Modern Economy, 2025, 16(7), 1091-1108
https://www.scirp.org/journal/me
ISSN Online: 2152-7261
ISSN Print: 2152-7245

A Bayesian Vector Autoregression Model to Estimate Bilateral Trade Flows in Panel Data Setting

Prabuddha Sanyal[1*], Mark Ehlen[2]

[1]Straightdeal Mortgage, Huntington Beach, CA, USA
[2]Prism Analytics Corporation, Albuquerque, NM, USA
Email: *prabuddha@straightdealmortgage.com

How to cite this paper: Sanyal, P., & Ehlen, M. (2025). A Bayesian Vector Autoregression Model to Estimate Bilateral Trade Flows in Panel Data Setting. *Modern Economy, 16*, 1091-1108.
https://doi.org/10.4236/me.2025.167052

Received: February 23, 2025
Accepted: July 21, 2025
Published: July 24, 2025

Abstract

This study examines how trade shocks that start in one large economy ripple through other countries and how long those effects stick around. Using quarterly bilateral-trade data for 2012-2023, the authors estimate a Bayesian four-country VAR model (China, Germany, Japan, United States) identified by sign restrictions. Impulse-response functions show that a one-standard-deviation drop in Chinese exports cuts German and Japanese exports by about three percent on impact, whereas German shocks are one-third as large, Japanese shocks quickly rebound, and U.S. shocks barely travel abroad. Forecast-error variance decomposition at a ten-quarter horizon confirms China's central role: its shocks explain roughly one-third of the medium-run export volatility in Germany and Japan, while 90 percent of U.S. volatility remains home-grown. Persistence analysis finds that own shocks in China, Germany, and Japan stay near full strength for five years, whereas U.S. shocks fade faster.

Keywords

Bilateral Trade, Bayesian Analysis, Vector Autoregression

1. Introduction

The canonical gravity framework posits that bilateral trade flows rise with economic mass and decline with geographic distance. Yet, operationalizing this ostensibly parsimonious relationship reveals a series of empirical challenges. The "distance" term embodies far more than physical mileage; it conflates freight-corridor topologies, colonial and linguistic affinities, supply-chain interdependencies, and even the availability of direct air travel. Once these latent dimensions are disentangled, the estimated distance elasticity becomes markedly unstable. A further complication arises from the pervasive incidence of zero trade observations, often exceeding half of all country pairs in a given year. This renders the logarithmic specification undefined. Researchers must therefore adopt Poisson pseudo-maximum-likelihood estimators, zero-inflated alternatives, or Bayesian treatments merely to obtain consistent estimates.

Additional hurdles stem from multilateral resistance effects: a policy shock such as a bilateral tariff reduction between United States and China propagates through the entire network of trading partners. Empirically, accounting for these general-equilibrium spillovers necessitates exhaustive sets of exporter-year and importer-year fixed effects, whose sheer dimensionality can rival or surpass the sample size, heightening concerns over overfitting. Key trade-policy numbers—such as tariffs, indexes of non-tariff barriers, and measures of port delays come out late, are often imprecise, and are themselves shaped by political pressures rather than being purely outside influences. Mitigating these issues forces the analyst into a regime of instrumental-variable strategies, hierarchical priors, and heteroskedasticity-robust clustering. The gravity model is a mainstay of trade analysis for its pedagogical clarity, yet it also poses a continual econometric challenge because credible estimation demands far more empirical rigor than its elegant theory suggests.

In this study, we provide a novel approach to address this issue of bilateral trade flows, where there are lot of zero trade flows between countries. We consider only the top 25 countries in trade flows over the period 2012 to 2023. Then, we impose a Bayesian Vector Autoregression model with two lags. We provide an example where this method can be implemented for a sample of four countries. This method can thus be suitable for country pairs, where zero trade can be potentially challenging issue in econometric estimation.

This study is organized as follows: In section 2, we provide a brief overview of the existing literature on bilateral trade flows using panel data models in a gravity modeling framework. Section 3 provides theoretical formulation and empirical model estimated. Section 4 provides the main results of the study, while section 5 concludes.

2. A Brief Overview of the Literature

Santos Silva and Tenreyro (2006) were among the first studies that reevaluated the standard practice of estimating gravity equations by ordinary least squares (OLS) on log-linearized trade data. The authors show that this approach is inconsistent whenever the error term is heteroskedastic (an empirical regularity in trade datasets), and it cannot accommodate the pervasive zero-trade observations that arise for more than 50 percent of country pairs. To remedy these shortcomings, the study advocates the Poisson pseudo-maximum-likelihood (PPML)[1] estima-

[1]Poisson pseudo-maximum likelihood (PPML) is a regression estimator that treats the conditional mean of the dependent variable as Poisson but allows the data to be any non-negative distribution. It is consistent under heteroskedasticity, handles zero trade flows naturally, and is estimated by maximizing a Poisson log-likelihood with robust (sandwich) standard errors.

tor, which models bilateral trade in levels rather than logs. Monte Carlo simulations confirm that PPML yields unbiased and consistent parameter estimates under a wide range of heteroskedastic error structures, while log-OLS exaggerates the sensitivities of trade to distance, GDP, colonial ties, and preferential-trade agreements.

Applying PPML to both traditional and Anderson and E. van Wincoop (2003) augmented gravity specifications[2], the authors document materially smaller elasticities of trade with respect to GDP and distance than those reported in the earlier literature and find that the effects of colonial links and trade agreements are close to zero. Because many policy assessments—such as counterfactual tariff analyses or welfare calculations—depend directly on these elasticities, the results imply that previous studies relying on log-OLS may have overstated the gains from trade agreements and the costs of distance. The paper therefore recommends PPML (or similar count-data techniques) as the preferred estimator for gravity models, stressing that robust inference about trade determinants is essential for sound policy design .

Egger and Nigai (2015) investigates how incorporating firm-level heterogeneity and multilateral resistance terms affects gravity-model estimates of bilateral trade flows. Using a large panel of manufacturing trade among OECD economies, the study develops a structural gravity specification that embeds Melitz-type productivity dispersion within a general-equilibrium framework (Melitz, 2003).[3] Identification relies on exporter-year and importer-year fixed effects to purge common shocks, while bilateral fixed effects capture time-invariant pair-specific frictions. The author estimates the model with Poisson pseudo-maximum likelihood to accommodate zero trade and heteroskedasticity, and then contrasts the results with conventional log-linear OLS and panel-FE estimators.

