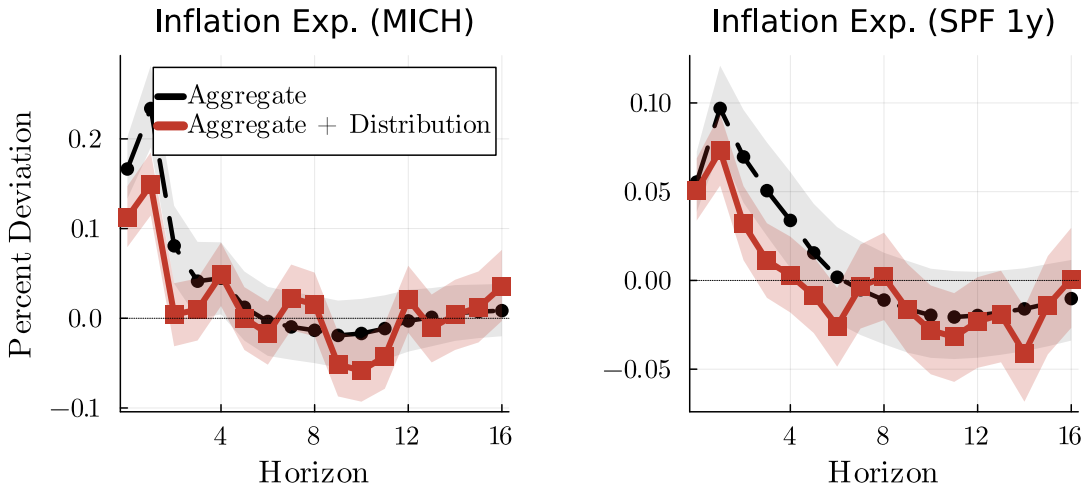
Notes: Solid: aggregate-only FAVAR. Dashed: FAVAR augmented with the eight smoothed distributional factors. Government transfers and personal income net of transfers. Shaded regions are 68% credible sets.

