

Power Prior for VARs^{*}

Martin Fankhauser[†], Massimiliano Marcellino[‡], Tommaso Tornese[§]

October 2, 2026

Abstract

We introduce the power-prior framework to vector autoregressions (VARs) as a flexible way to incorporate external data into a focal-country analysis. External likelihoods enter with estimated discount factors, allowing the degree of borrowing to adapt to the similarity between external and current data. We extend existing theoretical results on power priors to VARs, show their connection to simple hierarchical structures, and establish that the resulting priors preserve conjugacy and compatibility with modern structural identification methods. The power prior allows borrowing foreign information without estimating a full multicountry system. In an application to long-run Phillips curves, we show that cross-country borrowing substantially tightens posterior credible sets, while pseudo-out-of-sample exercises across two datasets and distinct evaluation periods demonstrate forecast performance competitive with established multicountry models.

Keywords: BVAR, shrinkage, forecasting, structural analysis

JEL Codes: C11, C15, C32, C52, C53, E37, E52

^{*}We thank Todd Clark, Jan Prüser, Christian Matthes, Roberto Casarin, Elmar Mertens, Mikkel Plagborg-Møller, Tobias Sheckel, Ryohei Oishi, Michele Piffer, Annika Camehl and seminar participants at the Annual Conference of the International Association for Applied Econometrics, the Nordic Econometric Meeting and the 7th Vienna Workshop on High-Dimensional Time Series in Macroeconomics and Finance for helpful comments and suggestions. In addition, we would like to thank Luca Benati for sharing his data and code with us.

[†]Bocconi University

[‡]Bocconi University, CEPR, Baffi

[§]Università Cattolica del Sacro Cuore

1 Introduction

The choice of prior is critical in Bayesian macroeconometrics, especially given the small sample sizes typically encountered in practice. The Minnesota prior, introduced by Doan, B. Litterman, and Sims (1984) and R. Litterman (1986), which shrinks VAR coefficients toward independent random walks (or white noise for growth rates), remains a widely used benchmark. A substantial body of literature, discussed in Section 2, has since refined or proposed alternatives to this prior. We contribute to this literature by introducing the power prior to macroeconometrics. The power prior allows researchers to augment any baseline prior with external evidence, such as data from other countries, earlier samples, or model-implied pseudo-data. Together with conjugate Minnesota-style priors, this enables external information borrowing while retaining conjugacy. More generally, however, the construction is modular and can be combined with essentially any baseline prior used in the VAR literature.

We focus on using the power prior to incorporate foreign data into the current data analysis. Our motivation is straightforward. Macroeconomic outcomes are linked across countries through common shocks, trade linkages, and financial spillovers, so data from foreign economies may contain useful signals for inference and forecasting in a given country. Existing multicountry approaches, such as panel VARs, hierarchical VARs, and global VARs, are designed to exploit such information. But these methods are often aimed at learning about the full cross-country system and typically require additional modeling choices about heterogeneity, pooling restrictions, and cross-sectional dependence. In many empirical settings, however, the object of interest is more limited: the researcher seeks posterior, predictive, or structural inference for one country, while using data from other countries only as an auxiliary source of information. This is precisely the problem for which the power prior is well suited. Rather than estimating the full joint distribution of a multicountry model, it provides a simple and transparent way to borrow information for a focal country through a discounting device.

Formally, the power prior is obtained by raising the likelihood of the external data to a discount factor $\delta \in [0, 1]$ and combining it with a baseline prior. When $\delta = 1$, the external data enter fully; when $\delta = 0$, they are ignored. Intermediate values allow for partial borrowing and thereby encode beliefs about the relevance of the external source for the current analysis. The framework, justified as an efficient information-processing rule by Zellner (2002) and developed by Ibrahim, Chen, and Sinha (2003), has become a widely used device for incorporating historical or external information in biostatistics and related fields; see Ibrahim, Chen, Gwon, et al. (2015) and Chen, Guan, et al. (2025) for recent overviews. To the best of our knowledge, however, it has not yet been systematically developed for VAR models, the workhorse specification of empirical macroeconometrics.

We extend the theory of normalized power priors to VARs. Building on results such as Duan, Ye, and Smith (2006), Carvalho and Ibrahim (2021), and Han, Ye, and Wang (2023), we derive conditions under which the resulting priors are proper under commonly used baseline specifications, obtain closed-form expressions for the marginal likelihood, and characterize the associated posterior moments. We further show that the power prior can be combined with the asymmetric prior of Chan (2022), which permits equation-specific discounting. This allows the researcher to borrow external information for selected variables while excluding it for others. In addition, we show that the power prior has a hedging property. In a simple normal-normal model, borrowing can improve inference even when no external information source is close to the data-generating process. What matters is whether combinations of external sources and a baseline prior are useful. Finally, we extend to the VAR setting the connection between hierarchical models and power priors established by Chen and Ibrahim (2006) and Shen et al. (2024). In particular, when external information comes from other countries, the power-prior posterior coincides under suitable conditions with the group-specific marginal posterior from a hierarchical model in which country-specific coefficients are drawn from a common distribution.

We document the empirical usefulness of the approach in an application to the long-run

Phillips curve analysis of Benati (2015). We first conduct a recursive pseudo-out-of-sample forecasting exercise comparing three power-prior specifications—using, respectively, diffuse, Minnesota-type, and asymmetric conjugate baseline priors—with their corresponding single-country benchmarks and with the multicountry models of Angelini et al. (2019) and Prüser and Blagov (2026). Cross-country borrowing improves both point and density forecasts relative to the original diffuse specification, while the power-prior models with informative baseline priors also improve on their corresponding single-country counterparts and perform competitively with the multicountry benchmarks. Importantly, these gains do not come at the cost of undercoverage: the empirical coverage of the 90% predictive intervals remains at or above its nominal level, providing no indication that the borrowing mechanism produces systematically overconfident predictive distributions. We then apply a deliberately parsimonious version of the power prior, with one discount factor for each donor country, to the sign- and zero-restricted SVAR of Benati (2015). Borrowing information across Australia, Canada, the Euro Area, the United Kingdom, and the United States substantially tightens posterior inference on the long-run inflation–unemployment trade-off and concentrates the estimates around economically small negative effects, without requiring the estimation and structural identification of a full multicountry system. Leave-one-donor-out results further show that the posterior locations and sign conclusions are not driven by any single foreign sample. As an external-validity exercise, the appendix reports complementary forecasting results using the global VAR dataset of Mohaddes and Raissi (2024), covering a broader country set and a different evaluation period, and finds particularly strong point- and density-forecast gains for GDP, with more heterogeneous performance across the remaining variables.

This application illustrates the main practical appeal of the power prior. It should not be viewed as a replacement for fully joint multicountry models when the objective is to learn about the entire cross-country system. Rather, it is a deliberately targeted alternative for settings in which the researcher seeks inference for one country while exploiting other countries as auxiliary information sources. In this sense, the power prior provides an operational

middle ground between standard single-country Bayesian VAR practice and full hierarchical or multicountry modeling. It is parsimonious—the degree of borrowing from each external source is summarized by a small number of discount factors—transparent to report, and easy to combine with standard Bayesian forecasting and structural-identification algorithms. In particular, because the power prior preserves the familiar single-country VAR form, standard Bayesian structural procedures, including the algorithms of Arias, Rubio-Ramirez, and Waggoner (2018) and Arias, Rubio-Ramirez, and Waggoner (2021), can be applied without modification.

This paper is structured as follows. Section 2 reviews the relevant literature. Section 3 develops the conjugate VAR results, relates power-prior borrowing to hierarchical pooling, derives the hedging and donor-redundancy results, and introduces equation-specific discounting. Subsequently, Section 4 presents the results of an extensive Monte Carlo simulation that provides practical guidance on the use of the power prior for VAR models. Section 5 presents the empirical application to the long-run Phillips curve of Benati (2015). Section 6 concludes.

2 Literature

The Minnesota prior introduced by Doan, B. Litterman, and Sims (1984) and R. Litterman (1986) shrinks VAR coefficients toward independent random walks (or white-noise processes for growth rates) and remains a popular choice for small- and medium-scale systems. Subsequent work extends and/or replaces this benchmark prior, including hierarchical schemes (Giannone, Lenza, and Primiceri (2015), Korobilis and Pettenuzzo (2019)), spike-and-slab formulations such as the stochastic-search variable-selection prior of George, Sun, and Ni (2008), global-local priors (e.g. Normal-Gamma or Horseshoe) studied in Huber and Feldkircher (2019), the prior formulations for time-varying parameter models in Chan (2021), or more economically motivated constructions as in Del Negro and Schorfheide (2004), Carriero, T. Clark, and Marcellino (2021) and Giannone, Lenza, and Primiceri (2019), to name only

a few. The power prior provides a way to incorporate external data, for example from other countries, into VARs in combination with any of the prior specifications above. In applications, we will focus on combining and comparing the power prior with basic Minnesota priors, because they are frequently applied in practice; however, the power prior is more flexible than this.

The potential merit of pooling information across countries or groups and/or including foreign variables in the VAR specification is well established in the literature. For instance, Canova and Ciccarelli (2013) review the literature on panel VARs and Chudik and Pesaran (2016) review that on global VARs. More recent developments include Korobilis (2016), Koop and Korobilis (2019), Angelini et al. (2019), Feldkircher et al. (2022), Pfarrhofer (2023). See also the application to univariate models in Müller and Watson (2025). Bai et al. (2022) provide a comparison of different prior choices for multi-country models. Prüser and Blagov (2026) propose a hierarchical prior allowing zero shrinkage and cross-country pooling. These developments, together with an ever greater availability of standardized data sources (e.g. Mohaddes and Raissi (2024) and Barigozzi, Lissona, and Tonni (2024)), increase the appeal of multicountry models substantially. The hierarchical model interpretation we give below for the power prior positions our paper as a user-friendly alternative to the hierarchical specifications above. This parsimony preserves conjugacy and seamless integration with structural algorithms (e.g. Arias, Rubio-Ramirez, and Waggoner (2018) and Arias, Rubio-Ramirez, and Waggoner (2021)), hence complementing the above contributions. In applications, for a fixed discount, implementing the power prior amounts to replacing current data moments with a weighted combination of current and external data moments. The price paid for this simplicity is that, unlike full hierarchical specifications, the power prior only returns the marginal-by-country posterior as opposed to a full cross-country joint distribution.

The power prior provides a unifying perspective on several strands of the macroeconometric literature, beyond the application to cross-country data. When the external information is generated by an estimated or calibrated structural model, it connects naturally to approaches

such as the DSGE-VAR of Del Negro and Schorfheide (2004) or the no-arbitrage priors for term-structure VARs studied by Carriero, T. Clark, and Marcellino (2021). Similarly, a series’ own past is often discounted in structural analysis; see, for instance, Baumeister and Hamilton (2019) and Kiley (2024). The presented theory puts these ideas into a common framework in which external information enters through a transparent discounting mechanism and can therefore be interpreted, estimated, and compared in a unified way.

As mentioned, excellent surveys of the power prior literature can be found in Ibrahim, Chen, Gwon, et al. (2015) and Chen, Guan, et al. (2025); here we focus on highlighting our contribution to this literature. First, we extend the results in Carvalho and Ibrahim (2021) and Han, Ye, and Wang (2023) on conditions for the normalized power prior to be proper and the conjugate posterior calculations in Banbata, Lesaffre, and Rosmalen (2022) to vector autoregressions. Similarly, the relation between hierarchical models and the power prior established in Chen and Ibrahim (2006) and Shen et al. (2024) is adapted to the multivariate case.¹ Although existing results already show that the power prior is the solution to a problem minimizing the expected Kullback–Leibler divergence between full and no borrowing (see Ibrahim, Chen, and Sinha (2003) for the joint power prior and Ye et al. (2022) for the normalized power prior), our hedging result for the normal-normal location model presented in Proposition 4 in Section 3 appears new to the literature. This theoretical result is illustrated in our VAR simulations and could be of interest beyond macroeconomics.

3 Power Prior for VAR Models

We start by introducing the basic ideas about the power prior. Next, we specialize the results to elicit power priors for VARs. Finally, we discuss the relationships between power priors for VARs that exploit information from other countries, hierarchical priors, and panel VARs. All proofs for the following propositions are provided in Appendix A.

¹Focusing on outlier detection instead of external data borrowing, Billio et al. (2025) independently develop several results with a mathematically similar structure. A similar structure is also prevalent in other tempered-likelihood approaches; see, for instance, T. Clark and Mertens (2026).

Throughout the theory section we use the following two reduced-form VAR(p) representations. For each sample s (e.g., external or current), define $x_{s,t} = (1, y'_{s,t-1}, \dots, y'_{s,t-p})' \in \mathbb{R}^k$ with $k = 1 + np$, and stack $Y_s = (y'_{s,1}, \dots, y'_{s,T_s})' \in \mathbb{R}^{T_s \times n}$ and $Z_s = (x'_{s,1}, \dots, x'_{s,T_s})' \in \mathbb{R}^{T_s \times k}$. The reduced-form model is

$$Y_s = Z_s A + U_s, \quad \text{vec}(U_s) \sim N(0, \Sigma \otimes I_{T_s}), \quad (1)$$

with $A = (b', A'_1, \dots, A'_p)' \in \mathbb{R}^{k \times n}$ and $\Sigma \in \mathbb{R}^{n \times n}$. The corresponding likelihood is

$$p(Y_s \mid A, \Sigma) = (2\pi)^{-\frac{T_s n}{2}} |\Sigma|^{-\frac{T_s}{2}} \exp\left(-\frac{1}{2} \text{tr}\left[\Sigma^{-1}(Y_s - Z_s A)'(Y_s - Z_s A)\right]\right). \quad (2)$$

A useful equivalent representation is obtained by vectorization: $y_s = \text{vec}(Y_s)$, $\beta = \text{vec}(A)$, and $X_s = I_n \otimes Z_s$, which yields

$$y_s = X_s \beta + \varepsilon_s, \quad \varepsilon_s \sim N(0, I_{T_s} \otimes \Sigma). \quad (3)$$

This identity provides the mapping to the hierarchical/panel-VAR section below.

3.1 Normalized Power Prior

Consider observing external data, denoted D_0 , and current data, denoted D_1 . Let the model be described by the likelihood $L(\theta \mid D_s)$, with θ collecting the parameters of interest. The basic power prior specification combines a baseline prior, $\pi_{\text{base}}(\theta)$, with the external likelihood function discounted by a parameter δ :

$$\pi(\theta \mid D_0, \delta) \propto L(\theta \mid D_0)^\delta \pi_{\text{base}}(\theta), \quad \delta \in [0, 1]. \quad (4)$$

If $\delta = 0$, we completely disregard the external data and the prior reduces to the baseline prior. If $\delta = 1$, we do not discount the external data at all; thus the power prior completely

pools the two data sources. Zellner (2002) shows that this generalized Bayesian updating is optimal from an information-theoretic point of view. A limitation of the basic power prior specification is that the discount factor is usually unknown. Duan (2005) and Duan, Ye, and Smith (2006) suggest the *normalized power prior* to conduct inference on (δ, θ) :

$$\pi(\theta, \delta \mid D_0) = \frac{L(\theta \mid D_0)^\delta \pi_{\text{base}}(\theta)}{\int_{\Theta} L(\theta \mid D_0)^\delta \pi_{\text{base}}(\theta) d\theta} \pi_0(\delta), \quad (5)$$

where $\pi_0(\delta)$ is a prior for the discount factor. The construction extends naturally to more than one external data source by replacing $L(\theta \mid D_0)^\delta$ with $\prod_{d=1}^J L(\theta \mid D_d)^{\delta_d}$. For instance, our structural application concerned with estimating the long-run Phillips curve uses, for each country, all four other countries as a donor pool, with four different estimated discount factors for each country. The theory presented next focuses, for expositional purposes, on the two-group case; extensions to more than one group and other external data types are immediate.²

3.2 Core Conjugate Results

Following the general discussion above, the normalized power prior for (A, Σ, δ) in a VAR based on external data (Y_0, Z_0) is

$$\pi(A, \Sigma, \delta \mid Y_0) = \frac{p(Y_0 \mid A, \Sigma)^\delta \pi_{\text{base}}(A, \Sigma)}{\iint p(Y_0 \mid A, \Sigma)^\delta \pi_{\text{base}}(A, \Sigma) dA d\Sigma} \pi_0(\delta). \quad (6)$$

This prior is proper whenever the baseline prior is proper (see Carvalho and Ibrahim (2021)) or the normalization constant is finite. The following proposition gives conditions for this to hold under different standard priors in the VAR literature. A proper prior is important to ensure that the marginal likelihood in δ is well defined.

Proposition 1 (Properness). *Let T_0 denote the sample size of the external data and assume*

²With the exception of potentially correlated priors for discount factors, as for instance discussed in Banbata, Lesaffre, and Rosmalen (2022).

$\text{rank}(Z_0) = k$ and $S_{Y_0} := Y_0' M_{Z_0} Y_0 \succ 0$, where $M_{Z_0} = I_{T_0} - Z_0(Z_0' Z_0)^{-1} Z_0'$. The normalized power prior (6) is proper under the following baseline priors:

(a) *Diffuse prior*: $\pi_{\text{base}}(A, \Sigma) \propto 1$ provided $\delta > \frac{k+2n}{T_0}$.

(b) *Jeffreys prior*: $\pi_{\text{base}}(A, \Sigma) \propto |\Sigma|^{-\frac{n+1}{2}}$ provided $\delta > \frac{k+n-1}{T_0}$.

(c) *Normal-inverse-Wishart (NIW)*: if $\pi_{\text{base}}(A, \Sigma)$ is proper NIW, then the normalized power prior is proper for all $\delta \geq 0$.