The preferred estimates reveal markedly lower distance elasticities and higher implied trade costs for small, less-productive firms than for large exporters, highlighting how aggregate gravity coefficients mask underlying firm-level margins. Counterfactual simulations indicate that trade-cost reductions disproportionately raise the export participation of mid-productive firms, thereby amplifying welfare gains beyond what a representative-firm model would predict. From a policy perspective, the findings underscore that ignoring heterogeneity can lead to underestimation of both the extensive-margin response to trade liberalization and the resulting welfare improvements, suggesting that trade agreements targeting fixed

[2]Anderson-van Wincoop-augmented gravity" is just the gravity equation equipped with multilateral-resistance controls, typically implemented via exporter-year and importer-year fixed effects (or their cross-sectional analogues) and estimated with PPML. That seemingly small tweak is what makes the empirical gravity model consistent with the micro-founded theory of trade costs and general equilibrium.

[3]Melitz-type productivity dispersion refers to the log-normal (or Pareto-like) spread of firm productivities first formalized by Melitz (2003). In that model each firm draws a productivity parameter from a continuous distribution—most firms have average (or low) productivity, while a thin upper tail of highly productive firms dominates exports. The variance (shape) of this distribution governs the minimum productivity cut-off for survival in a market.

export costs may yield larger aggregate benefits than previously thought.

Baltagi et al. (2015) provide an extensive summary to reassess the "gravity model" of international trade—an empirical workhorse that explains bilateral trade flows by the "mass" of trading partners (GDP, population) and the "distance" between them (geographic, cultural, institutional). The authors survey and implement modern econometric refinements that tackle longstanding problems: unobserved country-pair heterogeneity, the abundance of zero trade observations, and endogeneity of trade costs. They compare Poisson pseudo-maximum-likelihood (PPML), high-dimensional fixed effects, and instrumental-variables strategies, stressing that failure to control these issues can severely distort estimates of distance-related trade frictions and policy variables such as tariffs or free-trade agreements.

Applying these techniques to a large panel of world trade, the paper shows that once proper controls and identification strategies are in place, traditional gravity determinants remain robust but their magnitudes change meaningfully. In particular, unobserved bilateral factors (history, networks, supply chains) explain a sizable share of trade, while policy variables—preferential trade agreements, WTO membership, common currency—have effects that are smaller yet still economically and statistically significant. The authors conclude that well-specified gravity models provide more credible guidance for policymakers: they yield better forecasts of trade-creation versus trade-diversion, help identify the true gains from regional integration and improve cost–benefit assessments of future trade negotiations and infrastructure investments.

In the next section, we provide the theoretical framework for estimating a Bayesian Vector Autoregression model of bilateral trade with two lags in a panel data setting and provide the empirical context of this model that was implemented in this study.

3. Modeling Framework

Bayesian VAR (2) models offer a disciplined way to capture the rich, bidirectional dynamics that characterize bilateral trade flows while avoiding the over-parameterization that plagues classical VARs in panel settings. With two lags, the model is flexible enough to trace short-run propagation of shocks—such as tariff changes, exchange-rate swings, or demand shifts—across trading partners, yet parsimonious enough to remain computationally tractable when multiplied over dozens (or hundreds) of country pairs.

The Bayesian framework imposes sensible prior distributions (e.g., Minnesota or hierarchical shrinkage priors) that "shrink" the multitude of coefficients toward economically plausible values, stabilizing estimates when time-series length is short relative to the cross-section. This regularization sharply reduces estimation noise, guards against spurious impulse-response patterns, and delivers predictive densities that naturally quantify uncertainty.

Equally important, Bayesian VARs integrate seamlessly with panel-data structures that acknowledge heterogeneity across country pairs while still borrowing strength from the entire sample. Hierarchical priors allow coefficients to vary around a common mean, letting heavily traded dyads (e.g., USA and China) inform sparsely observed ones (e.g., Chile and Thailand) without forcing them to be identical. This partial pooling improves efficiency, mitigates small-sample bias from zero or infrequent trade observations, and accommodates latent factors—such as global supply-chain shifts or synchronized business cycles—via cross-equation covariance priors. By producing joint posterior distributions of all parameters and shocks, Bayesian VAR (2) models facilitate coherent structural analysis (e.g., sign-restricted or narrative identification) and scenario forecasting, delivering policy-relevant insights that classical fixed-effects gravity regressions or single-equation error-correction models cannot match.

3.1. Theoretical Framework

We represent the Vector form of the VAR (2) model as follows:

$$\boldsymbol{Y}_t = \boldsymbol{A}_1\boldsymbol{Y}_{t-1} + \boldsymbol{A}_2\boldsymbol{Y}_{t-2} + \boldsymbol{B}\boldsymbol{X}_t + \boldsymbol{u}_t,\ \ \boldsymbol{u}_t \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma}\right). \tag{1}$$

The component definitions are:

$$\begin{array}{ll} \boldsymbol{Y}_t = \left[y_{1t}, y_{2t}, \cdots, y_{Nt}\right]' & \left(\text{endogenous } N\times 1 \text{ vector}\right), \\ \boldsymbol{X}_t = \left[x_{1t}, x_{2t}, \cdots, x_{Kt}\right]' & \left(\text{exogenous } K\times 1 \text{ vector}\right), \\ \boldsymbol{A}_1, \boldsymbol{A}_2 \in \mathbb{R}^{N\times N} & \left(\text{lag matrices}\right), \\ \boldsymbol{B} \in \mathbb{R}^{N\times K} & \left(\text{exogenous-effect matrix}\right), \\ \boldsymbol{\Sigma} \in \mathbb{R}^{N\times N}, \boldsymbol{\Sigma} = \boldsymbol{\Sigma}' > 0 & \left(\text{covariance of shocks}\right). \end{array} \tag{2}$$