Figure 44: Rotating FAVAR — inflation expectations

Baumeister and Hamilton (2019) supply shock

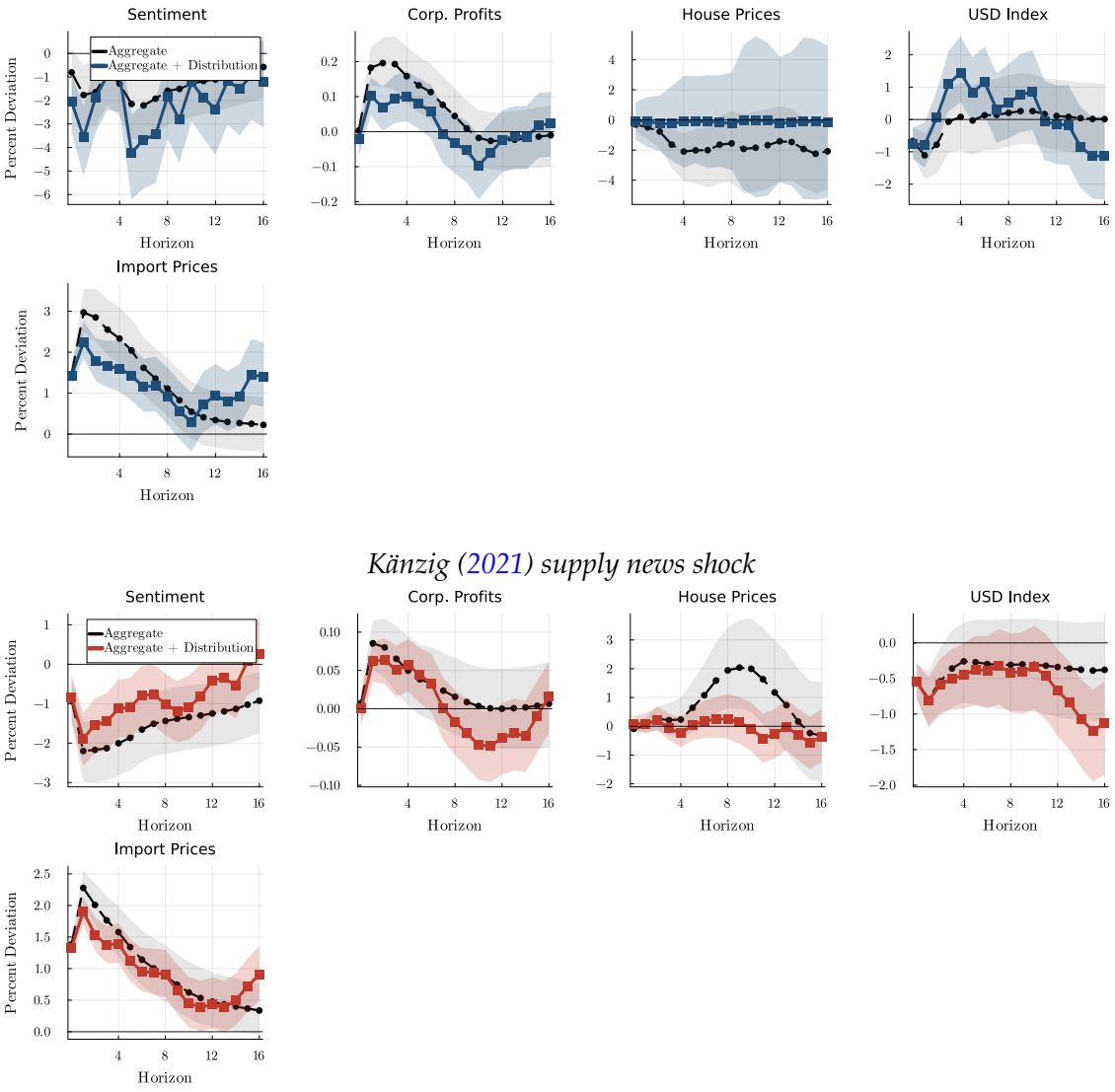


Känzig (2021) supply news shock



Notes: Solid: aggregate-only FAVAR. Dashed: FAVAR augmented with the eight smoothed distributional factors. Inflation expectations: University of Michigan 1-year household survey and SPF 1-year CPI forecasters. Shaded regions are 68% credible sets.

Figure 45: Rotating FAVAR — other



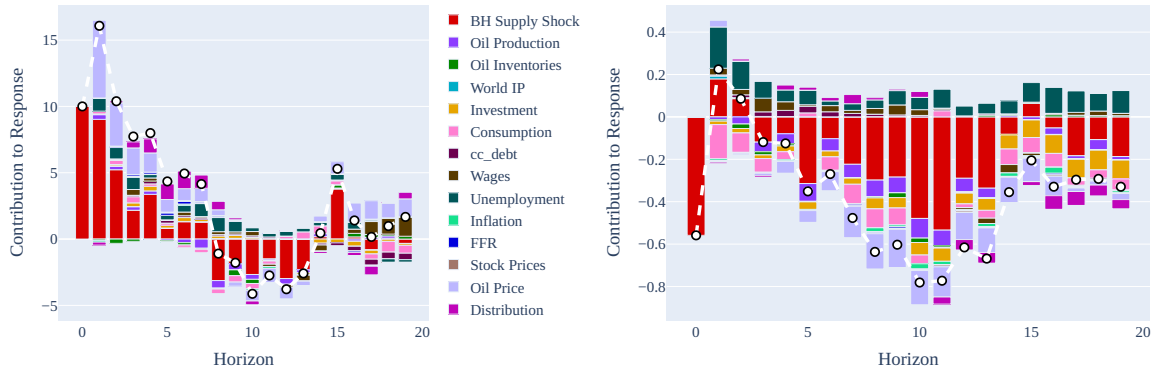
Notes: Solid: aggregate-only FAVAR. Dashed: FAVAR augmented with the eight smoothed distributional factors. Other macro and financial aggregates not classified above (e.g., trade-weighted dollar, corporate profits, house prices). Shaded regions are 68% credible sets.