These conditions essentially control the tails of the prior distribution and extend the univariate results in Han, Ye, and Wang (2023) to the multivariate case. For the normal-inverse-Wishart baseline prior, the proper baseline ensures that the construction is proper for all discount values, whereas for the other two priors the discount factor has to be bounded from below to control the thickness of the prior tails. Intuitively, in these last cases we need to borrow enough information, i.e., make the prior informative enough, for it to be a proper distribution. For instance, in a small three-variable VAR including four lags and considering an external sample size of one hundred observations, we have to choose a discount factor strictly greater than 0.19 if we want the power prior to be proper.³

Conditional on properness, we can use the marginal likelihood to estimate δ via an empirical Bayes procedure. Define the predictive density of the current sample (Y_1, Z_1) given the external sample as

$$ML(\delta) \equiv \iint p(Y_1 | A, \Sigma) \pi(A, \Sigma | Y_0, \delta) dA d\Sigma, \quad (7)$$

where $\pi(A, \Sigma | Y_0, \delta)$ denotes the normalized power prior conditional on δ . The following is a direct extension of the results in Banbata, Lesaffre, and Rosmalen (2022) to the VAR case.

Proposition 2 (Marginal likelihood). *For $s \in \{0, 1\}$ define $Z_s' Z_s$, $Z_s' Y_s$, $Y_s' Y_s$ as usual.*

³Note that, when we have multiple external sources, the requirement need only hold in total, not for each source, considerably loosening the constraint.

(a) Diffuse baseline $\pi_{base}(A, \Sigma) \propto 1$. Let

$$\begin{aligned}\hat{A}_0 &= (Z'_0 Z_0)^{-1} (Z'_0 Y_0), \\ \tilde{A}(\delta) &= (\delta Z'_0 Z_0 + Z'_1 Z_1)^{-1} (\delta Z'_0 Y_0 + Z'_1 Y_1), \\ S_0(\delta) &= \delta (Y'_0 Y_0 - \hat{A}'_0 (Z'_0 Z_0) \hat{A}_0), \\ S_1(\delta) &= Y'_1 Y_1 + \delta Y'_0 Y_0 - \tilde{A}(\delta)' (\delta Z'_0 Z_0 + Z'_1 Z_1) \tilde{A}(\delta), \\ v_0(\delta) &= \delta T_0 - k - n - 1, \quad v_1(\delta) = \delta T_0 + T_1 - k - n - 1.\end{aligned}$$

Then, up to a δ -independent constant,

$$ML(\delta) \propto \left(\frac{|\delta Z'_0 Z_0|}{|\delta Z'_0 Z_0 + Z'_1 Z_1|} \right)^{\frac{n}{2}} |S_0(\delta)|^{\frac{v_0(\delta)}{2}} |S_1(\delta)|^{-\frac{v_1(\delta)}{2}} \frac{\Gamma_n\left(\frac{v_1(\delta)}{2}\right)}{\Gamma_n\left(\frac{v_0(\delta)}{2}\right)}.$$

(b) Jeffreys baseline $\pi_{base}(A, \Sigma) \propto |\Sigma|^{-(n+1)/2}$. Same expression as in (a), but with

$$v_0(\delta) = \delta T_0 - k, \quad v_1(\delta) = \delta T_0 + T_1 - k.$$

(c) NIW baseline. Let $\pi_{base}(A \mid \Sigma) = MN(\bar{A}, \Omega, \Sigma)$ and $\pi_{base}(\Sigma) = IW(\nu_0, S_0)$. Define

$$\begin{aligned}\tilde{A}(\delta) &= (\Omega^{-1} + \delta Z'_0 Z_0)^{-1} (\Omega^{-1} \bar{A} + \delta Z'_0 Y_0), \\ \hat{A}(\delta) &= (\Omega^{-1} + \delta Z'_0 Z_0 + Z'_1 Z_1)^{-1} (\Omega^{-1} \bar{A} + \delta Z'_0 Y_0 + Z'_1 Y_1), \\ \tilde{S}(\delta) &= S_0 + \delta Y'_0 Y_0 + \bar{A}' \Omega^{-1} \bar{A} - \tilde{A}(\delta)' (\Omega^{-1} + \delta Z'_0 Z_0) \tilde{A}(\delta), \\ \hat{S}(\delta) &= S_0 + \delta Y'_0 Y_0 + Y'_1 Y_1 + \bar{A}' \Omega^{-1} \bar{A} - \hat{A}(\delta)' (\Omega^{-1} + \delta Z'_0 Z_0 + Z'_1 Z_1) \hat{A}(\delta), \\ \tilde{\nu}(\delta) &= \nu_0 + \delta T_0, \quad \hat{\nu}(\delta) = \nu_0 + \delta T_0 + T_1.\end{aligned}$$

Then, up to a δ -independent constant,

$$ML(\delta) \propto \left(\frac{|\Omega^{-1} + \delta Z_0' Z_0|}{|\Omega^{-1} + \delta Z_0' Z_0 + Z_1' Z_1|} \right)^{\frac{n}{2}} |\tilde{S}(\delta)|^{\frac{\tilde{\nu}(\delta)}{2}} |\hat{S}(\delta)|^{-\frac{\tilde{\nu}(\delta)}{2}} \frac{\Gamma_n\left(\frac{\tilde{\nu}(\delta)}{2}\right)}{\Gamma_n\left(\frac{\nu(\delta)}{2}\right)}.$$

In the simulations and in the structural application we use both an empirical-Bayes procedure and a fully Bayesian treatment of the parameter δ ; the forecasting exercises condition on empirical-Bayes discount estimates. The fully Bayesian treatment is unlikely to yield a full borrowing solution even if the past and current data are identical. This has previously been discussed in, e.g., Gravestock and Held (2017), Pawel et al. (2023), and Shen et al. (2026). In particular, Shen et al. (2026) show that the marginal posterior for the discount converges to a point mass at zero if there is any discrepancy between external and current data but fails to converge to a point mass at one if the DGPs are identical. The results in Shen et al. (2026) suggest that this problem is less severe in higher dimensions. In contrast, Ye et al. (2022) show under some conditions that the marginal posterior of the discount is close to one. Although maximizing the marginal likelihood avoids this excessive discounting, the empirical-Bayes implementation conditions on an estimated discount factor and therefore does not propagate discount uncertainty. We therefore additionally consider a fully Bayesian treatment of δ . Section 4 examines the resulting finite-sample trade-off. In our simulation design, treating the discount factor as random partially mitigates the coverage loss.

Finally, note that Proposition 2 also contains the posterior moments of the conjugate power prior model for the different baseline priors. For the normal-inverse-Wishart (NIW) case, the posterior mean of the slope coefficients is

$$\hat{A}(\delta) = (\delta Z_0' Z_0 + Z_1' Z_1 + \Omega(\lambda)^{-1})^{-1} (\delta Z_0' Y_0 + Z_1' Y_1 + \Omega(\lambda)^{-1} \bar{A}),$$

where we make explicit the dependence of the prior precision $\Omega(\lambda)$ on the hyperparameters λ . Implementing the power prior with fixed δ therefore amounts to replacing the usual cross-

products $Z_1'Z_1$ and $Z_1'Y_1$ by the weighted combinations $\delta Z_0'Z_0 + Z_1'Z_1$ and $\delta Z_0'Y_0 + Z_1'Y_1$. The posterior mean therefore combines current-data sufficient statistics, discounted external-data sufficient statistics, and the baseline prior mean, weighted by their respective precisions. The expression for the posterior variance, scale, and degrees of freedom follows a similar logic.

This representation also links the power prior to the dummy-observation approach commonly used to implement conjugate VAR priors; see, for example, Giannone, Lenza, and Primiceri (2015). For fixed δ , the powered external likelihood is fully characterized by the fractional sufficient statistics $\delta Z_0'Z_0$, $\delta Z_0'Y_0$, $\delta Y_0'Y_0$, and the effective sample size δT_0 . The external data can therefore be interpreted as generalized dummy observations, with each observation receiving weight δ . Conditional on Σ , this is equivalent to augmenting the VAR with the scaled observations $(\sqrt{\delta}Y_0, \sqrt{\delta}Z_0)$. When Σ is unknown, however, this scaling alone is not sufficient, because treating the scaled matrices as T_0 ordinary observations would contribute T_0 , rather than δT_0 , degrees of freedom to the covariance posterior. The exact interpretation is thus one of fractional, rather than ordinary, dummy observations. Moreover, as with priors implemented through dummy observations, the normalizing constant in (6) is irrelevant for posterior inference conditional on δ but essential when δ is estimated through the marginal likelihood. In this sense, the normalized power prior can be viewed as a data-dependent dummy-observation prior whose strength is governed by the estimated discount factor.

3.3 Interpretation as Hierarchical Model and Hedging Property

This section gives two complementary interpretations of the power-prior construction. The first links direct discounting to hierarchical pooling; the second interprets discount selection as a hedging problem relative to the baseline prior.

Chen and Ibrahim (2006) and Shen et al. (2024) delineate the connection between hierarchical modeling and the power prior. Here we first extend their results to VARs and point out the connection between the power prior and panel VARs.

Consider the vectorized representation (3) for each group j :

$$y_j = X_j \beta_j + \varepsilon_j, \quad \varepsilon_j \sim N(0, I_{T_j} \otimes \Sigma_j),$$

where $y_j = \text{vec}(Y_j)$, $X_j = I_n \otimes Z_j$, and $\beta_j = \text{vec}(A_j)$. Assume for simplicity $j \in \{a, b\}$ and that Σ_a, Σ_b are known. A simple hierarchical (random effects) structure is

$$\beta_j = \bar{\beta} + \eta_j, \quad \eta_j \sim N(0, \Omega),$$

with $\bar{\beta}$ a common hyper-mean. We assume $p(\bar{\beta}) \propto 1$, that Ω is positive definite, and that X_b has full column rank.

Proposition 3. *Assume Σ_a and Σ_b are known, Ω is positive definite, X_b has full column rank, and $p(\bar{\beta}) \propto 1$. Consider the corresponding power-prior specification with a flat baseline prior for β_a . Let $V_b := I_{T_b} \otimes \Sigma_b$. The marginal posterior of β_a under the hierarchical specification above coincides with the posterior from a power prior specification that uses $p(y_b \mid \beta_a, \Sigma_b)^\delta$ as an external likelihood contribution if and only if*

$$\delta(I_{nk} + 2\Omega X_b' V_b^{-1} X_b) = I_{nk}.$$

In particular, if $\Omega = g(X_b' V_b^{-1} X_b)^{-1}$ (a Zellner g -prior form), then the equivalent discount factor is $\delta = 1/(1 + 2g)$.

This result links the power prior to a simple hierarchical model. In a two-country setting, the posterior from a focal-country VAR that uses the other country as external information through a power prior coincides with the corresponding marginal posterior from a joint hierarchical model. When the object of interest is only one country, estimating the full hierarchical system may therefore be unnecessarily complex. By contrast, conditional on a given discount factor, the power prior is no more difficult to estimate than a standard single-country VAR. This is especially attractive in structural applications, because familiar

single-country identification algorithms (e.g. Arias, Rubio-Ramirez, and Waggoner (2018), Arias, Rubio-Ramirez, and Waggoner (2021)) can then be applied without modification.

The proposition should, however, be interpreted as a local equivalence result rather than as a general identity between power priors and hierarchical or panel VARs. A simple scalar example makes both the connection and the difference clear. Suppose

$$y_j \mid \beta_j \sim N(\beta_j, h_j^{-1}), \quad j \in \{a, b\},$$

and consider a pairwise pooling prior similar to Prüser and Blagov (2026), with an implicit flat prior on the common level,

$$p(\beta_a, \beta_b) \propto \exp \left[-\frac{\rho_{ab}}{2} (\beta_a - \beta_b)^2 \right].$$

Equivalently, conditional on β_a , this prior implies $\beta_b \mid \beta_a \sim N(\beta_a, \rho_{ab}^{-1})$. Integrating out β_b gives

$$y_b \mid \beta_a \sim N(\beta_a, h_b^{-1} + \rho_{ab}^{-1}).$$

Thus the donor contribution has precision $h_b \rho_{ab} / (h_b + \rho_{ab})$. A power-prior contribution $p(y_b \mid \beta_a)^{\delta_{a \leftarrow b}}$ has precision $\delta_{a \leftarrow b} h_b$, so the two coincide when

$$\delta_{a \leftarrow b} = \frac{\rho_{ab}}{h_b + \rho_{ab}}.$$

This is the scalar analogue of Proposition 3.⁴ It also shows the restriction imposed by a

⁴The scalar pooling prior above places all of its coupling precision on the contrast $\beta_a - \beta_b$, giving $\delta_{a \leftarrow b} = \rho_{ab} / (h_b + \rho_{ab})$. Proposition 3 instead routes the coupling through a common hyper-mean shared by both groups; integrating out that mean assigns only half the coupling precision to any single pairwise contrast, which is the origin of the factor 2 in the equivalence condition and of the g -prior discount $1/(1+2g)$. More generally, with J symmetric donor groups this constant is $(J+1)$ -dependent. The two mappings therefore differ, which is precisely the sense in which the power-prior/hierarchy correspondence is a specification-dependent *local* equivalence rather than an identity.

symmetric pairwise pooling relation: the reverse target problem would imply

$$\delta_{b \leftarrow a} = \frac{\rho_{ab}}{h_a + \rho_{ab}},$$

so the two effective discounts are tied through the same similarity parameter ρ_{ab} . In a power-prior specification, by contrast, $\delta_{a \leftarrow b}$ and $\delta_{b \leftarrow a}$ are target-specific objects and need not be linked.

The difference becomes sharper with more than two countries. Suppose there are three countries and a joint pairwise pooling prior links a to b and b to c , but not a directly to c :

$$p(\beta_a, \beta_b, \beta_c) \propto \exp \left[-\frac{\rho_{ab}}{2}(\beta_a - \beta_b)^2 - \frac{\rho_{bc}}{2}(\beta_b - \beta_c)^2 \right].$$

Combining this prior with the three Gaussian likelihoods gives posterior precision matrix

$$K = \begin{pmatrix} h_a + \rho_{ab} & -\rho_{ab} & 0 \\ -\rho_{ab} & h_b + \rho_{ab} + \rho_{bc} & -\rho_{bc} \\ 0 & -\rho_{bc} & h_c + \rho_{bc} \end{pmatrix}.$$

Hence the effect of y_c on the posterior mean of β_a is

$$\frac{\partial E(\beta_a \mid y_a, y_b, y_c)}{\partial y_c} = h_c(K^{-1})_{ac} = \frac{h_c \rho_{ab} \rho_{bc}}{\det(K)} > 0.$$

Country c therefore affects inference for country a through country b , even though there is no direct a - c pooling term. By contrast, a focal-country power prior for country a has the form

$$p(\beta_a \mid y_a, y_b, y_c) \propto p(y_a \mid \beta_a) p(y_b \mid \beta_a)^{\delta_{a \leftarrow b}} p(y_c \mid \beta_a)^{\delta_{a \leftarrow c}} p_{\text{base}}(\beta_a).$$

If $\delta_{a \leftarrow c} = 0$, then y_c has no effect on the posterior of β_a . Thus borrowing is direct and target-specific rather than induced through a joint pooling network.⁵

⁵More general normalized power priors can introduce dependence across discount factors, which would

This distinction is useful when comparing the power prior with hierarchical or panel VAR specifications such as Prüser and Blagov (2026). Pairwise pooling priors can be more granular in coefficient space, since the pooling strength may vary coefficient by coefficient. The direct-discount specification used in this paper instead provides a focal-country borrowing device: it does not estimate the full cross-country joint distribution and it does not generate transitive borrowing. Instead, the discounts can be chosen separately by target country, donor country, and, in the asymmetric extension below, by equation.

Apart from the hierarchical interpretation, the structure of the normalized power prior gives it an alternative hedging interpretation. Let π_{base} denote the baseline prior and let π_{full} denote the prior obtained by updating π_{base} with D_0 at full weight ($\delta = 1$). The normalized power prior $\pi_\delta(\cdot)$ interpolates between them as δ varies. Ye et al. (2022) show that the normalized power prior minimizes the expected weighted Kullback–Leibler divergence between the two extremes. Instead, the following result gives conditions under which no borrowing is suboptimal within a normal-normal location model.

Proposition 4. *Let $y_i = \theta + \varepsilon_i$ with $\varepsilon_i \sim N(0, \sigma^2)$ and denote by D_0 the external sample. Let the oracle distribution be $\theta \sim \pi^* = N(\tau^*, \sigma^2)$. Let the baseline prior be $N(\tau, \sigma^2)$ and define the normalized power prior*

$$\pi_\delta(\theta \mid D_0, \tau) = \frac{L(\theta \mid D_0)^\delta N(\theta; \tau, \sigma^2)}{\int L(\vartheta \mid D_0)^\delta N(\vartheta; \tau, \sigma^2) d\vartheta}, \quad \delta \in [0, 1].$$

Denote by μ_0 the sample mean of D_0 . If $\mu_0 < \tau^* < \tau$ or $\mu_0 > \tau^* > \tau$, then

$$\exists \delta \in (0, 1] \quad \text{s.t.} \quad \text{KL}(\pi^* \parallel \pi_\delta(\cdot \mid D_0, \tau)) < \text{KL}(\pi^* \parallel \pi_0(\cdot \mid D_0, \tau)),$$

i.e. $\delta = 0$ is suboptimal.

introduce a similar network/transitivity property. For example, BHM-matching normalized power priors defined in Shen et al. (2024) define dataset-specific discounts as functions of a common latent borrowing parameter, thereby reproducing the marginal posterior of a hierarchical model in normal-data settings. Such constructions create indirect dependence through the discount parameters rather than through a joint coefficient-pooling network.