The companion (state-space form) is represented by the following system:

$$\underbrace{\begin{bmatrix} \boldsymbol{Y}_t \\ \boldsymbol{Y}_{t-1} \end{bmatrix}}_{\tilde{\boldsymbol{Y}}_t} = \underbrace{\begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{I}_N & \boldsymbol{0} \end{bmatrix}}_{\boldsymbol{A}_c} \underbrace{\begin{bmatrix} \boldsymbol{Y}_{t-1} \\ \boldsymbol{Y}_{t-2} \end{bmatrix}}_{\tilde{\boldsymbol{Y}}_{t-1}} + \underbrace{\begin{bmatrix} \boldsymbol{B} \\ \boldsymbol{0} \end{bmatrix}}_{\boldsymbol{B}_c} \boldsymbol{X}_t + \underbrace{\begin{bmatrix} \boldsymbol{u}_t \\ \boldsymbol{0} \end{bmatrix}}_{\tilde{\boldsymbol{u}}_t}. \tag{3}$$

The stacked form for estimation of Equation (3) is given as follows:

$$\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\beta} + \boldsymbol{u},\ \ \boldsymbol{u} \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{I}_T \otimes \boldsymbol{\Sigma}\right) \tag{4}$$

with

$$\begin{aligned} \boldsymbol{y} &= \text{vec}\left(\left[\boldsymbol{Y}_3, \cdots, \boldsymbol{Y}_T\right]\right), \\ \boldsymbol{Z} &= \boldsymbol{I}_{T-2} \otimes \left[\boldsymbol{Y}'_{t-1}, \boldsymbol{Y}'_{t-2}, \boldsymbol{X}'_t\right], \\ \boldsymbol{\beta} &= \text{vec}\left(\left[\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{B}\right]\right). \end{aligned}$$

The Normal-Inverse-Wishart prior (conjugate) is given by:

$$\boldsymbol{\beta} \mid \boldsymbol{\Sigma} \sim \mathcal{N}\left(\boldsymbol{\beta}_0, \boldsymbol{\Sigma} \otimes \boldsymbol{V}_0\right),\ \ \boldsymbol{\Sigma} \sim \mathcal{W}^{-1}\left(\boldsymbol{S}_0, \nu_0\right) \tag{5}$$

The analytical posterior would thus be given by the following expressions:

$$\begin{aligned} \boldsymbol{\Sigma} \mid \text{data} &\sim \mathcal{W}^{-1}\left(\boldsymbol{S}_n, \nu_n\right), \\ \boldsymbol{\beta} \mid \boldsymbol{\Sigma}, \text{data} &\sim \mathcal{N}\left(\boldsymbol{\beta}_n, \boldsymbol{\Sigma} \otimes \boldsymbol{V}_n\right), \end{aligned} \tag{6}$$

With

$$\boldsymbol{V}_n = \left(\boldsymbol{V}_0^{-1} + \boldsymbol{Z}'\boldsymbol{Z}\right)^{-1},$$
$$\boldsymbol{\beta}_n = \boldsymbol{V}_n\left(\boldsymbol{V}_0^{-1}\boldsymbol{\beta}_0 + \boldsymbol{Z}'\boldsymbol{y}\right),$$
$$\boldsymbol{S}_n = \boldsymbol{S}_0 + \left(\boldsymbol{y} - \boldsymbol{Z}\boldsymbol{\beta}_n\right)'\left(\boldsymbol{y} - \boldsymbol{Z}\boldsymbol{\beta}_n\right) + \left(\boldsymbol{\beta}_n - \boldsymbol{\beta}_0\right)'\boldsymbol{V}_0^{-1}\left(\boldsymbol{\beta}_n - \boldsymbol{\beta}_0\right),$$
$$\nu_n = \nu_0 + \left(T - 2\right).$$

The impulse-response function (IRF) at horizon h is:

$$\Psi_h = \begin{cases} I_N, & h = 0, \\ A_1\Psi_{h-1} + A_2\Psi_{h-2}, & h \geq 1. \end{cases} \tag{7}$$

Forecast error variance decomposition (FEVD):

$$\text{FEVD}_{i\leftarrow j}\left(H\right) = \frac{\sum_{h=0}^{H-1}\left(e_i'\Psi_h\Sigma e_j\right)^2}{\sum_{h=0}^{H-1} e_i'\Psi_h\Sigma\Psi_h' e_i},\ e_i = i^{\text{th}}\text{ selector vector} \tag{8}$$

For model comparison, we can use the Log-marginal likelihood function as follows:

$$\begin{aligned} &-\frac{TN}{2}\log\pi + \frac{\nu_0}{2}\log\left|\boldsymbol{S}_0\right| - \frac{\nu_n}{2}\log\left|\boldsymbol{S}_n\right| \\ &+\sum_{i=0}^{N-1}\left[\log\Gamma\left(\nu_n - i/2\right) - \log\Gamma\left(\nu_0 - i/2\right)\right] + \frac{1}{2}\log\left|\boldsymbol{V}_0\right| - \frac{1}{2}\log\left|\boldsymbol{V}_n\right|. \end{aligned} \tag{9}$$

3.2. Empirical Implementation

This section provides a framework for empirical implementation[4] of the above theoretical model step-by-step. This is given as follows:

Step 1: Data Layout

For every ordered country pair (i, j), we stack the two trade directions ~log exports and log imports into a (2 × 1) vector as follows:

$$y_{ij,t} = \begin{bmatrix} \ln\left(\text{Exports}_{ij,t}\right) \\ \ln\left(\text{Imports}_{ij,t}\right) \end{bmatrix} \tag{10}$$

All such vectors are piled into the tall block:

$$Y_t = \left[y_{12,t}, y_{13,t}, \cdots, y_{NN,t}\right]^\top \in \mathbb{R}^K,\ K = 2\left(N^2 - N\right) \tag{11}$$

The vector X_t collects bilateral frictions: geographic distance, common language, colonial history, RTA status, and the tariff rates.[5]

Step 2: The VAR Framework

[4]We do not provide the estimation results for the sake of brevity. These results can be obtained from the authors' upon reasonable request. Our preliminary results indicate that one-unit rise in linguistic proximity is associated with e1.77−1 ≈ 5.9e1.77−1 ≈ 5.9 times larger trade. Common_language retains a positive conditional effect (~ 0.38 log points, ≈ 46 % higher trade). Colonial ties is only marginally significant, while distance and the log GDP terms drop out because exporter & importer fixed effects soak up all their cross-sectional variation; with no meaningful within-country change, coefficients are numerically zero.