J Channel Decomposition under the Extended Macro VAR

This appendix reports the impact-horizon ($n = 0$) channel decomposition of the oil-price and aggregate-consumption responses under the extended-macro specification, which augments the baseline VAR with a richer US block (`inv`, `cons`, `cc_debt`, `wages`, `unemp`, `infl`, `ffr`, `sp500`) plus the oil-market block. The decomposition follows the same Dufour–Wang construction used for the baseline (Section 5) but operates on the larger 12-variable macro block, so additional channels (credit, labor, monetary, financial) appear alongside the oil-market and inequality channels.

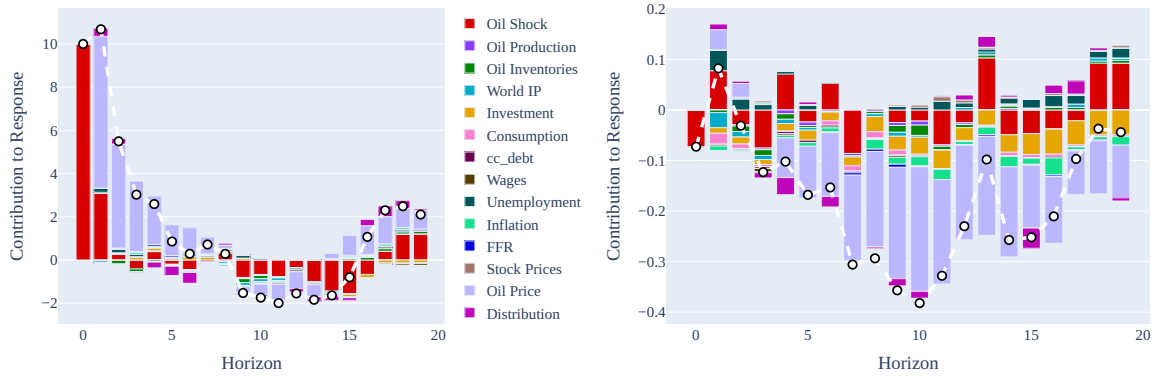
Figure 46: Channel decomposition at $n = 0$, extended-macro VAR — oil price (`poil`) and aggregate consumption (`cons`)

Panel A: Baumeister and Hamilton (2019) realized supply shock



Oil price (`poil`) Aggregate consumption (`cons`)