Although the setting is simple, the above proposition provides a valuable insight. Even with a badly chosen baseline prior and an external source not matching the current data-generating process, borrowing can still be beneficial. For instance, suppose a random walk prior is chosen for an AR(1) process, but the true autoregressive coefficient is 0.8. If we augment the prior with external data generated from a less persistent process, any positive discount will push the prior toward the true 0.8. On the other hand, if the baseline prior is already well chosen, all an external data source can do is increase precision if its data-generating process aligns with the current data. We believe this hedging property might be especially beneficial for higher lag coefficients in VARs. For instance, the traditional Minnesota prior shrinks the VAR toward random-walk processes, setting the prior mean of higher-order lags to zero. The power prior will pull these coefficients away from zero toward the respective lag structure of the donor data, which, depending on the donor pool, could improve estimation.

If no combination of one external data source and the baseline prior is optimal, the power prior might hurt, even abstracting from estimation issues. The hedging result above suggests, however, that with an increasing number of donor datasets, the hedging property is more likely to hold. The following result isolates the geometry of the posterior center. We therefore hold the covariance matrix fixed across candidate priors. This is the homoskedastic idealization in which donor weights are interpreted as precision shares on the simplex: the pooled precision is $\sum_j \lambda_j \Sigma^{-1} = \Sigma^{-1}$. In the VAR application, donor likelihoods also affect posterior precision; the proposition is therefore best read as a statement about the direction of shrinkage rather than about the total amount of information.

Proposition 5 (Multi-donor hedging and redundancy). *Let $\pi^* = N_r(\tau^*, \Sigma)$ and let $\pi_m = N_r(m, \Sigma)$, with Σ positive definite. Let τ denote the baseline center and let $\mu_1, \dots, \mu_J$ denote donor centers. Define*

$$\mathcal{H}_J = \text{conv}\{\tau, \mu_1, \dots, \mu_J\}, \quad R_J = \inf_{m \in \mathcal{H}_J} \text{KL}(\pi^* \parallel \pi_m).$$

Let $P_J(\tau^*)$ denote the projection of τ^* onto $\mathcal{H}_J$ under the norm $\|x\|_{\Sigma^{-1}}^2 = x'\Sigma^{-1}x$. Then:

- (a) Adding a donor weakly improves the attainable hedge: $R_{J+1} \leq R_J$.
- (b) The improvement is strict if and only if the projection onto the enlarged donor hull lies outside the old donor hull: $R_{J+1} < R_J \iff P_{J+1}(\tau^*) \notin \mathcal{H}_J$.
- (c) In particular, if $\mu_{J+1} \in \mathcal{H}_J$, then $\mathcal{H}_{J+1} = \mathcal{H}_J$ and hence $R_{J+1} = R_J$.

Proposition 5 gives two complementary implications. First, expanding the donor pool weakly improves the best attainable hedge. This is the sense in which multiple donors can be valuable even when no single donor exactly matches the target: a combination of donor centers may move the posterior center closer to the oracle than the baseline prior alone. Second, the result identifies when an additional donor is redundant. If the new donor does not move the Σ^{-1} -projection of the oracle center, it cannot improve the best attainable posterior center. In particular, donors inside the existing convex hull are redundant for this purpose. Donors outside the hull are only potentially useful; they improve the hedge only if they expand the hull in a direction relevant for τ^* .

This is a statement about the direction of shrinkage, not about posterior precision. A second donor that is close to an already selected donor may still add precision if its likelihood contribution is independent. In macroeconomic panels, however, such independence is often questionable because similar countries are exposed to common shocks. The result therefore provides a rationale for retaining donors that add new directions relative to the baseline and already selected donors, while screening out donors that are approximately redundant.

This is precisely where the power prior with independent discounts differs from joint pairwise pooling. In a pairwise pooling model, if country a is linked to b and b is linked to c , then c can affect inference for a through the latent pooling graph. Similar donors are therefore mutually reinforcing. In the direct-discount power prior used here, by contrast, donor c affects the posterior for a only if an explicit discount $\delta_{a \leftarrow c}$ is assigned to its likelihood. Similar donors are therefore candidates for redundancy rather than automatic indirect channels.

3.4 Asymmetric Prior and Equation-Specific Discounting

The construction above is restrictive in the sense that one global discount factor is applied to all coefficients. A natural extension is to allow the discount to vary across VAR equations. For instance, borrowing information may be more attractive for financial variables in the Euro Area than for real variables. In this section we show how the power prior can be combined with the asymmetric prior of Chan (2022) to achieve equation-specific borrowing. Consider the structural representation of the VAR following Chan (2022):

$$A_0 y_t = b + B_1 y_{t-1} + \cdots + B_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim N(0, D), \quad D = \text{diag}(\sigma_1^2, \dots, \sigma_n^2).$$

Let A_0 be unit lower-triangular. Define $\tilde{w}_{i,t} = (-y_{1,t}, \dots, -y_{i-1,t})$ and $\tilde{x}_t = (1, y'_{t-1}, \dots, y'_{t-p})$, and set $x_{i,t} = (\tilde{w}_{i,t}, \tilde{x}_t)'$. Collect the equation-specific parameters in $\theta_i = (\beta'_i, \alpha'_i)'$.

Stacking over t yields, for each $i = 1, \dots, n$,

$$y_i = X_i \theta_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma_i^2 I_T),$$

and because D is diagonal the likelihood factorizes:

$$p(y \mid \theta, \sigma^2) = \prod_{i=1}^n (2\pi\sigma_i^2)^{-\frac{T}{2}} \exp\left(-\frac{1}{2\sigma_i^2} (y_i - X_i \theta_i)' (y_i - X_i \theta_i)\right).$$

Proposition 6. *Under the structural VAR above with equation-specific discount factors $\delta = (\delta_1, \dots, \delta_n)$, the normalizing constant and marginal likelihood of the normalized power prior factorize across equations. In particular, the normalization constant is*

$$C(\delta) = (2\pi)^{-\frac{T_0}{2} \sum_{i=1}^n \delta_i} \prod_{i=1}^n \left[|V_i|^{-\frac{1}{2}} |\tilde{K}_{\theta,i}|^{-\frac{1}{2}} \frac{\Gamma(\nu_i + \frac{\delta_i T_0}{2}) S_i^{\nu_i}}{\Gamma(\nu_i) \tilde{S}_i^{\nu_i + \frac{\delta_i T_0}{2}}} \right],$$

where

$$\begin{aligned}
\tilde{K}_{\theta,i} &= V_i^{-1} + \delta_i X'_{0,i} X_{0,i}, \\
\tilde{\theta}_i &= \tilde{K}_{\theta,i}^{-1} (V_i^{-1} m_i + \delta_i X'_{0,i} y_{0,i}), \\
\tilde{S}_i &= S_i + \frac{1}{2} \left(\delta_i y'_{0,i} y_{0,i} + m'_i V_i^{-1} m_i - \tilde{\theta}'_i \tilde{K}_{\theta,i} \tilde{\theta}_i \right), \\
\tilde{\nu}_i &= \nu_i + \frac{\delta_i T_0}{2}.
\end{aligned}$$

The marginal likelihood for the current sample is

$$ML(\delta) = (2\pi)^{-\frac{nT_1}{2}} \prod_{i=1}^n \left(\frac{|\hat{K}_{\theta,i}|}{|\tilde{K}_{\theta,i}|} \right)^{-\frac{1}{2}} \frac{\Gamma(\hat{\nu}_i) \tilde{S}_i^{\hat{\nu}_i}}{\Gamma(\tilde{\nu}_i) \hat{S}_i^{\tilde{\nu}_i}},$$

where

$$\begin{aligned}
\hat{K}_{\theta,i} &= V_i^{-1} + \delta_i X'_{0,i} X_{0,i} + X'_{1,i} X_{1,i}, \\
\hat{\theta}_i &= \hat{K}_{\theta,i}^{-1} (V_i^{-1} m_i + \delta_i X'_{0,i} y_{0,i} + X'_{1,i} y_{1,i}), \\
\hat{S}_i &= S_i + \frac{1}{2} \left(\delta_i y'_{0,i} y_{0,i} + y'_{1,i} y_{1,i} + m'_i V_i^{-1} m_i - \hat{\theta}'_i \hat{K}_{\theta,i} \hat{\theta}_i \right), \\
\hat{\nu}_i &= \nu_i + \frac{\delta_i T_0 + T_1}{2}.
\end{aligned}$$

Proposition 6 shows that, under the asymmetric prior, the power prior inherits an equation-specific structure. Because the likelihood factorizes across equations, foreign information can be borrowed selectively across variables, and the corresponding empirical-Bayes problem remains tractable.

3.5 Computation and Fully Bayesian Algorithm

This subsection describes posterior simulation for the random-discount specifications with multiple donors, both with common and equation-specific borrowing. The conjugate and properness results above extend by replacing a single discounted external cross-product with

the sum of donor-specific discounted cross-products.

For the non-equation-specific specification, let $\delta = (\delta_1, \dots, \delta_D)'$ denote the donor discounts and define $s = 1 - \sum_{d=1}^D \delta_d$. This auxiliary variable is used to impose the simplex restriction $\sum_{d=1}^D \delta_d < 1$, which is the multi-donor analogue of the single-donor restriction $\delta < 1$ and caps the total donor weight at one. Sampling is conducted on the unconstrained additive-log-ratio scale $z_d = \log(\frac{\delta_d}{s})$. This maps every proposal into the simplex, thereby enforcing the constraints. We place a Dirichlet $(\delta_1, \dots, \delta_D, s) \sim \text{Dirichlet}(a, \dots, a, b)$ with $a = b = 1$ on the discount factors and the slack parameter jointly. The posterior target is proportional to the normalized power-prior marginal likelihood multiplied by the Dirichlet density and, on the transformed scale, the corresponding Jacobian. Draws are additionally restricted to the region in which the normalized power prior is proper given above. We use a componentwise Gaussian random-walk Metropolis algorithm, initialized at the empirical-Bayes marginal-likelihood estimate, and adapt the proposal scales during burn-in before fixing them for posterior sampling. The equation-specific sampler follows the same logic equation-by-equation.

The Dirichlet prior is a natural choice for the discount factors as it is a multivariate extension of the Beta distribution. For $D = 1$, our choice $a = b = 1$ corresponds to a uniform prior over $[0, 1]$. More generally, the prior implies an expected total donor mass of $\frac{D}{D+1}$, so that it becomes increasingly permissive toward total borrowing as the number of potential donors increases.

4 Monte Carlo Simulation

The simulation is designed to answer three practical questions. First, does the empirical-Bayes power prior discount external data when donor dynamics are far from the target? Second, does adaptive borrowing improve estimation and forecasting without behaving like naive pooling? Third, when several donors are available, does the unrestricted power prior use them selectively, and are multi-donor allocations associated with gains over a best-single-donor

benchmark?

The Monte Carlo design is based on a stationary three-variable VAR(1) without intercepts.

For each economy j we generate data from the reduced-form system

$$y_{j,t} = B_j y_{j,t-1} + u_{j,t}, \quad u_{j,t} \sim N(0, \Sigma_j),$$

where $B_j = A_{1,j} A_0^{-1}$ and $\Sigma_j = A_0^{-1} D_j A_0^{-'}$. The contemporaneous matrix is fixed at

$$A_0 = \begin{pmatrix} 1.0 & 0.0 & 0.0 \\ 0.3 & 1.0 & 0.0 \\ -0.2 & 0.5 & 1.0 \end{pmatrix}.$$

The lag matrix $A_{1,j}$ is drawn around convex combinations of two reference matrices,

$$A_1^{(1)} = \begin{pmatrix} 0.5 & 0.4 & 0.2 \\ 0.3 & 0.4 & 0.4 \\ -0.2 & 0.3 & 0.5 \end{pmatrix}, \quad A_1^{(2)} = \begin{pmatrix} 0.9 & 0.0 & 0.0 \\ 0.1 & 0.8 & 0.0 \\ 0.0 & 0.1 & 0.85 \end{pmatrix}.$$

For each parameter draw, we first draw a common perturbation scale $\sigma_A^2 \sim 10^{U(\log_{10}(10^{-4}), \log_{10}(10^{-1}))}$.

Then, independently for the target economy and for each donor economy, we draw $\alpha_j \sim U(0, 1)$, set $\bar{A}_{1,j} = \alpha_j A_1^{(1)} + (1 - \alpha_j) A_1^{(2)}$, and draw $A_{1,j} = \bar{A}_{1,j} + \sigma_A E_j$, where E_j is a matrix with i.i.d. $N(0, 1)$ entries. Draws that imply non-stationary VAR dynamics are discarded and redrawn.

The innovation variances are generated as $D_j = D = \text{diag}(10^{u_1}, 10^{u_2}, 10^{u_3})$ with $u_i \sim U(-0.12, 0.12)$, using the same D for the target and all donors within a parameter draw. Hence differences between the target and donor economies are driven primarily by differences in the dynamic coefficient matrices rather than by differences in innovation variances. For each parameter draw we simulate one target sample of length $T_1 = 100$ and donor samples of length $T_0 = 100$, after discarding a burn-in of 200 observations. We consider $M = 1000$ parameter draws and, conditional on each draw, $R = 200$ Monte Carlo replications. The donor

distance reported on the horizontal axis of the figures is computed from the equation-specific lag coefficient vectors of the target and donor VARs. Small values therefore correspond to donor dynamics close to the target dynamics, whereas large values correspond to increasingly dissimilar donor VARs.

We consider two main simulation designs. The first design contains a single donor, denoted by $D = 1$. This design isolates the basic behavior of the discount factor as donor relevance changes and allows us to study how donor distance affects borrowing, posterior accuracy, interval length, and coverage in the simplest possible setting. The second design contains three donors, denoted by $D = 3$. This design is used to study whether the unrestricted power prior allocates borrowing selectively across multiple external sources and whether multi-donor allocations are associated with gains over a best-single-donor benchmark.

For each simulated sample we fit several competing specifications. The baseline model is the no-borrowing ACP/Minnesota-type prior, with its shrinkage hyperparameters selected by marginal likelihood. The one-donor power prior augments this baseline with a single external likelihood and estimates the corresponding discount factor by empirical Bayes. In addition, we provide results for a fully Bayesian treatment of the discount parameter using a uniform prior. In the three-donor design we estimate an unrestricted power prior with donor- and equation-specific discounts, allowing the procedure either to concentrate borrowing on one donor or to spread mass across several donors. We compare this specification with a best-single-donor model, which selects the donor that maximizes the marginal likelihood among the three one-donor restrictions. This comparison separates the value of adaptive borrowing from the value of simply having more donor data available.

Figure 1 reports the one-donor design. Panel A shows that the empirical-Bayes estimate of the discount factor, shown in black, declines sharply with the ex ante distance between the donor and target VARs. When donor dynamics are close to the target, the empirical-Bayes criterion assigns substantial weight to the external likelihood. As the donor becomes less similar, the discount rapidly approaches zero. This confirms that the power prior does not

mechanically use external data; it borrows only when the donor is sufficiently informative. Panel A also presents the posterior mean of the discount under the fully Bayesian specification in red. Those discounts are less extreme; the posterior does not concentrate on the full borrowing solution, and the discounts converge more slowly to zero for fairly different DGPs.

Panels B–D show the performance consequences of this adaptive discounting. IRF RMSE and posterior interval length improve most when donor distance is small, precisely where the estimated discount is largest. As donor distance increases, the gains fade and the power prior approaches the no-borrowing benchmark. Panel D highlights the main trade-off: borrowing can reduce frequentist coverage, especially at intermediate donor distances where the donor is close enough to receive non-negligible weight but still different enough to introduce shrinkage bias. The random-discount version instead mitigates this concern. For close donors, the random-discount specification retains RMSE and interval-length gains while coverage remains close to the no-borrowing benchmark. Thus, the one-donor experiment shows both the benefit and the cost of adaptive borrowing: the procedure delivers accuracy gains and reductions in uncertainty when donors are relevant, while limiting damage from clearly dissimilar donors.