[5]We do not include the RTA status and the tariff rates since consistent information is not available for all country pairs.

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + BX_t + c_{nt},\ c_{nt} \sim \mathcal{N}(0,\Sigma) \tag{12}$$

where,

- A_1 and A_2 propagate lags of the entire network;
- B translates contemporaneous trade-cost shocks;
- Σ is the reduced-form covariance of the residuals.

Step 3: Country-aggregated illustration

A didactic four-country version rewrites the same idea at the macro level (e.g. CHN, USA, DEU, JPN):

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \varepsilon_t,\ \varepsilon_t \sim \mathcal{N}(0,\Sigma). \tag{13}$$

Step 4: Prior structure for Bayesian estimation

$$\Phi_{k\ell,t} \sim \mathcal{N}\left(1_{k=\ell}\rho^{\ell}, \lambda_1^{\ell}\sigma_k^2 \big/ \left(\ell^2\sigma_\ell^2\right)\right),\ \rho = 0.8,\ \lambda_1 = 0.20. \tag{14}$$

In Equation (14), the off-diagonal elements are shrunk towards zero; the diagonal elements start near ρ^{ℓ} and

$$\Sigma \sim \mathcal{W}^{-1}(S_0, \nu_0),\ \nu_0 = 6 \tag{15}$$

Equation (15) gives a loose inverse-Wishart prior on the covariance.

Step 5: Structural shocks via a Cholesky-rotation step

$$L = \mathrm{chol}(\Sigma),\ LL^{\top} = \Sigma,\ Q^{\top}Q = I,\ B = LQ,\ \varepsilon_t = Bu_t,\ u_t \sim \mathcal{N}(0,I). \tag{16}$$

The orthonormal matrix Q can help in exploring alternative shock orderings without re-estimating the entire VAR model.

Step 6: Moving-average representation and impulse responses

The companion matrix is given as follows:

$$A = \begin{bmatrix} \Phi_1 & \Phi_2 \\ 0 & I \end{bmatrix},\ \mathrm{IRF}(h) = A^h B. \tag{17}$$

Then, the cumulative effects through horizon h are:

$$\mathrm{CumIRF}(h) = \sum_{\tau=0}^{h} \mathrm{IRF}(\tau). \tag{18}$$

Step 7: Forecast-error variance decomposition (FEVD)

For variable i, and shock s, at horizon h, the share of forecast error variance attributed to the shock is:

$$\begin{aligned}&\mathrm{FEVD}_{is}(h)\\ &= \frac{\sum_{\tau=0}^{h}\left(e_i^{\top}A^{\tau}Be_s\right)^2}{\sum_{s'=1}^{4}\sum_{\tau=0}^{h}\left(e_i^{\top}A^{\tau}Be_{s'}\right)^2},\ 0 \le \mathrm{FEVD}_{is}(h) \le 1,\ \sum_{s=1}^{4}\mathrm{FEVD}_{is}(h) = 1.\end{aligned} \tag{19}$$

With the above building blocks, we can simulate impulse responses, trace spillovers along the global trade supply chains, and quantify how much of a bilateral trade flow's forecast error originates from a US tariff shock versus a Chinese demand shock. In the next section, we provide some of the model capabilities.

3.3. Data Sources and Measurement

Our data comes from four main sources. The bilateral trade flow data comes from the CEPII website (https://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=3). The real GDP data comes from the World Bank (https://data.worldbank.org/).[6] The colonial ties data comes from the Correlates of War Project (2025) at: https://correlatesofwar.org/data-sets/colonial-dependency-contiguity/. The language similarity data comes from the United States International Trade Commission at: https://www.usitc.gov/data/gravity/dicl.htm

Finally, the dyadic distance data also comes from CEPII at: https://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id = 6

All trade and macro variables are quarterly observations. Every step along the horizontal axis of an impulse-response curve represents one calendar quarter: $h = 0$ shows the immediate "on-impact" reaction in the same quarter the shock hits,

$h = 1$ is the response one quarter later,

$h = 4$ is one year later,

$h = 8$ is two years later, and so on.

The persistence ratio (as described in Appendix 2) is the heat-map, which simply compares how large the response is two years after the shock ($h = 8$) to how large it was one quarter after the shock ($h = 1$). A ratio greater than 1 means the effect has grown or stayed strong over time; a ratio below 1 means the effect has faded.

4. Main Results

This section provides the main results of this study. They are organized as follows: 1) The impulse response functions are first mapped of a one-standard deviation effect of trade shocks on each individual country; 2) Trade spillovers are mapped at the medium run ($h = 4$) of the row affected country to its own shocks and foreign shocks; 3) Forecast-Error Variance Decomposition (FEVD) is shown over a longer period ($h = 10$); (and 4) Pre and Post Covid Shocks to countries are developed. We consider only 4 countries (Germany, China, Japan, and USA) for the purpose of our illustration.