Panel B: Känzig (2021) oil supply news shock



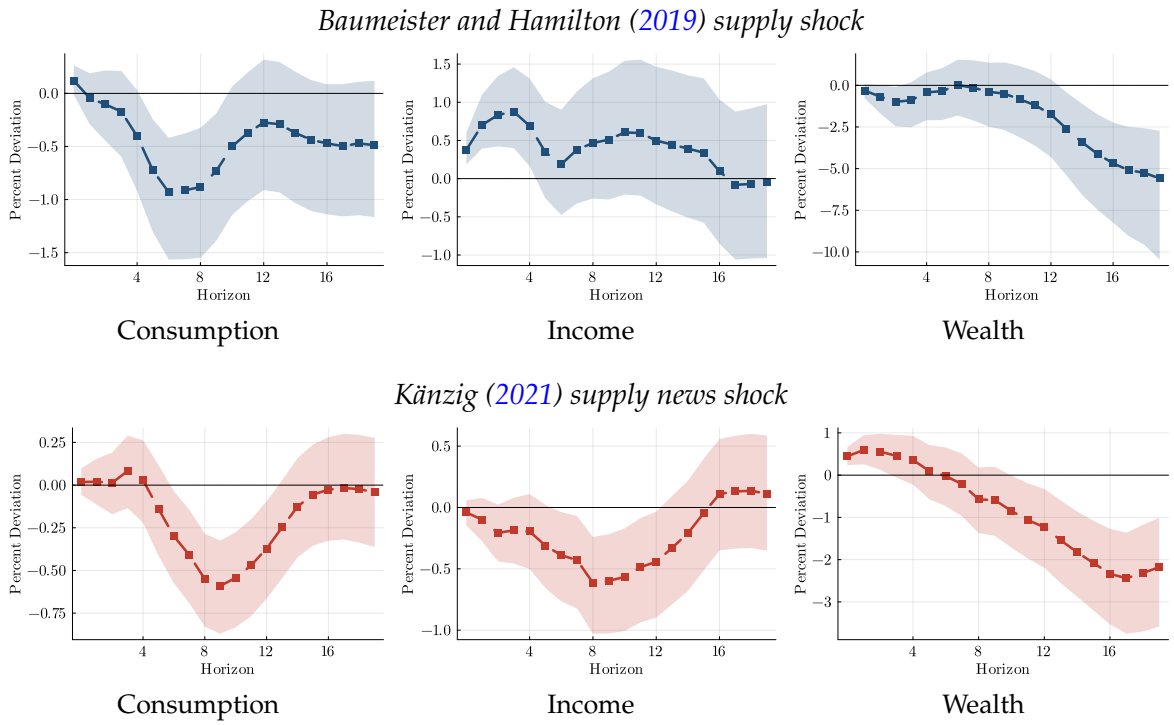
Oil price (`poil`) Aggregate consumption (`cons`)

Notes: Channel contributions $c_{m\to i}^{(h,0)}$ at impact ($n = 0$) under the extended-macro VAR. Each colored bar is the total contribution of channel m to the response of the outcome i at horizon h, evaluated from impact. Channels are grouped following the propagation classification in Section 4: oil market (oil price, oil production, inventories, world activity), real activity (investment, consumption, wages, unemployment), prices (PCE inflation), monetary/financial (federal funds rate, S&P 500), credit (`cc_debt`), and the eight smoothed distributional factors. Bars sum to the total impulse response at each horizon by construction.

K Additional Inequality Measures

This appendix complements the Gini-coefficient responses reported in Section 6 with three additional summary measures of dispersion: the log $90/10$ and $90/50$ percentile ratios and the standard deviation of logs. Each is reported for the consumption, income, and wealth marginals and for both identification schemes. The percentile ratios summarize between-tail spread (a rise in $90/10$ indicates a widening tail) while the standard deviation of logs summarizes overall dispersion.