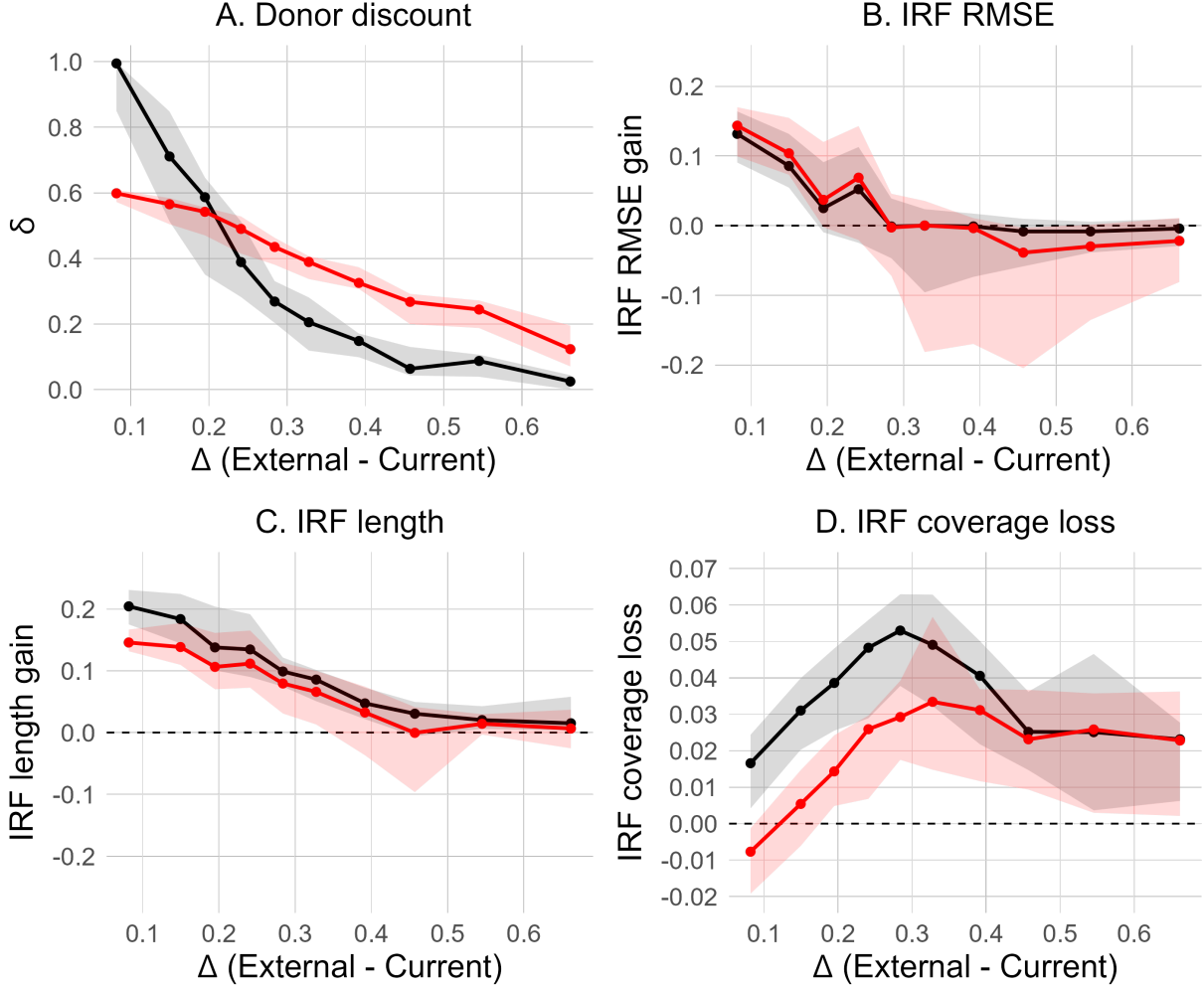


Figure 1: One-donor Monte Carlo design: discounting and estimation performance

Note: Panel A: estimated empirical-Bayes discount factor (black) and posterior mean under the fully Bayesian specification (red) against donor-target distance. Panel B reports IRF RMSE gain relative to no borrowing; Panel C reports interval-length gain relative to no borrowing; Panel D reports IRF coverage loss relative to no borrowing. Positive values in Panels B and C indicate improvement; positive values in Panel D indicate lower coverage under the power prior. Lines denote Monte Carlo averages; shaded areas denote uncertainty bands across parameter draws.

Figure 2 examines the three-donor design. Panel A ranks donors by their ex ante distance from the target DGP and reports the estimated donor-specific discounts. The allocation is strongly selective: the closest donor receives by far the largest discount, the second-closest donor receives substantially less, and the farthest donor receives essentially no weight on average. Thus, even when several external samples are available, the power prior does not mechanically pool all donors. Instead, it allocates borrowing toward the external sources whose dynamics are closest to the target.

Panel B asks whether the cases in which the unrestricted power prior departs from

single-donor-like borrowing are also the cases in which it improves over the ML-best-single benchmark. Forecast gains are computed at the DGP level by averaging forecast MSE over replications before forming RMSE ratios. The figure splits parameter draws into low- and high- N_{eff} regimes, where N_{eff} is the effective number of donors implied by the estimated donor weights. Forecast gains over best-single selection are concentrated in the high- N_{eff} regime. This comparison is descriptive rather than causal, since N_{eff} is itself an outcome of the EB allocation, but it shows that the unrestricted power prior tends to use multiple donors precisely in designs where this additional flexibility is associated with better forecasting performance.

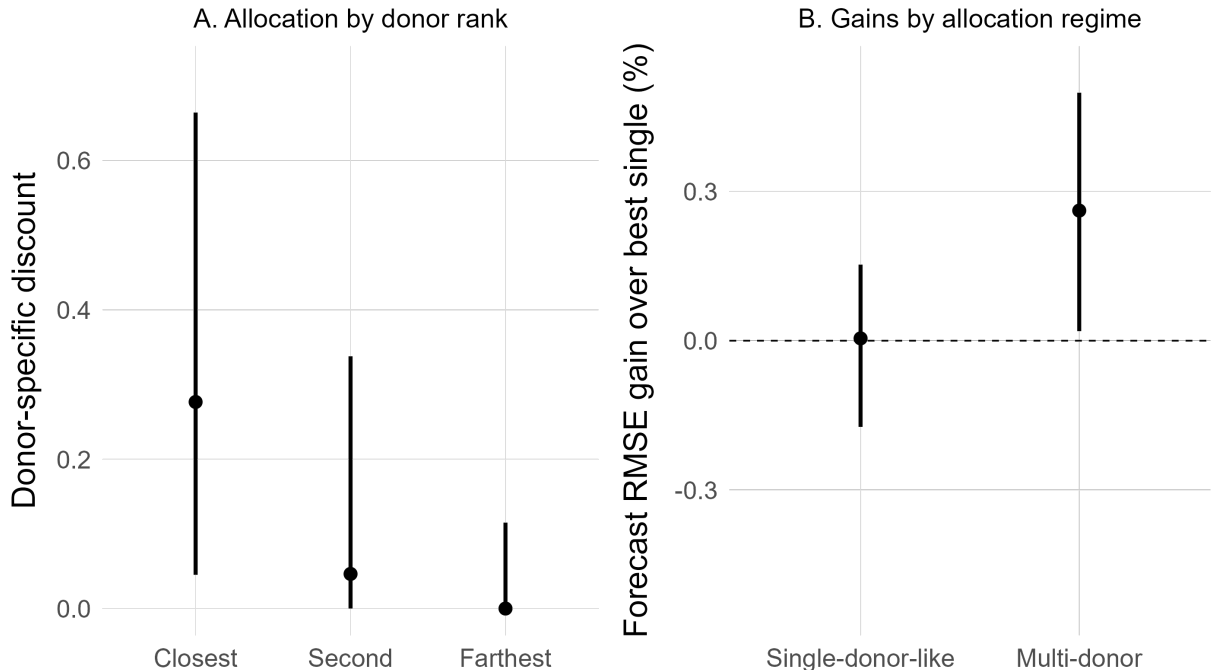


Figure 2: Three-donor Monte Carlo design: selective discount allocation and forecast gains

Note: Panel A: estimated empirical-Bayes donor-specific discounts ranked by ex ante distance to the target (closest, second, and farthest). Panel B: forecast RMSE gains of the unrestricted power prior over the best-single-donor benchmark, split into low- and high- N_{eff} regimes. Dots denote averages across parameter draws; vertical segments denote uncertainty bands.

Overall, the simulation results support three conclusions. First, the empirical-Bayes discount factor behaves sensibly: it assigns high weight to close donors and rapidly discounts donors whose dynamics differ from the target VAR. Second, adaptive borrowing improves estimation accuracy and shortens posterior intervals when donors are relevant, but can reduce frequentist coverage at intermediate donor distances. This overconfidence is partly

expected under the empirical-Bayes procedure, which ignores discount uncertainty, and is partially mitigated in fully Bayesian models. Third, the multi-donor results are consistent with the hedging mechanism: gains over best-single-donor selection are concentrated in designs in which the empirical-Bayes allocation places meaningful weight on several donors. This comparison is descriptive rather than a direct test of Proposition 5.

5 Application

In this section we revisit the empirical analysis in Benati (2015), which estimates long-run Phillips curves for five different countries using a sign- and zero-restricted SVAR. As summarized by Benati (2015), estimating the long-run relationship between inflation and economic activity is difficult for at least two reasons. First, drawing inference on low-frequency behavior from short and finite samples is inherently challenging. Second, identifying long-run relationships requires sufficient low-frequency variation in the relevant variables. In the case of inflation, however, the main source of persistent variation in several countries is concentrated in relatively short episodes such as the Great Inflation. As a result, evidence on the long-run Phillips curve tends to be hidden under substantial econometric uncertainty. In his study, Benati (2015) finds limited and economically small evidence against a vertical long-run relation and concludes: *“The extent of uncertainty surrounding the estimates is however substantial, thus implying that a researcher holding alternative priors about what a reasonable slope of the long-run trade-off might be will likely not see her views falsified.”* (page 16, Benati (2015)).

Here we show that the power prior can substantially sharpen posterior estimates and thus mitigate the expressed concern about posterior precision. We do so by replicating Benati (2015)’s analysis while allowing information to be borrowed across countries through different power-prior constructions using the theory developed above. The remainder of this section first focuses on a pseudo-out-of-sample forecasting exercise demonstrating the merits of the

power prior as a forecasting tool relative to the baseline specification in Benati (2015) and to multicountry benchmark models. Afterwards, we show that the structural conclusions in Benati (2015) can also be materially sharpened once cross-country information is included via the power prior.

After some preliminary unit-root and cointegration analysis, Benati (2015)⁶ estimates six-variable VARs for the following country samples: US (1959Q1–1979Q3), Euro Area (1970Q1–1998Q4), UK (1972Q1–1992Q3), Canada (1961Q1–1990Q4), and Australia (1969Q3–1994Q2). We use the same data and largely retain his specification. The VAR contains a measure of the output gap, inflation, the unemployment rate, a short-term interest rate, the consumption–GDP ratio, and a long–short interest-rate spread. Following Benati (2015), inflation and unemployment enter in first differences, while the output gap, the consumption–GDP ratio, and the spread enter in levels. The short rate enters in first differences except in the UK, where Benati (2015)’s unit-root evidence supports using the level of the short rate. The original quarterly specifications use four lags; the original US model is estimated at monthly frequency with twelve lags, whereas we estimate the US model at quarterly frequency and use four lags for comparability across countries. In the structural exercise, we follow Benati (2015)’s weak-form long-run Phillips-curve specification: permanent and transitory shocks are separated by zero restrictions on the long-run effect on inflation, and the four permanent inflation shocks are disentangled using the impact sign restrictions.

5.1 Forecasting with the Power Prior

We assess the prior constructions in a recursive pseudo-out-of-sample forecasting exercise. All country samples are first aligned to a common start date of 1972Q4. For each focal country, we estimate the model using observations available up to the forecast origin and evaluate forecasts over 1985Q1–2000Q1, thereby directly extending the original sample. Forecasts are computed at horizons $h = 1, \dots, 8$ for the six variables in the original specification, using

⁶We kindly thank Luca Benati for providing us with his dataset.

quarterly VARs with four lags. When foreign countries are used as donors, their samples are lagged by one quarter relative to the focal country at each forecast origin, ensuring that no future donor information enters the forecasts. Forecast performance is summarized by root mean squared forecast errors based on posterior-mean forecasts, log predictive scores computed from the posterior predictive distribution, and the empirical coverage of nominal 90% predictive intervals.

We compare three power-prior specifications with several benchmarks. The first benchmark is the original single-country VAR specification of Benati (2015), estimated under a diffuse prior. Its power-prior counterpart retains this diffuse baseline but augments it with discounted information from the other countries, thereby isolating the value of cross-country borrowing within the original specification. We additionally estimate a standard Minnesota-type prior and a corresponding power-prior version, which separates the contribution of foreign information from that of conventional own-country shrinkage. Finally, to allow both asymmetric prior shrinkage and equation-specific borrowing, we compare the adaptive conjugate prior (ACP) of Chan (2022) with its power-prior counterpart. We also include two multicountry benchmarks: the panel VAR of Angelini et al. (2019), as a parsimonious joint multicountry extension of Minnesota-style shrinkage, and the hierarchical VAR of Prüser and Blagov (2026), which is conceptually closest to our approach because it combines Minnesota-type shrinkage with cross-country pooling. The objective is not to exhaust all possible prior specifications, but to determine whether targeted power-prior borrowing improves a given single-country prior and remains competitive with established multicountry approaches.

Table 1: Pseudo-out-of-sample forecast performance relative to the diffuse single-country benchmark

Model	Median RMSE ratio	RMSE win share	Median LPS difference	LPS win share	90% PI coverage
<i>Power-prior variants</i>					
Power prior, Minnesota baseline	0.864	32.5%	0.216	34.2%	96.3%
Power prior ACP	0.869	19.2%	0.235	28.7%	94.2%
Power prior, diffuse baseline	0.897	13.3%	0.147	4.2%	97.5%
Power-prior family	—	65.0%	—	67.1%	—
<i>Single-country shrinkage benchmarks</i>					
Baseline Minnesota	0.916	3.8%	0.180	8.3%	93.7%
Baseline ACP	0.901	8.3%	0.202	9.2%	93.7%
<i>Multicountry benchmarks</i>					
Hierach. VAR	0.881	17.9%	0.209	14.6%	94.4%
Panel VAR	0.960	5.0%	0.111	0.8%	94.6%

Note: Results are pooled across focal countries, variables, and forecast horizons. The RMSE ratio is calculated relative to the diffuse single-country specification of Benati (2015); values below one favor the model in the corresponding row. The LPS difference is the model’s log predictive score minus that of the diffuse benchmark; positive values indicate better density forecasts. RMSE and LPS win shares report the percentage of country–variable–horizon cells in which the model has, respectively, the lowest RMSE and highest log predictive score among the individual models considered. For the “Power-prior family” row, the win share records the percentage of cells in which at least one of the three power-prior variants is best. Predictive-interval coverage is the empirical coverage of nominal 90% posterior predictive intervals, averaged over the full out-of-sample evaluation periods and then across country–variable–horizon cells.

Table 1 summarizes the aggregate forecasting results, while Figure 3 displays the corresponding medians and interquartile ranges across country–variable–horizon cells. RMSE ratios below one and positive log-score differences indicate improvements relative to the diffuse specification of Benati (2015). Detailed results by country, variable, and horizon, together with predictive-accuracy tests, are reported in Appendix B.2.

All three power-prior specifications improve on the diffuse benchmark in terms of both median RMSE and median log predictive score. Even the diffuse-baseline power prior reduces the median RMSE by approximately 10% and raises the median log predictive score by 0.147, showing that cross-country information can be useful even without an additional shrinkage prior. Combining the power prior with a Minnesota baseline produces a median RMSE ratio of 0.864 and a median log-score improvement of 0.216. Relative to the corresponding single-country Minnesota model, this represents an additional RMSE reduction of approximately 6%. Similarly, the power-prior ACP improves on the single-country ACP, indicating that the gains cannot be attributed solely to conventional within-country shrinkage.

Across the three variants, at least one power-prior model has the lowest RMSE in 65.0% of the country–variable–horizon cells and the highest log predictive score in 67.1% of the

cells. These family-level figures should be interpreted as the combined performance of the three specifications rather than as the win share of a single model. Among the individual specifications, the Minnesota-baseline power prior is best in 32.5% of RMSE comparisons and 34.2% of density-forecast comparisons, while the equation-specific ACP version provides particularly strong median density-forecast performance.

Importantly, the forecasting gains do not arise from predictive distributions that are too narrow. The empirical coverage of nominal 90% predictive intervals ranges from 94.2% to 97.5% across the power-prior variants. Thus, there is no evidence of predictive undercoverage or systematic overconfidence; if anything, the resulting predictive intervals remain conservative. The ACP power-prior specification delivers coverage closest to the nominal level.

The power-prior specifications are also competitive with the multicountry benchmarks. The panel VAR of Angelini et al. (2019) improves on the diffuse benchmark in terms of median RMSE and log predictive score, but all three power-prior specifications have better median performance on both measures. The hierarchical VAR of Prüser and Blagov (2026) performs better than the diffuse-baseline power prior on the aggregate RMSE and log-score measures, but is slightly outperformed by the power-prior specifications with Minnesota and ACP baselines. Targeted focal-country borrowing can therefore achieve forecast performance comparable to that of a full hierarchical multicountry system.

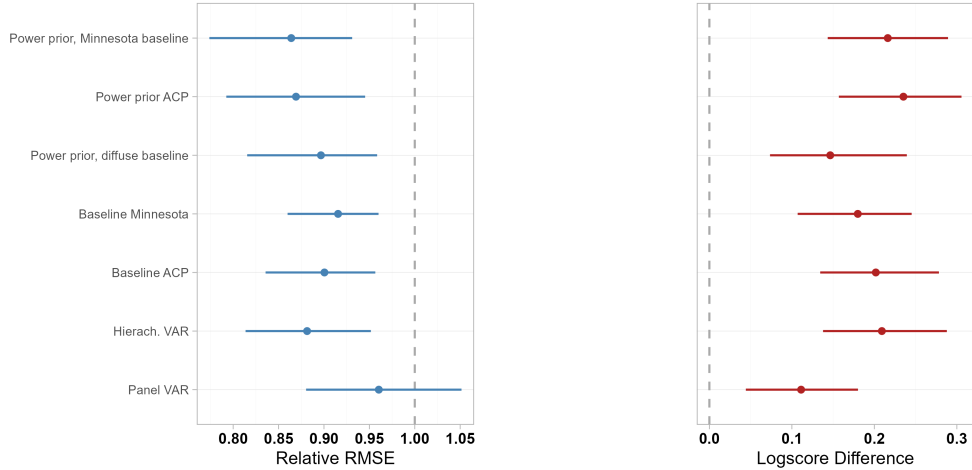


Figure 3: Forecast performance relative to the diffuse single-country benchmark

Note: The left panel reports RMSE ratios relative to the diffuse single-country specification of Benati (2015); values below one indicate improved point forecasts. The right panel reports log predictive score differences relative to the same benchmark; positive values indicate improved density forecasts. Blue and red dots denote medians and horizontal segments denote interquartile ranges across country-variable-horizon cells. Dashed vertical lines indicate equal performance with the benchmark.