4.1. Impulse Response Functions

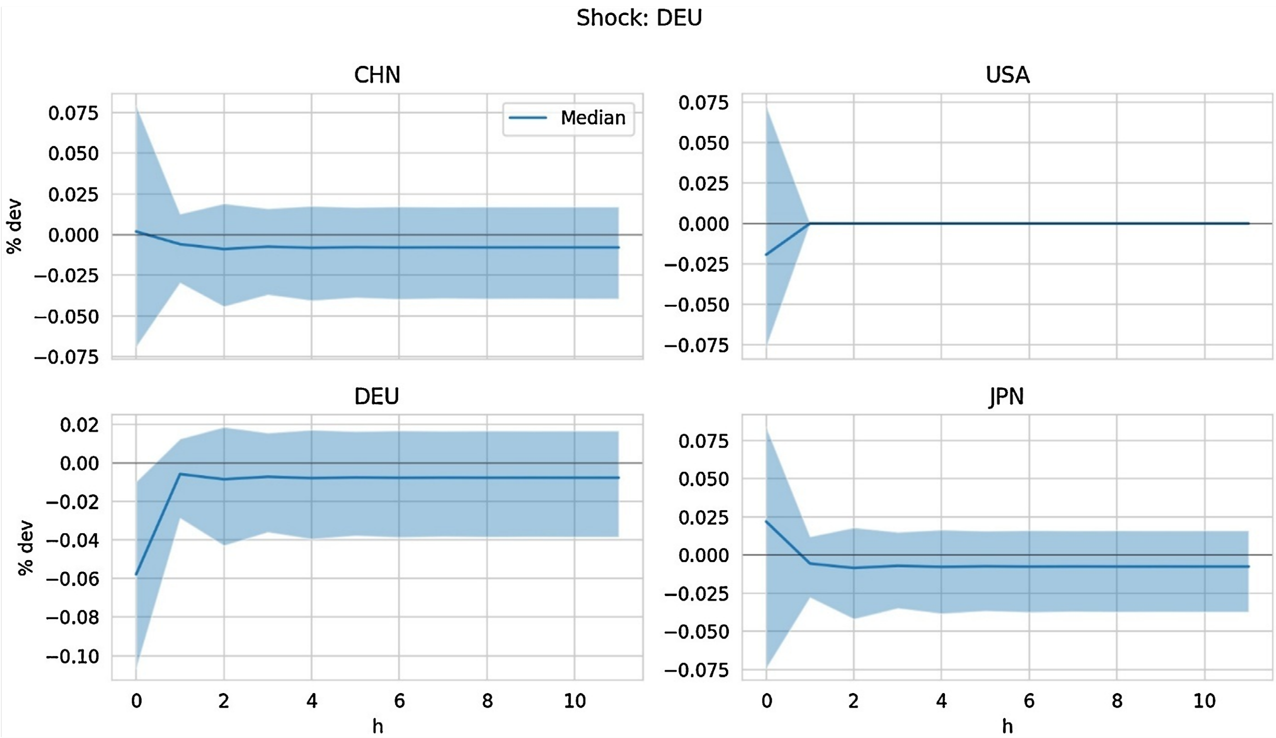
In this section, we demonstrate how the impulse response function (IRF) behaves with a structural trade shock of Germany to itself and to other countries. This is given in **Figure 1**.

An IRF traces the effect of a one-time, one-standard-deviation shock to one variable (here, German exports) on itself and other variables (exports of China, USA, Japan, and Germany) over time.

Mathematically, in a VAR (Vector Autoregression) system:

$$Y_t = A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + \epsilon_t \tag{20}$$

[6]We had initially chosen 50 countries in the sample. However, as the number of lagged parameter values A_1 and A_2 increases exponentially, this was very difficult to handle when running the estimation. For the purpose of illustration, we have chosen 4 countries to demonstrate our method, namely Germany, China, Japan, and USA.



Source: Authors' Own Computations.

Figure 1. Structural trade shocks from Germany (Period 2012-2023)

where, Y_t is a vector of export and import matrices, A_t are coefficient matrices, and ε_t are shocks, the IRF for horizon (h) is given by:

$$\mathrm{IRF}(h) = \frac{\partial Y_{t+h}}{\partial \epsilon_{DEU,t}} \tag{21}$$

What Do the Plots Show?

In the German-shock IRF panel, Germany's own exports fall abruptly by roughly six percent on impact and, although they claw back part of that loss over subsequent horizons, they remain below baseline for the entire ten-period window. This is the classic "own-effect" that captures how the shock depresses the originating economy most strongly. The disruption is not limited to Germany: exports from China and Japan also fall—though by much smaller margins—highlighting how a downturn in German trade ripples through global supply chains. Those foreign responses persist over time, yet their magnitudes never match the depth of Germany's initial drop. By contrast, U.S. exports hardly budge—the response curve for the United States hovers indistinguishably around zero—signaling that, in this specification, American trade is effectively insulated from a German-specific shock. Finally, the shaded bands enveloping each line represent statistical uncertainty (a 68 or 95 percent credible interval); they widen as the horizon lengthens, reflecting the compounding imprecision in long-term projections.

Table 1 shows how a shock to one country propagates through the system.

Table 1. Immediate and long-run effects of trade shocks to German exports.

Country	Immediate Effect ($h = 0$)	Long-Run Effect ($h = 10$)	Spillover?
DEU (Germany)	Large negative	Persistent, less negative	Own shock
CHN (China)	Small negative	Persistent	Yes
JPN (Japan)	Small negative	Persistent	Yes
USA (United States of America)	Near zero	Near zero	No

Source: Authors' Own Computations.

Table 1 shows that a German shock mostly affects Germany, with smaller, persistent spillovers to China and Japan, and negligible effect on the USA. Appendix 3 provides the out-of-sample validation results for 3 countries, namely China, Germany, and Japan.

4.2. Why Was VAR(2) Chosen as the Model Specification?

Whether we use one, two, or three lags, the impulse-response curves sit on top of each other; peak effects, long-run impacts, and "fade-out" ratios differ only in the second decimal place. So the policy conclusions stay the same.

Letting the model look further back in time is tempting, but every added lag doubles the number of parameters it has to estimate. With quarterly data, that extra complexity quickly drains the limited information available, stretches the confidence intervals, and makes the forecasts behave erratically. Capping the model at two lags keeps it manageable and the predictions more stable.

Stopping at two lags keeps things in balance. Looking back half a year is long enough to capture how these trade shocks ripple through the economy, yet short enough to avoid over-complicating the model. In other words, two lags hit the "just right" point: rich enough to describe the data, lean enough to stay reliable. Appendix 4 provides robustness of the VAR results to different orderings of the variables.

4.3. Trade Spillovers in the Medium-Run

We consider the medium horizon i.e. $h = 4$ for the 4 countries examined in the following table.