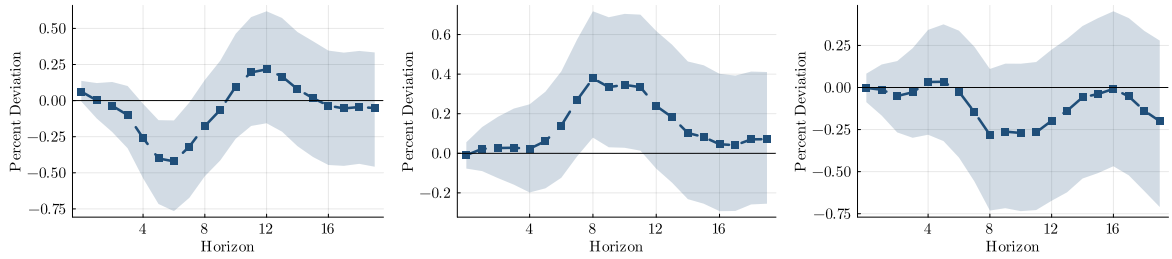
Figure 47: Log $90/10$ percentile ratio



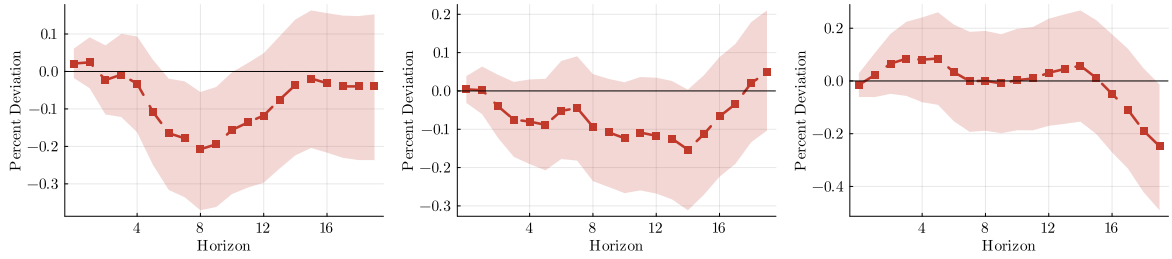
Notes: Impulse responses of the log $90/10$ percentile ratio for consumption, income, and wealth. Units: percentage point deviation from steady state. Solid lines: posterior median. Shaded areas: 68% credible sets.

Figure 48: Log 90/50 percentile ratio

Baumeister and Hamilton (2019) supply shock



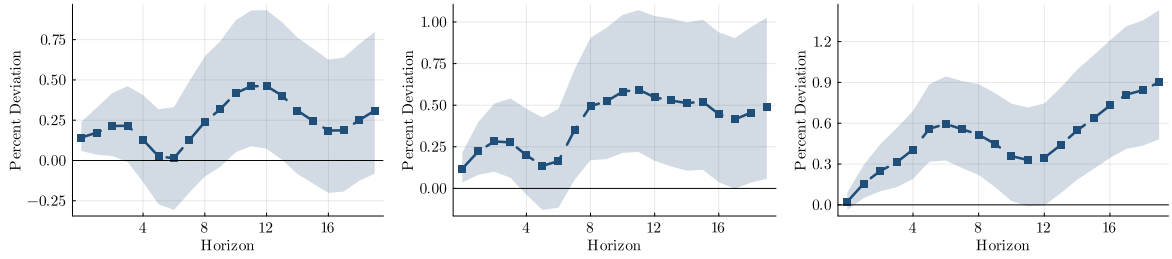
Känzig (2021) supply news shock



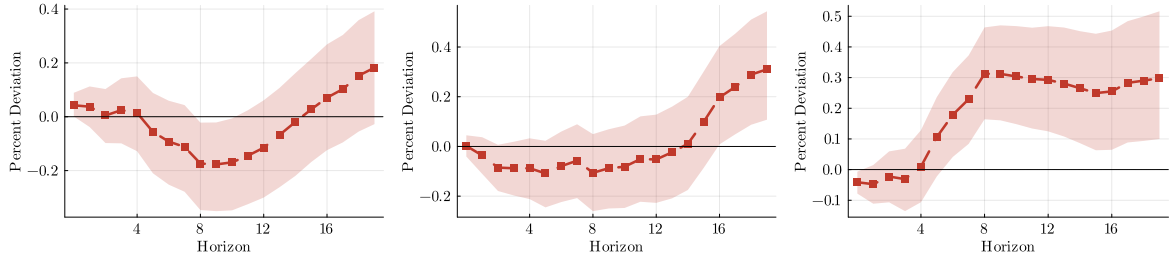
Notes: Impulse responses of the log 90/50 percentile ratio for consumption, income, and wealth. Units: percentage point deviation from steady state. Solid lines: posterior median. Shaded areas: 68% credible sets.