To isolate the value of flexible donor allocation, Table 2 compares the unrestricted power-prior specifications with two restricted borrowing alternatives. Under the first restriction, the model can borrow only from the single donor that maximizes the marginal likelihood. Under the second, all donor countries enter symmetrically through equal-share pooling. The reported RMSE ratios and log-score differences compare each unrestricted specification directly with its corresponding restricted version. Hence, RMSE ratios below one and positive log-score differences favor unrestricted empirical-Bayes borrowing.

Table 2: Unrestricted power-prior borrowing relative to restricted donor allocations

Restricted benchmark	Median RMSE ratio	RMSE win share	Median LPS difference	LPS win share
Best single donor	0.972	71.7%	0.044	80.4%
Equal-share pooling	1.008	41.2%	0.043	72.1%
ACP best single donor	0.985	65.0%	0.031	70.4%
ACP equal-share pooling	0.991	60.0%	0.035	67.1%

Note: Each row compares an unrestricted power-prior specification with the indicated restricted donor-allocation benchmark. The first two rows refer to the diffuse baseline and the final two rows to the ACP baseline. The best-single-donor restriction retains only the donor that maximizes the marginal likelihood, whereas equal-share pooling assigns the same total external weight symmetrically across donors. RMSE ratios below one and positive LPS differences favor the unrestricted specification. Win shares report the percentage of country-variable-horizon cells in which the unrestricted specification outperforms the corresponding restricted benchmark.

The unrestricted power prior clearly improves on the best-single-donor restriction. For the diffuse-baseline version, it produces a median RMSE ratio of 0.972, improves RMSE in

71.7% of cells, and improves the log predictive score in 80.4% of cells. The ACP specification displays the same qualitative pattern. These results indicate that useful foreign information is generally not concentrated in a single donor country. Equal-share pooling is more competitive for point forecasts. For the diffuse-baseline model, its median RMSE is slightly lower than that of unrestricted borrowing, although the unrestricted specification still improves RMSE in 41.2% of individual cells. For the ACP model, unrestricted borrowing produces a modest median RMSE improvement and wins in 60.0% of cells. The density-forecast results are clearer: both unrestricted specifications achieve positive median log-score gains and win in more than two-thirds of cells relative to equal pooling. Thus, simple pooling captures part of the point-forecast benefit of foreign information, whereas flexible donor discounting appears particularly valuable for predictive densities.

Table 3 reports discount-allocation diagnostics for the unrestricted diffuse-baseline power prior. The empirical-Bayes procedure uses substantial foreign information for all focal countries, but the allocation is far from uniform. Average total donor weight is close to one for the Euro Area, Canada, Australia, and the US, and is somewhat lower for the UK. The breadth of the donor allocation also differs across countries. Australia and Canada borrow from relatively broad donor sets, whereas borrowing by the Euro Area and the US is more concentrated. The Euro Area is the largest average donor for Canada, Australia, the UK, and the US, while Canada receives the largest average discount when the Euro Area is the focal economy. These patterns support the interpretation of the power prior as a targeted borrowing device rather than a mechanical full-pooling rule: the model learns both how much foreign information to use and how to allocate the borrowing across donors.

Table 3: Estimated allocation of donor weights by focal country

Focal country	Total donor weight	Active donors	Concentration	Top donors
Euro Area	1.00	1.92	0.80	Canada, US
Canada	1.00	3.00	0.49	Euro Area, US
Australia	0.98	3.79	0.43	Euro Area, Canada
UK	0.73	2.85	0.72	Euro Area, Australia
US	1.00	2.00	0.64	Euro Area, Canada

Note: Diagnostics refer to the unrestricted power-prior specification with a diffuse baseline. Total donor weight is the average, across forecast origins, of the sum of the estimated donor-specific discounts. “Active donors” is the corresponding average number of donors receiving positive weight. Concentration is the median Herfindahl index of the normalized donor weights, with larger values indicating that borrowing is concentrated on fewer countries. Top donors are the countries receiving the largest average discount weights for the indicated focal country.

Overall, the forecasting exercise points to a clear conclusion. Cross-country information is useful in this application, and the power prior provides a parsimonious way to incorporate it without estimating a full multicountry system. The comparisons with the matching single-country priors show that the gains cannot be explained solely by standard Minnesota-type or ACP shrinkage. The restricted specifications further show that the results are not merely a consequence of mechanically adding foreign observations: although equal pooling captures part of the point-forecast improvement, allowing the data to select the relevant donor weights is especially valuable for density forecasts. Moreover, the forecast improvements are not accompanied by predictive undercoverage. We next apply the same borrowing logic to structural inference on the long-run Phillips curve.

5.2 Structural Inference with the Power Prior

Reassured by the improved reduced-form performance, we next investigate how the power prior changes the structural results in Benati (2015). To this end, we impose the same identifying restrictions as in the original paper in order to recover a monetary policy shock. First, shocks are divided into transitory and permanent shocks through zero restrictions on the long-run effect on inflation. The four permanent shocks, $(\varepsilon_t^{TE}, \varepsilon_t^{MO}, \varepsilon_t^{TA}, \varepsilon_t^{MK})$ (technology, monetary policy, taste, markup), are then distinguished from the two transitory shocks $(\varepsilon_t^{TR,1}, \varepsilon_t^{TR,2})$ by the sign restrictions summarized in Table 4, following Canova and Paustian

(2011) and Benati (2015).

Table 4: Impact sign restrictions used to identify the permanent shocks

	ε_t^{TE}	ε_t^{MO}	ε_t^{TA}	ε_t^{MK}
Output gap	+	−	+	−
Inflation	−	−	+	+
Unemployment rate	+	+	−	+
Short rate	−	+	+	+

Note: The table reports the impact sign restrictions used to distinguish the four permanent shocks. The shocks are technology (ε_t^{TE}), monetary policy (ε_t^{MO}), taste (ε_t^{TA}), and markup (ε_t^{MK}) shocks. No impact sign restrictions are imposed on the two transitory shocks.

We estimate the model using the zero- and sign-restriction algorithm of Arias, Rubio-Ramirez, and Waggoner (2018)⁷. For fixed or empirical-Bayes discount factors, implementing the power prior within this framework amounts simply to replacing current-data moments with weighted combinations of current and external-data moments. This highlights a key advantage of the approach: it allows foreign information to be incorporated while preserving the single-country structural system, so standard identification algorithms apply without modification, which would not be true for the panel and hierarchical benchmarks.

Figure 4 and Table 5 show that the power prior changes both posterior location and posterior precision. The posterior medians are pulled toward a narrow range of small negative values, between roughly -0.14 and -0.11 , whereas the diffuse-prior medians are more dispersed, ranging from -0.22 for the US to slightly positive for the Euro Area. At the same time, the most pronounced effect is a substantial reduction in posterior uncertainty, especially for the Euro Area and the US. Credible-interval length is reduced by up to 65%. Hence the power prior does not overturn Benati’s qualitative conclusion that the long-run effect is economically small, but it makes the evidence more concentrated around a mildly

⁷Note that the zero and sign restrictions only set-identify the impulse responses and thus the long-run trade-off. We nevertheless abstract from reporting robust Bayes results in the sense of Giacomini and Kitagawa (2021) for two reasons. First, and most importantly, under sign restrictions on multiple columns, it is not clear that the procedure in Giacomini and Kitagawa (2021) has the desirable frequentist properties. Similarly, Granziera, Moon, and Schorfheide (2018) and Gafarov, Meier, and Montiel Olea (2018) are not applicable in our case. Second, the power prior borrows reduced-form dynamics and we view the choice of priors for the orthogonal matrix as an independent problem. Using external evidence for the latter is not desirable, since even the current likelihood provides no information about it.

negative relationship, thus considerably sharpening his conclusion.

The reduction in credible intervals could in general be due to the empirical-Bayes implementation, which ignores uncertainty about the discount factor and can cause the credible intervals to be badly calibrated from a frequentist point of view. However, Figure 4 and Table 5 additionally display the results for a fully Bayesian approach. The long-run inflation–unemployment trade-off posteriors are remarkably similar under the empirical-Bayes and random-discount specifications. This similarity suggests that the tightening is not primarily an artifact of conditioning on empirical-Bayes discount estimates.

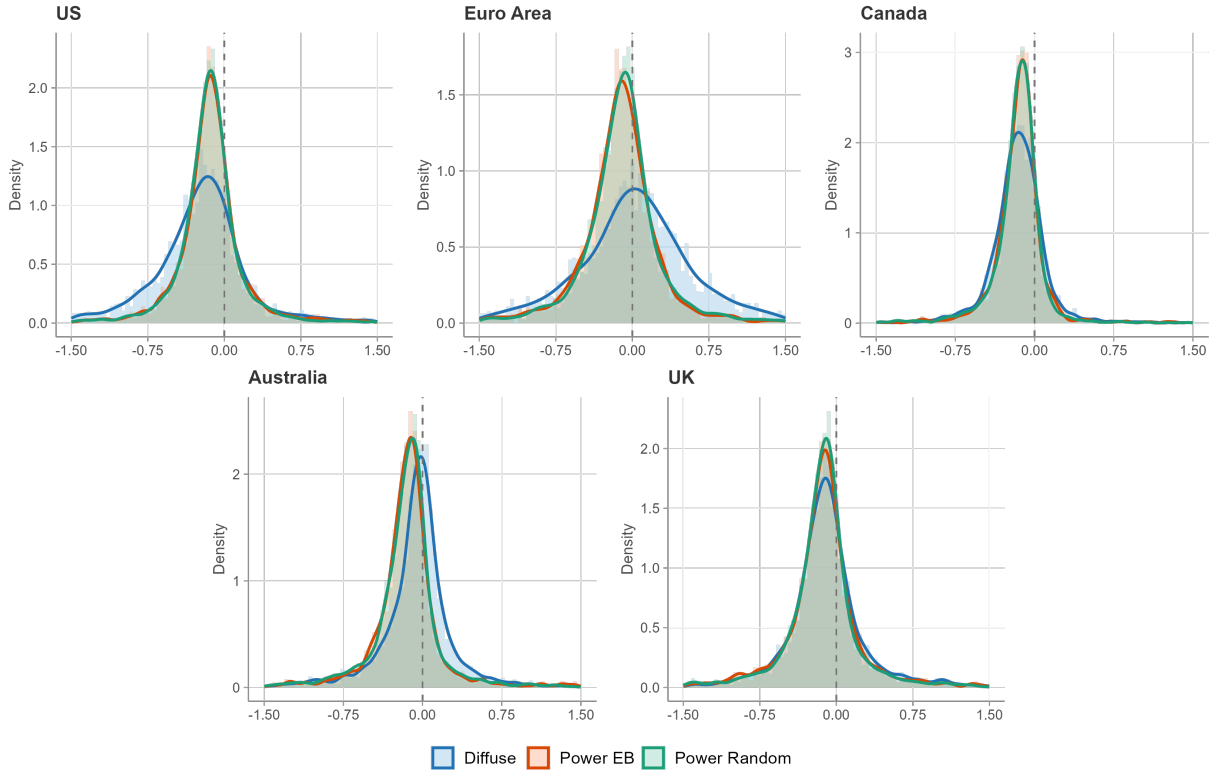


Figure 4: Posterior distribution of the long-run unemployment response to a monetary policy shock

Note: Posterior distribution of the long-run effect of a 1% monetary policy shock on unemployment for each focal country. Blue histograms show the diffuse baseline specification of Benati (2015). Red histograms show the power-prior specification with empirical-Bayes donor discounts. Green histograms show the power-prior specification with random donor discounts. Both power-prior specifications retain the same diffuse baseline prior and augment it with discounted data from the four remaining countries.

The posterior probability of a negative long-run unemployment response also increases in most countries, reaching roughly 65–85% under the power prior. The increase is especially visible for the Euro Area, where the diffuse-prior posterior assigns less than half of its mass to

negative values, while the power-prior posterior assigns about two thirds. Thus, cross-country borrowing both regularizes the location of the estimates and substantially sharpens inference.

Table 5: Long-run unemployment responses under diffuse and power-prior specifications

	Diffuse		Power EB				Power Random			
	$P(< 0)$	Median [5%, 95%]	$P(< 0)$	Median [5%, 95%]	CI reduction (%)		$P(< 0)$	Median [5%, 95%]	CI reduction (%)	
Australia	0.579	-0.037 [-1.055, 0.624]	0.796	-0.140 [-0.846, 0.413]	25.0		0.790	-0.130 [-0.855, 0.434]	23.2	
Canada	0.790	-0.149 [-0.600, 0.217]	0.841	-0.133 [-0.542, 0.163]	13.6		0.848	-0.135 [-0.586, 0.126]	12.7	
Euro Area	0.467	0.042 [-2.403, 2.598]	0.663	-0.110 [-1.018, 0.755]	64.6		0.654	-0.093 [-0.961, 0.756]	65.7	
UK	0.690	-0.113 [-0.972, 0.790]	0.731	-0.126 [-1.007, 0.640]	6.6		0.742	-0.124 [-0.937, 0.600]	12.8	
US	0.747	-0.221 [-1.753, 0.938]	0.756	-0.138 [-0.757, 0.502]	53.2		0.775	-0.145 [-0.773, 0.401]	56.4	

Note: Each row corresponds to a focal country. For the diffuse specification of Benati (2015), the table reports the posterior probability that the long-run unemployment response is negative ($P(< 0)$), the posterior median, and the 90% credible interval (5th–95th percentiles). The power-prior columns report the corresponding quantities under empirical-Bayes (Power EB) and random-discount (Power Random) donor weighting. CI reduction (%) is the percentage reduction in 90% credible-interval length relative to the diffuse specification.

The discount factors in Figure 5 show that this tightening is not the result of mechanical full pooling. The figure shows, in green, the marginal posterior of the donor-specific discount factor for each focal-country/donor-country pair, together with the empirical-Bayes estimate as a dashed vertical line. Donor relevance and uncertainty differ substantially across focal countries. Whereas the Euro Area borrows mainly from the US and vice versa, Australia, Canada, and the UK spread donor mass across countries.

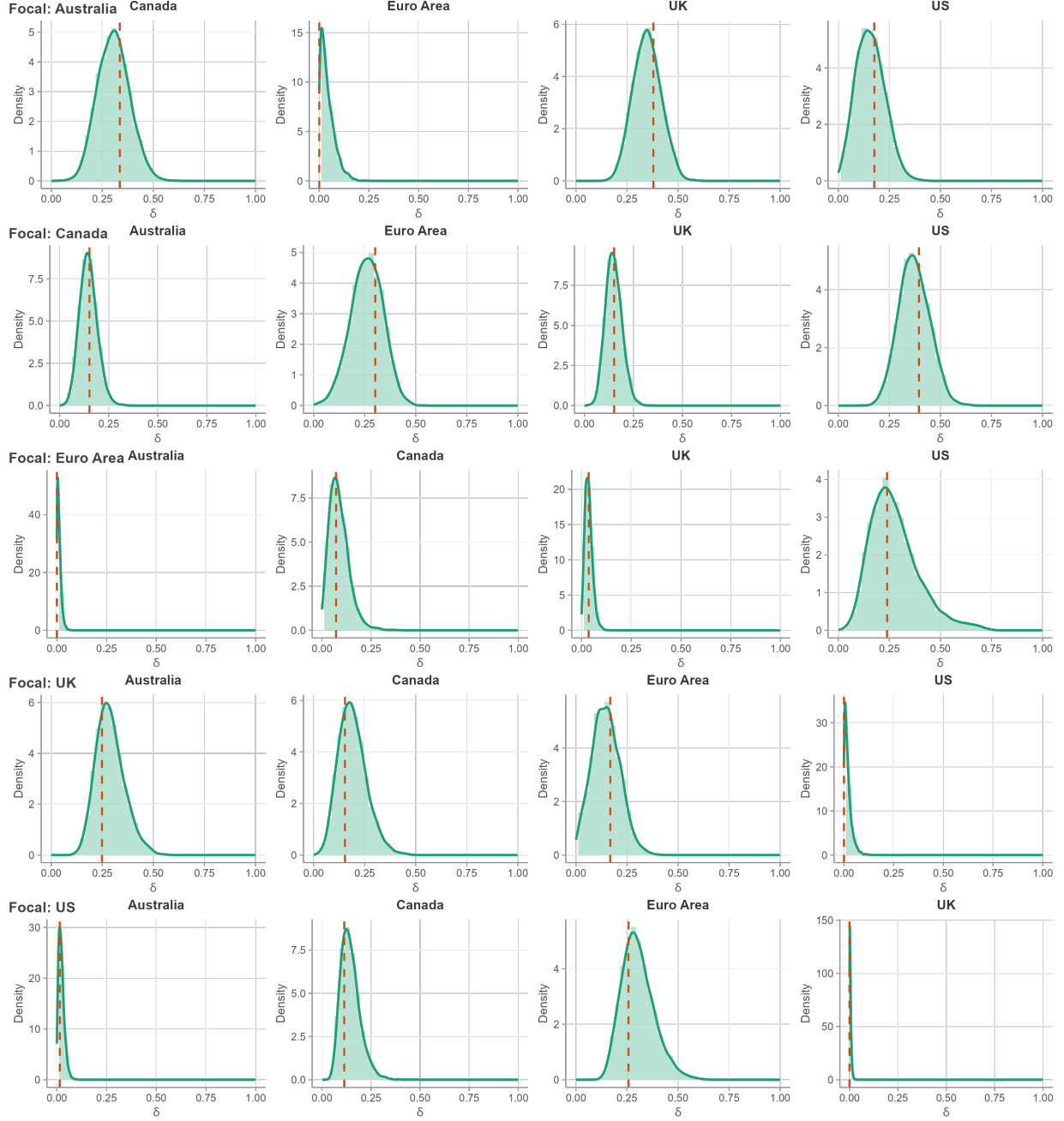


Figure 5: Posterior distribution of discount factors

Note: Each panel corresponds to a focal-country/donor-country pair; focal-country self-donor combinations are omitted. The green density is the marginal posterior of the donor-specific discount factor under the random-discount specification. The vertical dashed line marks the corresponding empirical-Bayes estimate.