Table 2 provides some interesting results. When a downturn begins in China, it does not stay at home. Within four quarters, the drop in Chinese exports drags German and Japanese exports down by almost three percent each, and China itself remains just as badly hit. A shock that starts in Germany is felt abroad too, but much less dramatically: foreign exports shrink by only about eight-tenths of a percent. Japan's own shock behaves differently; after an initial dip it swings upward, so by the fourth quarter Japanese exports and those of its partners have edged slightly above their pre-shock paths. A shock that starts in the United States is largely contained within its own borders—foreign exports show no meaningful movement at all. These results line up with the impulse-response charts, which show the largest immediate falls for China and Germany, a mild rebound for Japan, and almost no reaction to a U.S. shock.

Table 2. Trade spillover matrix in the medium-run ($h = 4$).

Country	Germany	China	Japan	USA
DEU (Germany)	−2.85	−0.8	1.26	−0.58
CHN (China)	−2.94	−0.82	1.3	−0.6
JPN (Japan)	−2.77	−0.77	1.23	−0.56
USA (United States of America)	0	0	0	0

Source: Authors' Own Computations.

4.4. Forecast-Error Variance Decomposition

Forecast-error variance decomposition (FEVD) tells us where the uncertainty in each country's trade outlook actually comes from once the initial shock has had time to ricochet through the system. Appendix 1 provides the mathematical reasoning behind the FEVD. At a 10-quarter horizon the numbers break down as follows (**Table 3**):

Table 3. Forecast-error variance decomposition analysis.

Country	Germany	China	Japan	USA
DEU (Germany)	38	42	13	7
CHN (China)	54	18	15	13
JPN (Japan)	35	20	40	5
USA (United States of America)	4	3	3	90

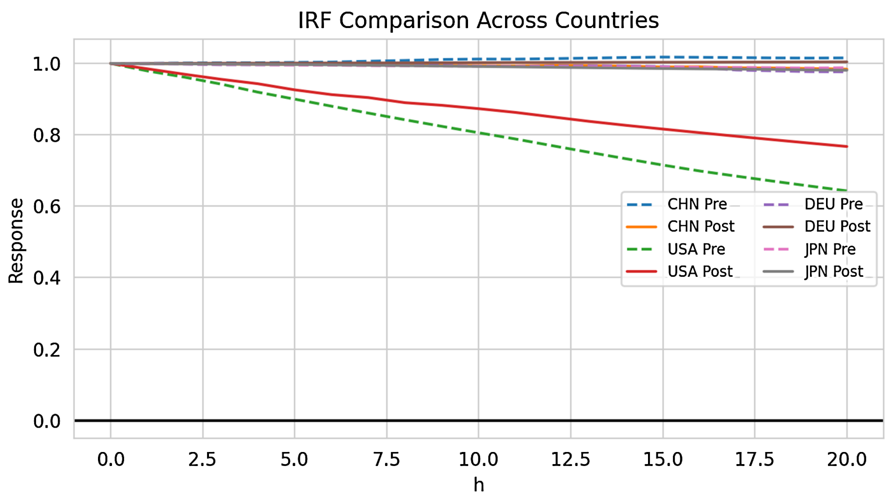
Source: Authors' Own Computations.

Impulse-response functions (IRFs) reveal the average directional effect of a shock—China's export contraction, for instance, pulls German and Japanese exports down by roughly three percent on impact. What IRFs cannot show is how much of each partner's future uncertainty that same shock will generate. Forecast-error variance decomposition (FEVD) fills in that risk dimension: ten quarters after the disturbance, nearly one-third of Germany's and Japan's forecast volatility can still be traced back to the original Chinese shock, even though the level effect has largely played out. Put differently, IRFs map the path of the mean, whereas FEVD maps the turbulence around that path; together they provide a complete picture of both the average damage and the lingering unpredictability embedded in global trade links.

The policy message that flows from combining the two tools is direct. Because Chinese shocks dominate not only the initial export decline but also the medium-run volatility of its partners, measures that stabilize Chinese trade—whether through internal demand support, swap lines, or diversified supply-chain architecture—would deliver the largest payoff to foreign exporters, especially in Germany and Japan. German stabilization carries secondary regional benefits, while U.S.-focused measures mostly help USA itself, given the negligible spillovers of American shocks.

4.5. Pre-Covid and Post-Covid Shocks to Individual Countries

We also split the sample for two sub-periods: 2012-2019 and 2020-2023 for pre-covid and post-covid results. The main objective of this exercise is to ascertain whether the shocks had different effects on bilateral trade flows prior to the Covid-19 period and post Covid-19 period. Figure 2 and Figure 3 provide the graphs of these results.



Source: Authors' Own Computations.

Figure 2. IRF comparisons across countries.

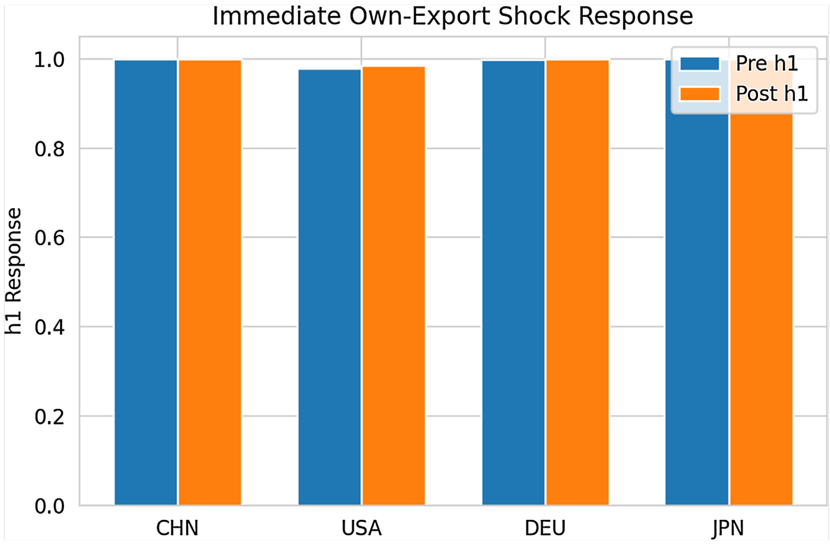
Figure 2 shows how the response to a country's own export shock evolves over time for four countries—China (CHN), USA, Germany (DEU), and Japan (JPN)—under two scenarios: "Pre" and "Post" (likely before and after a structural change or policy event).