Figure 49: Standard deviation of logs

Baumeister and Hamilton (2019) supply shock



Känzig (2021) supply news shock



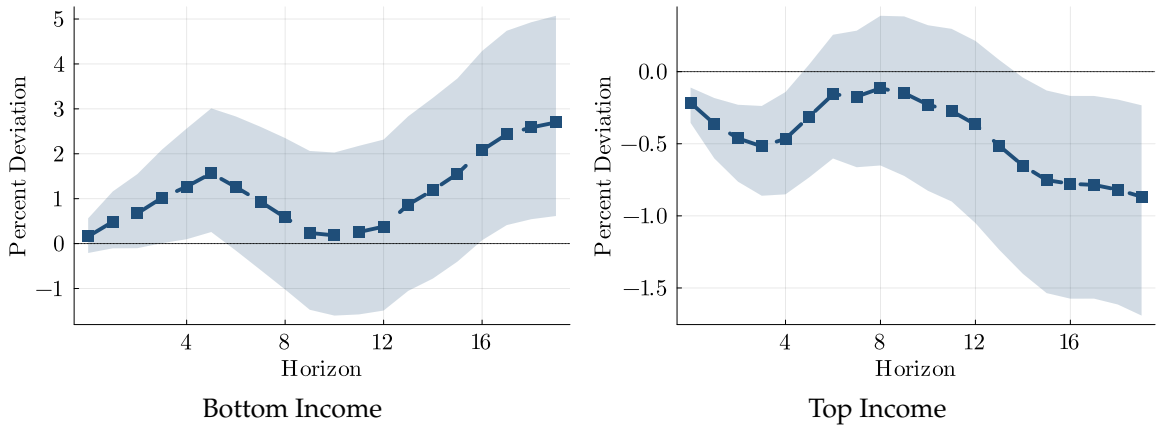
Notes: Impulse responses of the standard deviation of logs for consumption, income, and wealth. Units: percentage point deviation from steady state. Solid lines: posterior median. Shaded areas: 68% credible sets.

L Consumption Responses for the Distribution Tails

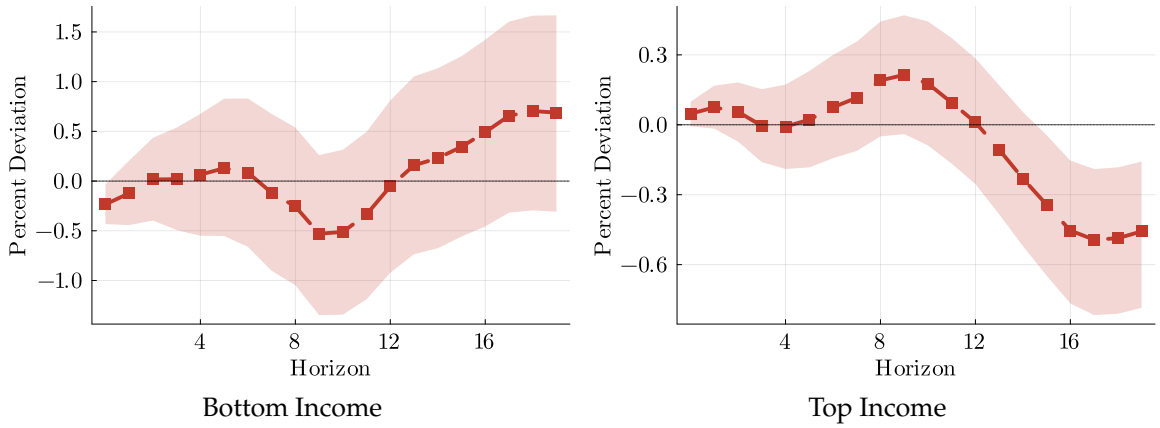
Figures 9 and 10 in the body report consumption responses by the Low (0–50%), Middle (50–90%), and Top (90–100%) groups of the income and wealth distributions. This appendix complements those plots with the responses of households at the *tails* of the marginal distributions: the Bottom (0–20%) and the Top (90–100%) of each margin. The qualitative picture extends naturally from the broad cuts, but the magnitudes are sharper at the tails, particularly for the bottom-income and top-wealth groups.