6 Conclusion

This paper introduces a new class of data-based priors for VAR models that allows external information to be incorporated into the analysis of a focal country. For instance, the proposal

permits the use of US data in the analysis of the Euro Area, or more generally the use of foreign information in a way that remains targeted at marginal inference for one country at a time. We extend theoretical results from the power-prior literature and provide conditions under which the construction yields a proper prior for VARs under several common baseline prior choices.

The theoretical results also clarify the relationship between the power prior and existing approaches to cross-country pooling. Under suitable conditions, the focal-country posterior generated by the power prior coincides with the corresponding marginal posterior of a simple hierarchical model. This equivalence is local and specification-dependent, however. Unlike a joint hierarchical or panel model, the power prior does not estimate the full cross-country distribution and does not generate indirect borrowing through a pooling network. Instead, borrowing is direct and target-specific: each external dataset affects a focal-country posterior only through its own discount factor. The hedging results provide a complementary interpretation. External information can improve upon a misspecified baseline prior even when no individual donor matches the target data-generating process, provided that the baseline and donor information jointly move the prior in a useful direction. At the same time, donors that do not expand the relevant set of shrinkage directions may be redundant.

The Monte Carlo analysis illustrates how these mechanisms operate in finite samples. The empirical-Bayes discount declines sharply as donor dynamics move away from those of the target, so the procedure borrows substantially from similar donors but approaches the no-borrowing specification when the external data become clearly incompatible. When donors are informative, borrowing reduces estimation and impulse-response errors and shortens posterior intervals. These gains can be accompanied by some loss of frequentist coverage at intermediate donor distances, highlighting that adaptive borrowing is not costless and that posterior precision should be interpreted together with coverage. In the multi-donor design, the estimated discounts concentrate on the most relevant external sources and assign little weight to distant donors. Moreover, gains over the best-single-donor restriction are

concentrated in cases in which the empirical-Bayes allocation makes meaningful use of several donors, consistent with the multi-donor hedging mechanism.

The empirical application revisits the long-run Phillips-curve analysis of Benati (2015). In the recursive forecasting exercise, all three power-prior specifications improve point and density forecasts relative to the original diffuse single-country model. The versions based on Minnesota-type and asymmetric conjugate priors also improve upon their corresponding no-borrowing counterparts and perform competitively with the panel and hierarchical multi-country benchmarks. The gains are not obtained by producing excessively narrow predictive distributions: empirical coverage of the nominal 90% predictive intervals remains at or above the nominal rate. Comparisons with restricted donor allocations further show that useful foreign information is generally not confined to a single country. Equal pooling captures part of the point-forecast benefit, but unrestricted donor-specific discounting is particularly valuable for density forecasting.

Cross-country borrowing also materially changes structural inference. Applying the power prior within the sign- and zero-restricted SVAR concentrates the posterior estimates of the long-run unemployment response to a monetary policy shock around economically small negative values and substantially reduces posterior uncertainty, especially in the countries for which the original estimates are least precise. Thus, the results reinforce the conclusion that the long-run trade-off is small while providing more informative evidence about its likely sign and magnitude. Leave-one-donor-out estimates show that these conclusions are not driven by any individual foreign sample. Complementary forecasts based on the broader global VAR dataset and a different evaluation period, reported in the appendix, provide external-validity evidence, especially for GDP, while also showing that relative performance remains variable-specific.

Taken together, the results position the power prior as an operational middle ground between standard single-country Bayesian VARs and fully joint multicountry models. It is most useful when the object of interest is inference or forecasting for a particular country and

foreign datasets are intended to serve as auxiliary information rather than as components of a jointly identified international system. In such settings, the approach is parsimonious—its borrowing decisions are transparent through the estimated discount factors—and it preserves compatibility with familiar forecasting and structural-identification algorithms. It should therefore be viewed not as a replacement for hierarchical, panel, or global VARs, but as a modular extension of single-country VAR practice that can be added to a wide range of baseline priors.

Several extensions remain promising. The power prior could be combined with richer shrinkage specifications or applied to external information generated by structural economic models. Extending the framework to stochastic-volatility, time-varying-parameter, and other latent-variable VARs would broaden its empirical scope, although the normalizing constant and the empirical-Bayes criterion will generally no longer be available in closed form. Further work could also model dependence among donor likelihoods or discount factors, which may be important when foreign datasets contain overlapping information because of common shocks and tightly integrated economies. These extensions would retain the central principle developed here: external information should be borrowed selectively, transparently, and in a manner tailored to the inferential object of interest.

References

- Angelini, E. et al. (2019). “Mind the gap: A multi-country BVAR benchmark for the Eurosystem projections”. In: *International Journal of Forecasting* 35, pp. 1658–1668.
- Arias, J., J. Rubio-Ramirez, and D. Waggoner (2018). “Inference based on Structural Vector Autoregressions identified with Sign and Zero Restrictions: Theory and Applications”. In: *Econometrica* 86, pp. 685–720.
- (2021). “Inference in Bayesian Proxy-SVARs”. In: *Journal of Econometrics* 225, pp. 88–106.

- Bai, Y. et al. (2022). “Macroeconomic forecasting in a multicountry context”. In: *Journal of Applied Econometrics* 37.6, pp. 1230–1255.
- Banbetta, A., E. Lesaffre, and J. van Rosmalen (2022). “The power prior with multiple historical controls for the linear regression model”. In: *Pharmaceutical Statistics* 21, pp. 418–438.
- Barigozzi, M., C. Lissona, and L. Tonni (2024). “Large datasets for the Euro Area and its member countries and the dynamic effects of the common monetary policy”. In.
- Baumeister, C. and J. Hamilton (2019). “Structural Interpretation of Vector Autoregressions with Incomplete Identification: Revisiting the Role of Oil Supply and Demand Shocks”. In: *American Economic Review* 109.5, pp. 1873–1910.
- Benati, L. (2015). “The long-run Phillips curve: A structural VAR investigation”. In: *Journal of Monetary Economics* 76, pp. 15–28.
- Billio, M. et al. (2025). “Bayesian Outlier Detection for Matrix-variate Models”. In.
- Canova, F. and M. Ciccarelli (2013). “Panel Vector Autoregressive Models: A Survey”. In: *VAR models in macroeconomics—new developments and applications: Essays in honor of Christopher A. Sims*.
- Canova, F. and M. Paustian (2011). “Business cycle measurement with some theory”. In: *Journal of Monetary Economics* 58.4, pp. 345–361.
- Carriero, A., T. Clark, and M. Marcellino (2021). “No-arbitrage priors, drifting volatilities, and the term structure of interest rates”. In: *Journal of Applied Econometrics* 36.5.
- Carvalho, L. and J. Ibrahim (2021). “On the normalized power prior”. In: *Statistics in Medicine*.
- Chan, J. (2021). “Minnesota-type adaptive hierarchical priors for large Bayesian VARs”. In: *International Journal of Forecasting* 37.3, pp. 1212–1226.
- (2022). “Asymmetric conjugate priors for large Bayesian VARs”. In: *Quantitative Economics* 13, pp. 1145–1169.

- Chen, M., Z. Guan, et al. (2025). “Power Priors for Leveraging Historical Data: Looking Back and Looking Forward”. In: *Journal of Data Science* 23.1, pp. 1–30. DOI: 10.6339/24-JDS1161.
- Chen, M. and J. Ibrahim (2006). “The Relationship Between the Power Prior and Hierarchical Models”. In: *Bayesian Analysis* 1.3, pp. 551–574.
- Chudik, A. and H. Pesaran (2016). “Theory and Practice of GVAR Modelling”. In: *Journal of Economic Surveys* 30.1, pp. 165–197.
- Clark, T. and E. Mertens (2026). “Entropic Tilting of Forecasts to SPF Histograms: Analytics and Applications”. In.
- Clark, Todd E. and Kenneth D. West (2007). “Approximately Normal Tests for Equal Predictive Accuracy in Nested Models”. In: *Journal of Econometrics* 138.1, pp. 291–311. DOI: 10.1016/j.jeconom.2006.05.023.
- Del Negro, M. and F. Schorfheide (2004). “Priors from General Equilibrium Models for VARs”. In: *International Economic Review* 45.2.
- Diebold, F. and R. Mariano (2002). “Comparing Predictive Accuracy”. In: *Journal of Business and Economic Statistics* 20.1, pp. 134–144.
- Doan, T., B. Litterman, and C. Sims (1984). “Forecasting and Conditional Projection Using Realistic Prior Distributions”. In: *Econometric Reviews* 3, pp. 1–100.
- Duan, Y. (2005). “A modified Bayesian power prior with applications in water quality evaluation”. In.
- Duan, Y., K. Ye, and E. Smith (2006). “Evaluating water quality using power priors to incorporate historical information”. In: *Environmetrics* 17, pp. 95–106.
- Feldkircher, M. et al. (2022). “Approximate bayesian inference and forecasting in huge-dimensional multicountry VARs”. In: *International Economic Review* 63.4, pp. 1625–1658.
- Gafarov, B., M. Meier, and J. Montiel Olea (2018). “Delta-method inference for a class of set-identified SVARs”. In: *Journal of Econometrics* 203, pp. 316–327.

- George, E., D. Sun, and S. Ni (2008). “Bayesian stochastic search for VAR model restrictions”. In: *Journal of Econometrics* 142.1, pp. 553–580.
- Giacomini, R. and T. Kitagawa (2021). “Robust Bayesian Inference for Set-Identified Models”. In: *Econometrica* 89, pp. 1519–1556.
- Giannone, D., M. Lenza, and G. Primiceri (2015). “Prior Selection for Vector Autoregressions”. In: *The Review of Economics and Statistics* 97.2, pp. 436–451.
- (2019). “Priors for the Long Run”. In: *Journal of the American Statistical Association* 114.526, pp. 565–580.
- Granziera, E., H. Moon, and F. Schorfheide (2018). “Inference for VARs identified with Sign Restrictions”. In: *Quantitative Economics* 9.3, pp. 1087–1121.
- Gravestock, I. and L. Held (2017). “Adaptive power priors with empirical Bayes for clinical trials”. In: *Pharmaceutical Statistics* 16, pp. 349–360.
- Han, Z., K. Ye, and M. Wang (2023). “A Study on the Power Parameter in Power Prior Bayesian Analysis”. In: *The American Statistician* 77.1, pp. 12–19.
- Huber, F. and M. Feldkircher (2019). “Adaptive Shrinkage in Bayesian Vector Autoregressive Models”. In: *Journal of Business and Economic Statistics* 37.1, pp. 27–39.
- Ibrahim, J., M. Chen, Y. Gwon, et al. (2015). “The power prior: theory and applications”. In: *Statistics in Medicine* 34, pp. 3724–3749.
- Ibrahim, J., M. Chen, and D. Sinha (2003). “On Optimality Properties of the Power Prior”. In: *Journal of the American Statistical Association* 98.461, pp. 204–213.
- Kiley, M. (2024). “Anchored or Not: How Much Information Does 21st Century Data Contain on Inflation Dynamics?” In: *International Journal of Central Banking* 20.1, pp. 239–261.
- Koop, G. and D. Korobilis (2019). “Forecasting with High-Dimensional Panel VARs”. In: *Oxford Bulletin of Economics and Statistics* 81.5, pp. 937–959.
- Korobilis, D. (2016). “Prior Selection for Panel Vector Autoregressions”. In: *Computational Statistics and Data Analysis* 101, pp. 110–120.

- Korobilis, D. and D. Pettenuzzo (2019). “Adaptive Hierarchical Priors for High-Dimensional Vector Autoregressions”. In: *Journal of Econometrics* 212.1, pp. 241–271.
- Litterman, R. (1986). “Forecasting with Bayesian vector autoregressions — five years of experience”. In: *Journal of Business and Economic Statistics* 4.1, pp. 25–38.
- Mohaddes, K. and M. Raissi (2024). “Compilation, Revision and Updating of the Global VAR (GVAR) Database, 1979Q2-2023Q3.” In: *University of Cambridge: Judge Business School*.
- Müller, U. and M. Watson (2025). “Forecasting Related Time Series”. In: *Journal of Applied Econometrics*.
- Pawel, S. et al. (2023). “Normalized power priors always discount historical data”. In: *Stat* 12.
- Pfarrhofer, M. (2023). “Measuring International Uncertainty Using Global Vector Autoregressions with Drifting Parameters”. In: *Macroeconomic Dynamics* 27, pp. 770–793.
- Prüser, J. and B. Blagov (2026). “Improving inference and forecasting in VAR models using cross-sectional information”. In: *Economic Modelling* 160.107618.
- Shen, Y. et al. (2024). “Exploring the Connection Between the Normalized Power Prior and Bayesian Hierarchical Models”. In: *arXiv preprint arXiv:2404.02453*.
- (2026). “Optimal Priors for the Discounting Parameter of the Normalized Power Prior”. In: *Statistica Sinica* 36, pp. 843–866.
- Ye, K. et al. (2022). “Normalized power prior Bayesian analysis”. In: *Journal of Statistical Planning and Inference* 216, pp. 29–50.
- Zellner, A. (2002). “Information processing and Bayesian analysis”. In: *Journal of Econometrics* 107, pp. 41–50.

A Proofs for Section 3

In the following, we write $A \succ 0$ to denote that A is positive definite.

Proof of Proposition 1. We begin by deriving the normalizing constant

$$C(\delta) := \iint p(Y_0 \mid A, \Sigma)^\delta \pi_0(A, \Sigma) dA d\Sigma.$$

Throughout, let $\hat{A}_0 = (Z_0' Z_0)^{-1} Z_0' Y_0$ and $S_{Y_0} = Y_0' Y_0 - \hat{A}_0' Z_0' Z_0 \hat{A}_0 = Y_0' M_{Z_0} Y_0$, with $M_{Z_0} = I_{T_0} - Z_0 (Z_0' Z_0)^{-1} Z_0'$. We use the n -variate gamma function $\Gamma_n(a) = \pi^{n(n-1)/4} \prod_{i=1}^n \Gamma(a - (i-1)/2)$, defined for $a > (n-1)/2$. Let $Y_0 \in \mathbb{R}^{T_0 \times n}$, $Z_0 \in \mathbb{R}^{T_0 \times k}$ (full column rank), $A \in \mathbb{R}^{k \times n}$, and $\Sigma \in \mathbb{S}_{++}^n$. Write the powered likelihood

$$p(Y_0 \mid A, \Sigma)^\delta = (2\pi)^{-\frac{\delta T_0 n}{2}} |\Sigma|^{-\frac{\delta T_0}{2}} \exp\left(-\frac{\delta}{2} \text{tr}[\Sigma^{-1}(Y_0 - Z_0 A)'(Y_0 - Z_0 A)]\right).$$

Let us first integrate out A . Using $\text{tr}(B'C) = \text{vec}(B)'\text{vec}(C)$ and $\text{vec}(BAC) = (C' \otimes B)\text{vec}(A)$, one obtains

$$\begin{aligned} \text{tr}(\Sigma^{-1}(Y_0 - Z_0 A)'(Y_0 - Z_0 A)) &= \text{tr}(\Sigma^{-1} Y_0' Y_0) - 2 \text{vec}(A)' \text{vec}(Z_0' Y_0 \Sigma^{-1}) \\ &\quad + \text{vec}(A)' (Z_0' Z_0 \otimes \Sigma^{-1}) \text{vec}(A). \end{aligned}$$

Let $K := Z_0' Z_0 \otimes \Sigma^{-1}$ and $b := \text{vec}(Z_0' Y_0 \Sigma^{-1})$. Completing the square with $\mu := K^{-1}b = \text{vec}(\hat{A}_0)$ yields

$$\exp\left(-\frac{\delta}{2} \text{tr}(\Sigma^{-1}(Y_0 - Z_0 A)'(Y_0 - Z_0 A))\right) = e^{-\frac{\delta}{2} \text{tr}(\Sigma^{-1} S_{Y_0})} \exp\left(-\frac{1}{2}(\text{vec}(A) - \mu)'(\delta K)(\text{vec}(A) - \mu)\right).$$

Therefore

$$\begin{aligned} & \int_{\mathbb{R}^{k \times n}} \exp\left(-\frac{\delta}{2} \text{tr}(\Sigma^{-1}(Y_0 - Z_0 A)'(Y_0 - Z_0 A))\right) dA \\ &= (2\pi)^{\frac{kn}{2}} |\delta K|^{-\frac{1}{2}} e^{-\frac{\delta}{2} \text{tr}(\Sigma^{-1} S_{Y_0})} = (2\pi)^{\frac{kn}{2}} |\delta Z_0' Z_0|^{-\frac{n}{2}} |\Sigma|^{\frac{k}{2}} e^{-\frac{\delta}{2} \text{tr}(\Sigma^{-1} S_{Y_0})}. \end{aligned} \quad (8)$$

The last equality uses $|Z_0' Z_0 \otimes \Sigma^{-1}| = |Z_0' Z_0|^n |\Sigma^{-1}|^k$.