Each country has two lines:

- "Pre" (dashed or colored) = before the event.
- "Post" (solid or colored) = after the event.

Across the entire twenty-quarter horizon, a trade shock that begins in China, Germany, or Japan hardly fades at all—it remains embedded in those countries' trade flows from start to finish, no matter which policy regime you look at. The United States behaves differently: its own shock gradually loses force, although the fade-out is slower once the new regime takes hold. Put simply, the regime shift we study does almost nothing to the staying power of shocks in China, Germany, and Japan, yet it does make U.S. trade disturbances linger longer than they used to.

Economically, this means most major exporters carry the imprint of a negative shock for years, whereas the United States normally shakes it off—just not as quickly after the policy change. The following bar-chart also confirms the above result.



Source: Authors' Own Computations.

Figure 3. Effect of immediate trade shocks (Pre-Covid and Post-Covid).

In **Figure 3**, each group of bars represents a country: China (CHN), USA, Germany (DEU), and Japan (JPN). For each country, there are two bars: one for the period before the regime change ("Pre h1," in blue) and one for after ("Post h1," in orange). All bars are very close to 1, meaning that the immediate effect of a trade shock on a country's own exports is nearly the same size as the shock itself, both before and after the regime change.

The main takeaway is that the immediate impact of an export shock on each country's own exports is extremely strong and barely changes between the pre- and post-regime periods. In other words, regardless of the regime, when a country experiences a trade shock, the initial effect on its own exports is almost fully felt right away, and this pattern is consistent across China, the USA, Germany, and Japan. **Table 4** summarizes this finding.

Table 4. Persistence ratios (h8/h1).

Country	Pre	Post
CHN	1.01	1.00
USA	0.86	0.90
DEU	1.00	1.00
JPN	0.99	0.99

Source: Authors' Own Computations.

When Chinese exports fall, their trading partners take the biggest hit over the medium term. A downturn that starts in Germany still hurts abroad but not as much. Shocks from Japan start out negative but soon reverse and even give partners a slight lift. When the United States experiences an export shock, the impact on other countries is so small and erratic that it is hard to pin down with confidence.

5. Conclusions and Policy Implications

The main result of this study from the IRF reveal the following result. First, stabilizing Chinese trade would cut German and Japanese trade volatility by more than one-third, yielding great dividend. Equivalent action in the U.S. would have almost no foreign payoff. Germany's own stabilization offers medium returns to Asia, while damping Japanese volatility mainly benefits Japan.

Second, the FEVD decomposition analysis answers the question of how big the level hit is. The main finding that China knocks partner exports down for other countries by almost 3 percent would imply that policies that stabilize Chinese trade flows would cut not only average losses but also the volatility their partners face. In short, FEVD shows that China is the single largest source of medium-run export uncertainty its major trading partners, while the United States is effectively a volatility island.

From a policy standpoint, the following are important: First, China focused insurance could deliver the largest reduction in external trade risk, especially for Japan and Germany; Second, German centered measures give more modest but still tangible benefits abroad; and finally, U.S. oriented stabilization tools mostly pay off at home and do little for the rest of the world.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Anderson, J. E., & Wincoop, E. V. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review, 93,* 170-192. https://doi.org/10.1257/000282803321455214

Baltagi, B. H., Egger, P. H., & Pfaffermayr, M. (2015). Panel Data Gravity Models of International Trade. In B. H. Baltagi (Ed.), *The Oxford Handbook of Panel Data*. Oxford Academic. https://doi.org/10.1093/oxfordhb/9780199940042.001.0001

Correlates of War Project (2025). *Colonial Contiguity Data, 1816-2016. Version 3.1.* https://correlatesofwar.org/data-sets/colonial-dependency-contiguity/

Egger, P. H., & Nigai, S. (2015). Structural Gravity with Dummies Only: Constrained ANOVA-Type Estimation of Gravity Models. *Journal of International Economics, 97,* 86-99. https://doi.org/10.1016/j.jinteco.2015.05.004

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica, 71,* 1695-1725.

https://doi.org/10.1111/1468-0262.00467

Santos Silva, J. M. C., & Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics, 88,* 641-658. https://doi.org/10.1162/rest.88.4.641

Appendix 1: Mathematical Reasoning Behind the FEVD Results

This appendix shows the mathematical formulation behind the FEVD results. They are given as follows:

Step 1: $\text{VAR}(p)$ for trade flows

$$Y_t = \begin{pmatrix} \Delta x_{DEU,t} \\ \Delta x_{CHN,t} \\ \Delta x_{JPN,t} \\ \Delta x_{USA,t} \end{pmatrix} \tag{A.1}$$

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + u_t \tag{A.2}$$

$$u_t = B\varepsilon_t, \quad \varepsilon_t \sim N(0, I_4) \tag{A.3}$$

Step 2: Moving-average (MA) Representation

$$\varphi_0 = B \tag{A.4}$$

$$\varphi_h = \sum\nolimits_{k=1}^{p} A_k \varphi_{h-k}, \quad h \ge 1 \tag{A.5}$$

$$Y_{t+h} = \sum\nolimits_{k=0}^{\infty} \varphi_k \varepsilon_{t+h-k} \tag{A.6}$$

Step 3: Forecast-error variance at horizon H

$$e_{t+H} = \sum\nolimits_{k=0}^{H-1} \varphi_k \varepsilon_{t+H-k} \tag{A.7}$$

$$\omega_H = \sum\nolimits_{k=0}^{H-1} \varphi_k \varphi_k^{\mathrm{T}} \tag{A.8}$$

Step 4: FEVD share (destination i from source j)

$$\text{FEVD}_{i \leftarrow j}(H) = \frac{\sum_{k=0}^{H-1} \left(e_i^{\mathrm{T}} \varphi_k e_j\right)^2}{\sum_{k=0}^{H-1} \sum_{l=1}^{4} \left(e_i^{\mathrm{T}} \varphi_k e_l\right)^2}, \quad e_i = \text{unit vector}(i) \tag{A.9}$$

Appendix 2: Persistence Ratios across Periods (Pre-Covid and Post Covid)

The following **Figure A1** shows the persistence ratios (h8/h1) during the pre-covid and post-covid periods for each of the countries.