Figure 50: Consumption responses by extreme positions in the income distribution

Panel A: Baumeister and Hamilton (2019) realized supply shock



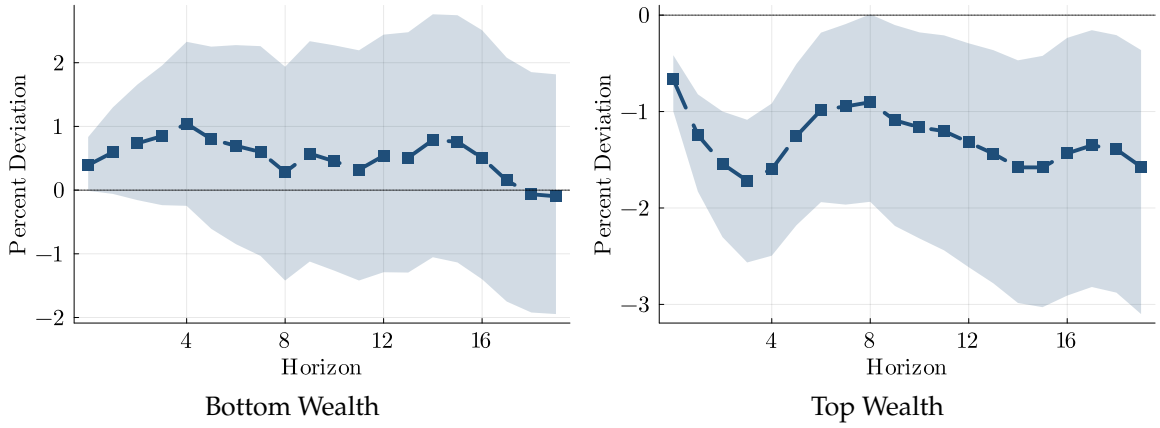
Panel B: Känzig (2021) oil supply news shock



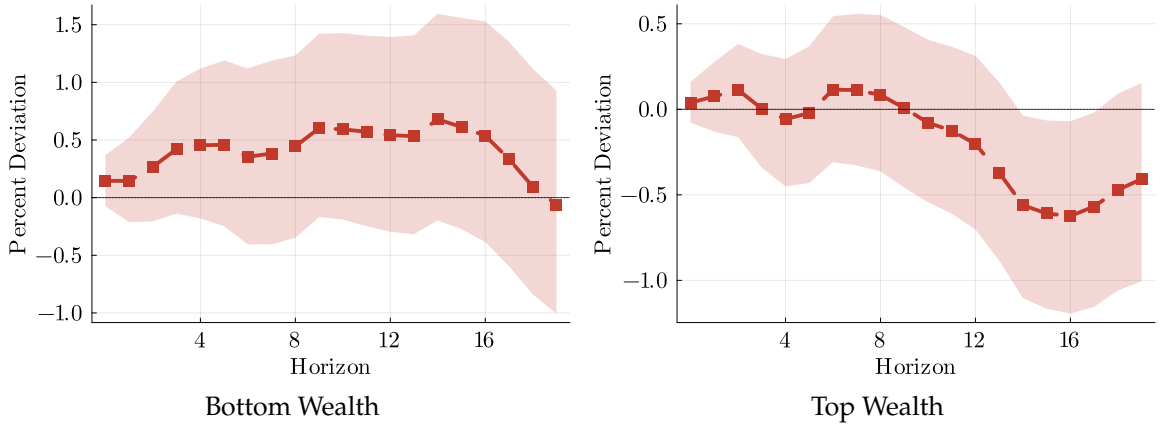
Notes: Impulse responses of mean consumption for households at the bottom and top of the marginal income distribution. Solid lines: posterior median. Shaded areas: 68% credible sets. Horizon in quarters.

Figure 51: Consumption responses by extreme positions in the wealth distribution

Panel A: Baumeister and Hamilton (2019) realized supply shock



Panel B: Känzig (2021) oil supply news shock



Notes: Impulse responses of mean consumption for households at the bottom and top of the marginal wealth distribution. Solid lines: posterior median. Shaded areas: 68% credible sets. Horizon in quarters.