Next we integrate out Σ . We will use the standard identity, valid for $S \succ 0$ and $\nu > n - 1$,

$$\int_{S_{++}^n} |\Sigma|^{-\frac{\nu+n+1}{2}} \exp\left(-\frac{1}{2} \text{tr}(S \Sigma^{-1})\right) d\Sigma = |S|^{-\frac{\nu}{2}} 2^{\frac{n\nu}{2}} \Gamma_n\left(\frac{\nu}{2}\right). \quad (9)$$

Combining (8) with the baseline prior $\pi_0(A, \Sigma)$ gives $C(\delta)$ case by case:

(a) *Diffuse baseline* $\pi_0(A, \Sigma) \propto 1$. From (8),

$$C(\delta) = (2\pi)^{-\frac{\delta T_0 n}{2}} (2\pi)^{\frac{kn}{2}} |\delta Z_0' Z_0|^{-\frac{n}{2}} \int |\Sigma|^{-\frac{\delta T_0 - k}{2}} e^{-\frac{\delta}{2} \text{tr}(\Sigma^{-1} S_{Y_0})} d\Sigma.$$

Match the integrand to (9) with $\nu = \delta T_0 - k - n - 1$ and $S = \delta S_{Y_0}$. Then

$$C(\delta) = (2\pi)^{-\frac{\delta T_0 n}{2}} (2\pi)^{\frac{kn}{2}} |\delta Z_0' Z_0|^{-\frac{n}{2}} |\delta S_{Y_0}|^{-\frac{\nu}{2}} 2^{\frac{n\nu}{2}} \Gamma_n\left(\frac{\nu}{2}\right), \quad \nu = \delta T_0 - k - n - 1. \quad (10)$$

(b) *Jeffreys baseline* $\pi_0(A, \Sigma) \propto |\Sigma|^{-(n+1)/2}$. Relative to (a), the power on $|\Sigma|$ becomes $-(\delta T_0 - k + n + 1)/2$. Thus with $S = \delta S_{Y_0}$ and

$$\nu = \delta T_0 - k \quad (11)$$

$$C(\delta) = (2\pi)^{-\frac{\delta T_0 n}{2}} (2\pi)^{\frac{kn}{2}} |\delta Z_0' Z_0|^{-\frac{n}{2}} |\delta S_{Y_0}|^{-\frac{\nu}{2}} 2^{\frac{n\nu}{2}} \Gamma_n\left(\frac{\nu}{2}\right). \quad (12)$$

(c) *NIW baseline* $\pi_0(A | \Sigma) = MN(\bar{A}, \Omega, \Sigma)$ and $\pi_0(\Sigma) = \text{IW}(\nu_0, S_0)$. Multiplying $p(Y_0 | A, \Sigma)^\delta$

by $MN(\bar{A}, \Omega, \Sigma)$, the quadratic form in A has precision $\Omega^{-1} + \delta Z_0' Z_0$ and mean

$$\tilde{A}(\delta) = (\Omega^{-1} + \delta Z_0' Z_0)^{-1} (\Omega^{-1} \bar{A} + \delta Z_0' Y_0).$$

Integrating out A and multiplying by the inverse-Wishart prior $\text{IW}(\nu_0, S_0)$, define

$$\tilde{S}(\delta) = S_0 + \delta Y_0' Y_0 + \bar{A}' \Omega^{-1} \bar{A} - \tilde{A}(\delta)' (\Omega^{-1} + \delta Z_0' Z_0) \tilde{A}(\delta),$$

$$\tilde{\nu}(\delta) = \nu_0 + \delta T_0.$$

Using (9) with $S = \tilde{S}(\delta)$ and $\nu = \tilde{\nu}(\delta)$, we obtain, up to a δ -independent constant,

$$C(\delta) \propto (2\pi)^{-\delta T_0 n/2} |\Omega^{-1} + \delta Z_0' Z_0|^{-n/2} |\tilde{S}(\delta)|^{-\tilde{\nu}(\delta)/2} 2^{n\tilde{\nu}(\delta)/2} \Gamma_n\left(\frac{\tilde{\nu}(\delta)}{2}\right). \quad (13)$$

The δ -independent factor collects terms such as $|\Omega|$ and the fixed normalizing constant of $\text{IW}(\nu_0, S_0)$.

The normalized power prior exists iff $0 < C(\delta) < \infty$, see Carvalho and Ibrahim (2021). From (10)–(13) and the fact that $\Gamma_n(a)$ is finite iff $a > (n-1)/2$, we obtain:

(a) *Diffuse prior*. Finiteness requires $\nu > n-1$ with $\nu = \delta T_0 - k - n - 1$, i.e.

$$\delta > \frac{k + 2n}{T_0}.$$

(b) *Jeffreys prior*. Finiteness requires $\nu > n-1$ with $\nu = \delta T_0 - k$, i.e.

$$\delta > \frac{k + n - 1}{T_0}.$$

(c) *NIW prior*. Here $\tilde{\nu}(\delta) = \nu_0 + \delta T_0$ and the baseline is proper, so

$$C(\delta) < \infty \text{ for all } \delta \geq 0.$$

□

Proof of Proposition 2. The marginal likelihood is

$$ML(\delta) \equiv \iint p(Y_1 \mid A, \Sigma) \pi(A, \Sigma \mid Y_0, \delta) dA d\Sigma.$$

Equivalently, $ML(\delta)$ is the ratio of the normalization integral based on the powered external likelihood plus the current likelihood to the external-data normalization constant $C(\delta)$:

$$ML(\delta) = \frac{1}{C(\delta)} \iint p(Y_0 \mid A, \Sigma)^\delta p(Y_1 \mid A, \Sigma) \pi_0(A, \Sigma) dA d\Sigma.$$

The integrations over A and Σ proceed exactly as in the derivation of $C(\delta)$, with δT_0 replaced by $\delta T_0 + T_1$ and the cross-products updated as $Z'_0 Z_0 \mapsto \delta Z'_0 Z_0 + Z'_1 Z_1$, $Z'_0 Y_0 \mapsto \delta Z'_0 Y_0 + Z'_1 Y_1$. The resulting expressions are stated up to a δ -independent constant, matching Proposition 2.

(a) *Diffuse baseline.* With the quantities $S_0(\delta)$, $S_1(\delta)$, $v_0(\delta)$, $v_1(\delta)$, and $\tilde{A}(\delta)$ defined in Proposition 2,

$$ML(\delta) \propto \left(\frac{|\delta Z'_0 Z_0|}{|\delta Z'_0 Z_0 + Z'_1 Z_1|} \right)^{\frac{n}{2}} |S_0(\delta)|^{\frac{v_0(\delta)}{2}} |S_1(\delta)|^{-\frac{v_1(\delta)}{2}} \frac{\Gamma_n\left(\frac{v_1(\delta)}{2}\right)}{\Gamma_n\left(\frac{v_0(\delta)}{2}\right)}.$$

(b) *Jeffreys baseline.* The same expression holds, with $v_0(\delta) = \delta T_0 - k$ and $v_1(\delta) = \delta T_0 + T_1 - k$.

(c) *NIW baseline.* With the quantities $\tilde{A}(\delta)$, $\hat{A}(\delta)$, $\tilde{S}(\delta)$, $\hat{S}(\delta)$, $\tilde{\nu}(\delta)$, and $\hat{\nu}(\delta)$ defined in Proposition 2,

$$ML(\delta) \propto \left(\frac{|\Omega^{-1} + \delta Z'_0 Z_0|}{|\Omega^{-1} + \delta Z'_0 Z_0 + Z'_1 Z_1|} \right)^{\frac{n}{2}} |\tilde{S}(\delta)|^{\frac{\tilde{\nu}(\delta)}{2}} |\hat{S}(\delta)|^{-\frac{\hat{\nu}(\delta)}{2}} \frac{\Gamma_n\left(\frac{\hat{\nu}(\delta)}{2}\right)}{\Gamma_n\left(\frac{\tilde{\nu}(\delta)}{2}\right)}.$$

□

Proof of Proposition 3. Let

$$V_j = I_{T_j} \otimes \Sigma_j, \quad Q_b = X_b' V_b^{-1} X_b, \quad q_b = X_b' V_b^{-1} y_b.$$

Under the flat prior $p(\bar{\beta}) \propto 1$, integrating out the common mean from the two Gaussian random effects gives

$$\beta_b \mid \beta_a \sim N(\beta_a, 2\Omega).$$

Therefore,

$$y_b \mid \beta_a \sim N(X_b \beta_a, V_b + 2X_b \Omega X_b').$$

Let $W_b = V_b + 2X_b \Omega X_b'$. Using the Woodbury identity,

$$\begin{aligned} X_b' W_b^{-1} X_b &= Q_b (I_{n_k} + 2\Omega Q_b)^{-1}, \\ X_b' W_b^{-1} y_b &= (I_{n_k} + 2Q_b \Omega)^{-1} q_b. \end{aligned}$$

The hierarchical donor contribution to the posterior kernel in β_a therefore has quadratic and linear terms determined by these two expressions. The powered donor likelihood $p(y_b \mid \beta_a)^\delta$ has corresponding terms δQ_b and δq_b . Because Q_b is nonsingular under the full-column-rank assumption, equality for all y_b is equivalent to

$$\delta(I_{n_k} + 2\Omega Q_b) = I_{n_k}.$$

The equivalent condition with $Q_b \Omega$ follows as well, because the displayed equality implies that ΩQ_b , and hence $Q_b \Omega$, is a scalar multiple of the identity. In particular, the g -prior $\Omega = gQ_b^{-1}$ yields $\delta = 1/(1 + 2g)$. □

Proof of Proposition 6. The result follows directly from the same conjugate calculations as above and the factorization in Chan (2022). $\square$

Proof of Proposition 4. Let n_0 be the number of external observations and μ_0 their sample mean. Under the model and baseline prior stated in the proposition,

$$\pi_\delta(\theta \mid D_0, \tau) = N(m(\delta), v(\delta)),$$

where

$$m(\delta) = \frac{\tau + \delta n_0 \mu_0}{1 + \delta n_0}, \quad v(\delta) = \frac{\sigma^2}{1 + \delta n_0}.$$

The KL divergence is

$$K(\delta) = \frac{1}{2} \left[\log \left(\frac{v(\delta)}{\sigma^2} \right) + \frac{\sigma^2 + (\tau^* - m(\delta))^2}{v(\delta)} - 1 \right].$$

A direct differentiation at $\delta = 0$ gives

$$K'(0) = \frac{n_0}{2\sigma^2} [(\tau^* - \tau)^2 - 2(\tau^* - \tau)(\mu_0 - \tau)].$$

Define $\Delta^* = \tau^* - \tau$ and $\Delta_0 = \mu_0 - \tau$. Then

$$K'(0) = \frac{n_0}{2\sigma^2} \Delta^* (\Delta^* - 2\Delta_0).$$

If $\tau < \tau^* < \mu_0$, then $0 < \Delta^* < \Delta_0$, so $\Delta^* - 2\Delta_0 < 0$ and $K'(0) < 0$. If $\tau > \tau^* > \mu_0$, then $\Delta_0 < \Delta^* < 0$, so $\Delta^* - 2\Delta_0 > 0$ and again $K'(0) < 0$. By continuity, there exists a sufficiently small $\delta \in (0, 1]$ for which $K(\delta) < K(0)$. $\square$

Proof of Proposition 5. For every $m \in \mathbb{R}^r$, the distributions $\pi^* = N_r(\tau^*, \Sigma)$ and $\pi_m =$

$N_r(m, \Sigma)$ have the same covariance matrix. Therefore,

$$\begin{aligned}\text{KL}(\pi^* \parallel \pi_m) &= \frac{1}{2}(\tau^* - m)' \Sigma^{-1}(\tau^* - m) \\ &= \frac{1}{2} \|\tau^* - m\|_{\Sigma^{-1}}^2.\end{aligned}$$

Hence

$$R_J = \inf_{m \in \mathcal{H}_J} \frac{1}{2} \|\tau^* - m\|_{\Sigma^{-1}}^2 = \frac{1}{2} \|\tau^* - P_J(\tau^*)\|_{\Sigma^{-1}}^2. \quad (14)$$

The last equality follows because $\mathcal{H}_J$ is a nonempty compact convex set. Moreover, since Σ is positive definite, the squared norm $\|\tau^* - m\|_{\Sigma^{-1}}^2$ is strictly convex in m . Consequently, the projection $P_J(\tau^*)$ exists and is unique.

For part (a), adding the donor μ_{J+1} enlarges the feasible set:

$$\mathcal{H}_J \subseteq \mathcal{H}_{J+1}.$$

It follows that

$$\begin{aligned}R_{J+1} &= \inf_{m \in \mathcal{H}_{J+1}} \frac{1}{2} \|\tau^* - m\|_{\Sigma^{-1}}^2 \\ &\leq \inf_{m \in \mathcal{H}_J} \frac{1}{2} \|\tau^* - m\|_{\Sigma^{-1}}^2 = R_J.\end{aligned}$$

For part (b), first suppose that $P_{J+1}(\tau^*) \in \mathcal{H}_J$. Since this point is feasible for the minimization defining R_J ,

$$R_J \leq \frac{1}{2} \|\tau^* - P_{J+1}(\tau^*)\|_{\Sigma^{-1}}^2 = R_{J+1}.$$

Part (a) gives the reverse inequality, and therefore $R_{J+1} = R_J$.

Conversely, suppose that $R_{J+1} = R_J$. Because $P_J(\tau^*) \in \mathcal{H}_J \subseteq \mathcal{H}_{J+1}$, equation (14)

implies

$$\frac{1}{2} \|\tau^* - P_J(\tau^*)\|_{\Sigma^{-1}}^2 = R_J = R_{J+1}.$$

Thus $P_J(\tau^*)$ also minimizes the distance from τ^* over $\mathcal{H}_{J+1}$. By uniqueness of the projection onto $\mathcal{H}_{J+1}$,

$$P_{J+1}(\tau^*) = P_J(\tau^*) \in \mathcal{H}_J.$$

We have therefore shown that

$$R_{J+1} = R_J \iff P_{J+1}(\tau^*) \in \mathcal{H}_J.$$

Together with $R_{J+1} \leq R_J$, this is equivalent to

$$R_{J+1} < R_J \iff P_{J+1}(\tau^*) \notin \mathcal{H}_J.$$

For part (c), if $\mu_{J+1} \in \mathcal{H}_J$, then, because $\mathcal{H}_J$ is convex,

$$\mathcal{H}_{J+1} = \text{conv}(\mathcal{H}_J \cup \{\mu_{J+1}\}) = \mathcal{H}_J.$$

The two optimization problems are therefore identical, and hence

$$R_{J+1} = R_J.$$

□

B Empirical Appendix

B.1 Benati Application: Leave-one-donor-out Robustness Exercise

The leave-one-out exercise reported in Table 6 shows that the structural findings are not driven by any single donor country. For each focal country, we re-estimate the empirical-Bayes discount factors and the structural posterior after omitting one donor at a time. The resulting posterior medians remain close to the full donor-pool estimates: the largest absolute median shift is 0.042, for the Euro Area when the US is omitted, and all other shifts are around 0.02 or less. The posterior probability of a negative long-run effect is also stable, ranging between 0.66 and 0.90 across all leave-one-out specifications. The most influential omitted donors are economically plausible and broadly correspond to the largest full-sample discount factors. Hence the power-prior results reflect information from the donor pool as a whole, rather than dependence on a single foreign sample.

Table 6: Leave-one-donor-out robustness of the structural estimates

Focal country	LOO median range	LOO 90% CI range	LOO $P(< 0)$ range	CI length range	Most influential omitted donor	Max median shift
Euro Area	[-0.149, -0.090]	[-1.257, 0.945]	[0.70, 0.80]	[0.702, 2.202]	US	0.042
Canada	[-0.154, -0.112]	[-0.596, 0.144]	[0.80, 0.90]	[0.627, 0.739]	UK	0.023
Australia	[-0.146, -0.124]	[-0.950, 0.660]	[0.72, 0.82]	[1.133, 1.611]	Euro Area	0.015
UK	[-0.138, -0.095]	[-1.232, 0.793]	[0.66, 0.79]	[1.226, 2.026]	Euro Area	0.022
US	[-0.169, -0.135]	[-0.986, 0.776]	[0.72, 0.80]	[0.778, 1.762]	Canada	0.019

Note: For each focal country, the power-prior model is re-estimated four times, each time omitting one of the available donor countries and re-optimizing the empirical-Bayes discount factors. The reported ranges give the minimum and maximum posterior median, 90% credible interval, posterior probability of a negative response, and credible-interval length across these four specifications. The most influential omitted donor is the omission producing the largest absolute change in the posterior median relative to the full donor-pool estimate. The final column reports the magnitude of that maximum absolute median shift.

B.2 Benati Application: Detailed Forecast Tables

This subsection reports the detailed pseudo-out-of-sample forecast comparison for the long-run Phillips curve application of Benati (2015). All tables refer to the evaluation period 1985Q1–2000Q1. The challenger is the power-prior specification with a Minnesota baseline. Each table compares this specification with one benchmark and reports relative RMSEs and log-score differences side by side. In the left block, entries are ratios of power-prior to benchmark RMSE; values below one favor the power prior. In the right block, entries are

differences in average log predictive scores (power prior minus benchmark); positive values favor the power prior.