Source: Authors' Own Computations.

Figure A1. Persistence ratios.

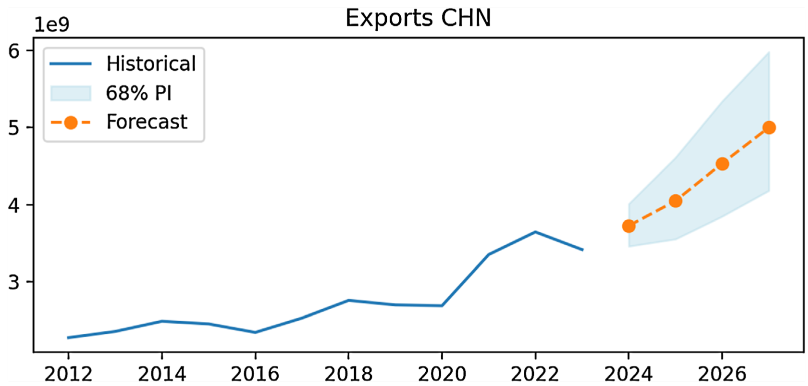
This heatmap shows the persistence ratio (h8/h1) for four countries across two periods: pre-covid and post-covid. This ratio compares the impulse response at horizon 8 to that of horizon 1, measuring how persistent the effect of a shock is over time. A value close to 1 means the effect at horizon 8 is about as strong as at horizon 1 implying high persistence. On the other hand, a value much less than 1 means the effect decays fast i.e. there is low persistence.

The bright yellow color indicates that the effect of a shock barely fades by the 8-period horizon. China, Germany and Japan exhibit these trends in both sub-periods. On the other hand, orange color indicates a noticeable decay. USA is in this band with the post-covid period signaling a modest increase in persistence from 0.86 to 0.9.

The key takeaway is that trade shocks to China, Germany, and Japan remain almost undiminished for eight periods, where as U.S. trade shocks diminish faster.

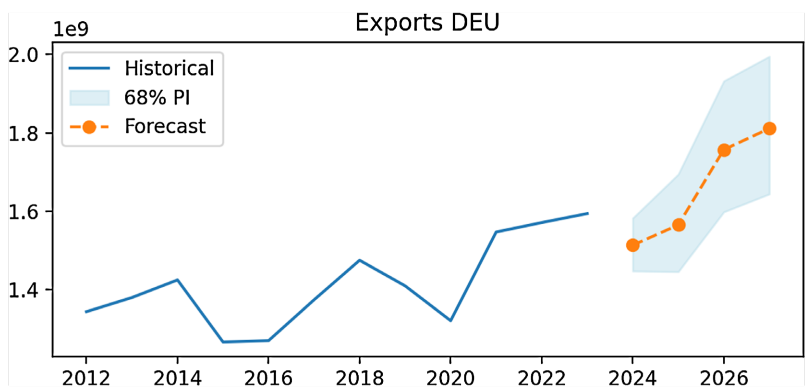
Appendix 3: Out-of-Sample Validation Results

This appendix provides the trade flows of China, Germany and Japan as examples with the 68% prediction bands in light blue and the dashed line showing the median forecast (Figures A2-A4).



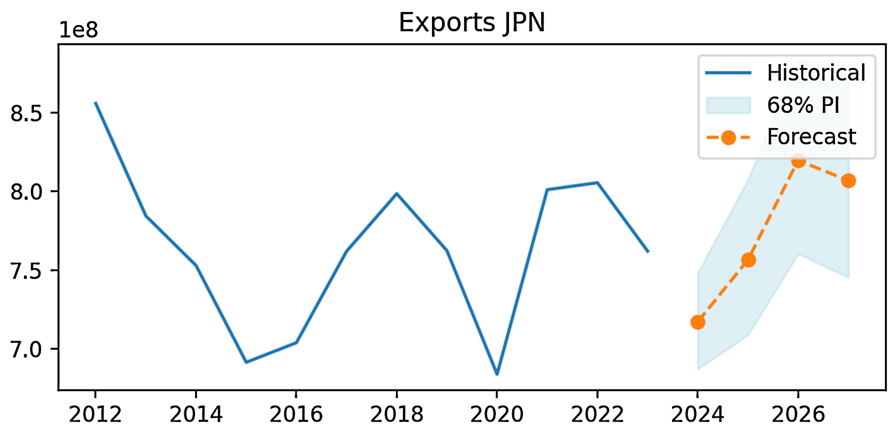
Source: Authors' Own Computations.

Figure A2. Out-of-sample prediction for China's trade flows.



Source: Authors' Own Computations.

Figure A3. Out-of-sample prediction for Germany trade flows.



Source: Authors' Own Computations.

Figure A4. Out-of-sample prediction for Japan trade flows.

Key points

- China continues its post-pandemic climb, but the fan widens sharply after 2025.
- Germany's recovery is gentler; its uncertainty band is narrower than China's.
- Japan remains nearly flat, reflecting subdued historical volatility.

Appendix 4: Robustness of Results to Different Shock Orderings in the VAR (2) Model

We ran the two-lag VAR three separate times, each time changing the order in which China, Germany and Japan appear in the Cholesky identification. Even after these rotations, the story hardly changed. China's own export shock always rises by roughly two to three percent in the first year and then settles to about half that size four years later.

Spillovers to Germany and Japan stay extremely small—well below half a percent—and those spillovers shift by only a fraction of a percentage point when we swap the ordering. Because the estimated coefficients themselves remain virtually unchanged, the medium-term forecasts and their fan charts look almost identical across all three runs. These results are shown in Table A1.

Table A1. 10-year forecast-error variance decomposition results.

Country	Own-shock share	Other-shock share
China	83% - 85%	15% - 17%
Germany	78% - 81%	19% - 22%
Japan	88% - 90%	10% - 12%

Source: Authors' Own Computations.

In short, most of the forecast uncertainty for each country still comes from its own shocks, and re-ordering the variables barely nudges that pattern.