Because the power-prior Minnesota specification nests the corresponding single-country Minnesota model when the external-data discount is set to zero, standard equal-predictive-accuracy tests based on squared forecast errors can be biased against the larger model. For this comparison, we therefore use the one-sided test of T. E. Clark and West (2007), which adjusts the squared-error differential for the additional estimation noise associated with the larger nested specification. We compute Newey–West standard errors with lag length $h - 1$ at forecast horizon h . The Clark–West adjustment is specific to squared-error loss and is not directly applicable to density forecasts. Significance for log predictive score differences is therefore assessed using two-sided tests of the mean score differential with HAC standard errors. For the remaining, non-nested benchmark comparisons, the same HAC equal-predictive-accuracy test is used for both squared-error and log-score differentials.

Table 7: Relative RMSE and log-score difference: power prior vs. Minnesota baseline

	Relative RMSE						LPS difference					
	Y	Dp	DU	DR	CY	S	Y	Dp	DU	DR	CY	S
Euro Area												
h=1	0.89 ***	0.99 *	1.14	1.03 *	0.89 ***	0.91 ***	0.12 **	-0.03	-0.28 ***	-0.19 ***	0.05	-0.03
h=2	0.9 ***	0.98	1.11	0.97 **	0.85 ***	0.89 **	0.08	-0.05 **	-0.23 **	-0.12 ***	0.06	-0.01
h=3	0.91 ***	0.99	1.11	0.97 **	0.81 ***	0.84 **	0.12	-0.06 ***	-0.23 *	-0.14 ***	0.11	0.03
h=4	0.91 **	0.99 **	1.11	0.94 **	0.78 ***	0.82 ***	0.14	-0.07 ***	-0.22 **	-0.12 ***	0.13	0.03
h=5	0.92 **	1	1.1	0.94 **	0.74 ***	0.81 ***	0.18	-0.08 ***	-0.2 **	-0.12 ***	0.17 *	0.07
h=6	0.93 *	1.01	1.1	0.94 **	0.73 ***	0.8 ***	0.17	-0.08 ***	-0.18 **	-0.13 ***	0.16 *	0.11
h=7	0.95	1	1.04	0.96 **	0.74 ***	0.8 ***	0.16	-0.06 ***	-0.14 *	-0.14 ***	0.16 *	0.12
h=8	0.97	1.01	1.04	0.96 **	0.76 ***	0.82 ***	0.13	-0.07 ***	-0.12	-0.14 ***	0.13 *	0.13
Canada												
h=1	0.94 **	1.04	0.95 *	1.04	0.94 **	0.96 **	-0.04	-0.06	0.04 *	-0.03	0	0.03
h=2	0.94 **	1.02	0.98 *	1	0.94 ***	0.97 **	-0.02	-0.09	0.04 **	-0.01	-0.02	0.04
h=3	0.94 **	1.02	1	0.97 ***	0.9 ***	0.95 ***	0.01	-0.11 **	0.03 *	0	-0.01	0.06 ***
h=4	0.96 *	1	1.04	0.99	0.9 ***	0.95 ***	0.01	-0.1	0.02	-0.02	0	0.08 ***
h=5	0.98	1	1.02	1.02	0.88 ***	0.94 ***	0.01	-0.08	0.02	-0.03 ***	0.03	0.08 **
h=6	1	1 **	1.02	1.01	0.86 ***	0.93 ***	-0.02	-0.05	0.03 *	-0.01	0.05	0.1 ***
h=7	1.02	1 **	1.01	1.01	0.86 ***	0.92 **	-0.04	-0.01	0.02 **	-0.02 ***	0.07	0.1
h=8	1.05	1	0.99	1	0.84 ***	0.92 **	-0.05	-0.05	0.03 **	-0.01 *	0.09	0.1
Australia												
h=1	0.9 ***	0.89 ***	1.03	0.95 **	0.96 ***	0.92 ***	0.08 **	0.11 ***	-0.04	0.05	0.02	0.09 ***
h=2	0.91 ***	0.93 ***	1.06	1.05	0.94 ***	0.91 ***	0.06	0.09 ***	-0.06 **	-0.03	0.01	0.12 **
h=3	0.94 ***	0.99 *	1.11	1.05	0.93 ***	0.91 ***	0.02	0.07 *	-0.08 ***	-0.05	0.02	0.14 ***
h=4	0.96 ***	0.99 **	1.09	1.12	0.91 **	0.92 ***	-0.01	0.06	-0.06 ***	-0.09 *	0.02	0.14 ***
h=5	1.01 **	0.98 ***	1.02	1.08	0.86 ***	0.93 ***	-0.04	0.05	-0.01	-0.07	0.08	0.13 ***
h=6	1.07	0.99 *	0.96	1.17	0.83 ***	0.93 ***	-0.06	0.07	0.02	-0.12 **	0.11	0.12
h=7	1.09	1	0.92 **	1.19	0.79 ***	0.95 ***	-0.07	0.09 ***	0.07 *	-0.16 ***	0.16	0.08
h=8	1.07	1	0.89 ***	1.29	0.77 ***	0.97 **	-0.06	0.1 ***	0.07 **	-0.2 ***	0.2	0.07 *
UK												
h=1	0.86 ***	0.93 ***	1.21	0.87 ***	0.99 **	0.93 ***	0.06 ***	0.1 ***	-0.16	0.03	0.01	0.06 **
h=2	0.86 ***	0.98 **	1.2	0.77 ***	0.98 **	0.92 ***	0.05 *	0.09 ***	-0.08	0.13	0	0.08 *
h=3	0.9 ***	0.95 ***	1.17	0.74 ***	0.95 **	0.9 ***	0.05	0.1 ***	-0.02	0.21	0.04	0.11
h=4	0.91 **	1.01	1.14	0.75 ***	0.9 **	0.89 **	0.04	0.1 ***	0.06	0.23	0.08	0.12
h=5	0.94 **	0.98 **	1.13	0.76 ***	0.89 **	0.87 **	0.04	0.11 ***	0.05	0.24	0.09	0.13
h=6	0.96 **	0.98 ***	1.08 *	0.76 ***	0.88 **	0.86 *	0.04	0.11 ***	0.07	0.25	0.11	0.13
h=7	0.97 *	0.97 ***	1.05 *	0.77 ***	0.88 **	0.86 *	0.03	0.11 ***	0.1	0.26	0.13	0.12
h=8	0.96 **	0.97 ***	1.02 **	0.78 **	0.87 **	0.86 *	0.03	0.11 ***	0.08	0.28	0.15	0.11
US												
h=1	0.89 ***	0.95 ***	0.84 ***	0.83 ***	0.88 ***	0.95 **	0.09 ***	-0.01	0.07 ***	0.1 ***	0.05 **	0
h=2	0.87 ***	0.96	0.85 ***	1.02	0.81 ***	1.01	0.11 ***	-0.04	0.07 ***	0.07 ***	0.11 ***	-0.01
h=3	0.85 ***	0.98	0.84 ***	0.99 **	0.83 ***	1.03	0.15 ***	-0.03 *	0.08 ***	0.06 ***	0.13 ***	-0.04 *
h=4	0.84 ***	1.01	0.9 ***	0.95 ***	0.81 ***	1.04	0.18 ***	-0.03 ***	0.06 ***	0.07 ***	0.13 ***	-0.04
h=5	0.84 ***	1	0.91 ***	0.98 *	0.77 ***	1.05	0.21 **	-0.02	0.05 ***	0.08 ***	0.18 ***	-0.04
h=6	0.84 **	1.01	0.96 ***	1	0.77 ***	1.06	0.22 **	-0.02 ***	0.05 ***	0.07 ***	0.18 ***	-0.04
h=7	0.85 **	1.02	0.99	1.02	0.8 ***	1.07	0.19 **	-0.02 ***	0.04 ***	0.07 ***	0.16 **	-0.06
h=8	0.85 **	1	0.99	0.99	0.81 **	1.06	0.2 **	-0.01	0.04 ***	0.08 ***	0.14 **	-0.05

Note: Left block: ratio of power-prior to Minnesota-baseline RMSE by variable, horizon, and country (values below one favor the power prior). Right block: difference in average log predictive score (power prior minus Minnesota baseline; positive values favor the power prior). *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. For RMSE, significance is based on the one-sided Clark–West test of T. E. Clark and West (2007), using Newey–West standard errors with lag length $h - 1$. For log predictive score differences, significance is based on a two-sided test of equal predictive accuracy using HAC standard errors.

Table 8: Relative RMSE and log-score difference: power prior vs. diffuse baseline

	Relative RMSE						LPS difference					
	Y	Dp	DU	DR	CY	S	Y	Dp	DU	DR	CY	S
Euro Area												
h=1	0.73 **	0.71 ***	0.92	0.93	0.77 ***	0.86 *	0.26 ***	0.23 ***	-0.08	0.02	0.17 **	0.09 **
h=2	0.75 ***	0.82 **	0.85	0.77 ***	0.76 ***	0.77	0.27 ***	0.21 ***	0.01	0.15 ***	0.17 ***	0.18 ***
h=3	0.76	0.85 **	0.9	0.84	0.72	0.75 **	0.27	0.24 ***	-0.06	0.15 ***	0.2 *	0.2 ***
h=4	0.81	0.83 ***	0.9	0.83	0.73	0.76 **	0.24	0.27 ***	-0.03	0.17 ***	0.19 *	0.2 ***
h=5	0.84	0.96	0.98	0.83	0.75	0.74	0.2	0.28 ***	-0.1	0.22 ***	0.17 *	0.25 **
h=6	0.88	1	1.02	0.9	0.81	0.74	0.17	0.29 ***	-0.13 *	0.23 ***	0.12	0.26 *
h=7	0.92	0.97	0.95	0.9	0.9	0.76	0.14	0.31 ***	-0.08	0.24 ***	0.07	0.26 *
h=8	0.95	0.98	0.95	0.93	0.98	0.78	0.12	0.31 ***	-0.06	0.24 ***	0.02	0.25 **
Canada												
h=1	0.86 *	0.9 *	0.83	0.94	0.9	0.9 *	0.05	0.09 *	0.19 ***	0.1 ***	0.11 **	0.16 ***
h=2	0.82 **	0.89 *	0.87	0.89	0.93	0.93	0.14 **	0.09 ***	0.22 ***	0.15 ***	0.07	0.17 ***
h=3	0.86	0.88 **	0.93	0.85 **	0.89	0.87 **	0.14 **	0.11 ***	0.24 ***	0.26 ***	0.09	0.21 ***
h=4	0.88 *	0.91 ***	1.02	0.93	0.91	0.87 **	0.16 ***	0.12 ***	0.23 ***	0.28 ***	0.09	0.24 **
h=5	0.9	0.95 *	0.99	1	0.9	0.9 *	0.18 ***	0.13 ***	0.26 ***	0.28 ***	0.11 **	0.21 **
h=6	0.93	0.97 **	1.03	0.99	0.91	0.92	0.15 **	0.17 ***	0.27 ***	0.32 ***	0.11 **	0.2 **
h=7	0.95	0.97 *	0.99	0.99	0.89	0.91	0.14	0.16 ***	0.31 ***	0.35 ***	0.14 **	0.19 *
h=8	0.99	0.99	1	1	0.86	0.91	0.13	0.17 ***	0.33 ***	0.35 ***	0.16 **	0.2 *
Australia												
h=1	0.73 ***	0.64 ***	0.88	0.82 ***	0.81 **	0.78 **	0.29 ***	0.34 ***	0.17 ***	0.15 ***	0.17 ***	0.21 ***
h=2	0.75 **	0.62 ***	0.9	0.88	0.87	0.78 *	0.29 ***	0.39 ***	0.14 ***	0.12 *	0.14 **	0.22 **
h=3	0.85	0.77 ***	1.03	0.86 *	0.9	0.76	0.23 ***	0.36 ***	0.11 **	0.13 ***	0.15 ***	0.23
h=4	0.87	0.88 **	1.07	1.06	0.89	0.79	0.2 **	0.34	0.13 ***	0.09 **	0.13 *	0.2
h=5	0.94	0.9 *	1.04	1.04	0.85	0.81	0.16 **	0.36 **	0.18 ***	0.12 ***	0.17 *	0.2 *
h=6	1.02	0.94 **	0.91	1.16	0.79	0.77 **	0.13 *	0.4 ***	0.24 ***	0.11 **	0.22 **	0.26 **
h=7	1.07	0.99	0.84 **	1.07	0.75	0.79 *	0.13 *	0.44 ***	0.29 ***	0.16 ***	0.26 **	0.24 *
h=8	1.05	1	0.8 **	1.12	0.71	0.84	0.15 **	0.47 ***	0.32 ***	0.16 ***	0.31 ***	0.22 **
UK												
h=1	0.61 **	0.67 ***	1.14	0.77 **	0.89	0.85 ***	0.29 ***	0.3 ***	-0.11	0.17 ***	0.15 ***	0.17 ***
h=2	0.62 **	0.66 **	1.12	0.7	0.91	0.86 *	0.31 ***	0.34 ***	-0.05	0.27 ***	0.13 ***	0.23 ***
h=3	0.63	0.75 ***	1.08	0.7	0.9	0.86	0.32 ***	0.39 ***	-0.02	0.35 **	0.14 **	0.31 ***
h=4	0.64	0.87 **	1.04	0.75	0.83	0.87	0.33 ***	0.45 ***	0.01	0.33 **	0.17 **	0.33 ***
h=5	0.68	0.89 **	1.06	0.78	0.83	0.83	0.31 ***	0.49 ***	-0.04	0.3 *	0.17 **	0.35 ***
h=6	0.7	0.89 **	1.06	0.82	0.81 *	0.79	0.29 ***	0.51 ***	-0.06	0.26	0.2 **	0.36 ***
h=7	0.73	0.89 ***	1.04	0.85	0.79 **	0.78	0.25 ***	0.54 ***	-0.04	0.23	0.22 **	0.34 **
h=8	0.76	0.91 **	1.03	0.86	0.78 *	0.77	0.23 **	0.56 ***	-0.04	0.21	0.24 ***	0.32 ***
US												
h=1	0.77 **	0.8 **	0.65 ***	0.58 ***	0.71 **	0.77 ***	0.24 ***	0.14 *	0.28 ***	0.37 ***	0.24 ***	0.2 ***
h=2	0.76 *	0.9	0.67 ***	0.7 ***	0.62 **	0.94	0.26 ***	0.09 **	0.28 ***	0.34 ***	0.3 **	0.16 ***
h=3	0.75 *	0.94	0.68 ***	0.68 ***	0.61	0.99	0.3 **	0.11 **	0.28 ***	0.35 ***	0.35	0.14 **
h=4	0.77 **	1.03	0.76 ***	0.71 ***	0.63 **	0.99	0.27 **	0.15 ***	0.29 ***	0.37 ***	0.32 **	0.15 **
h=5	0.79 ***	0.94	0.78 **	0.79 **	0.66 **	1.04	0.26 ***	0.21 ***	0.29 ***	0.41 ***	0.29 ***	0.15 **
h=6	0.8 **	0.93 *	0.81 *	0.85 **	0.68 **	1.04	0.24 ***	0.26 ***	0.33 ***	0.43 ***	0.27 ***	0.14
h=7	0.8 *	0.98	0.84 **	0.9 *	0.7	1.05	0.25 ***	0.31 ***	0.36 ***	0.46 ***	0.26 **	0.14
h=8	0.79	0.91 **	0.82	0.84 **	0.71	1.04	0.26 ***	0.35 ***	0.38 ***	0.49 ***	0.25	0.14

Note: Left block: ratio of power-prior to diffuse-baseline RMSE by variable, horizon, and country (values below one favor the power prior). Right block: difference in average log predictive score (power prior minus diffuse baseline; positive values favor the power prior). *, **, and *** indicate rejection of equal predictive accuracy at the 10%, 5%, and 1% significance levels, respectively. Significance is based on two-sided tests of the mean loss or score differential using HAC standard errors.

Table 9: Relative RMSE and log-score difference: power prior vs. ACP baseline

	Relative RMSE						LPS difference					
	Y	Dp	DU	DR	CY	S	Y	Dp	DU	DR	CY	S
Euro Area												
h=1	0.94	1.04	1.21 ***	1.04	0.92	0.9 **	0.04	-0.05 *	-0.34 ***	-0.19 ***	0.04	-0.01
h=2	0.95	1.04	1.21 **	1.01	0.95	0.9 *	0.04	-0.09 ***	-0.33 ***	-0.15 ***	-0.05	-0.01
h=3	0.96	1.03	1.23	0.97	0.92	0.87 **	0.05	-0.08 ***	-0.36 ***	-0.15 ***	-0.03	-0.01
h=4	0.96	1.01	1.22	0.95	0.93	0.85 *	0.06	-0.08 ***	-0.3 ***	-0.14 ***	-0.04	-0.01
h=5	0.95	1.01	1.18	0.96	0.89	0.85 **	0.11	-0.08 ***	-0.27 ***	-0.14 ***	-0.01	0.02
h=6	0.96	1.01	1.15	0.98	0.9	0.86 **	0.12	-0.08 ***	-0.23 ***	-0.16 ***	-0.06	0.03
h=7	0.97	1	1.05	0.99	0.9	0.87	0.12	-0.06 ***	-0.15 **	-0.17 ***	-0.07	0.03
h=8	0.99	1	1	1.01	0.92	0.89	0.1	-0.07 ***	-0.08	-0.17 ***	-0.1	0.03
Canada												
h=1	1.02	1.1 *	0.93 *	1.14 *	0.94	0.98	-0.09 ***	-0.1	0.09 ***	-0.09 ***	0	0.01
h=2	1	1.05 **	0.91 *	1.05	0.96	0.98	-0.05	-0.11	0.08 **	-0.03	-0.03	0.03
h=3	0.98	1.01	0.94	1	0.9 ***	0.98	-0.02	-0.1	0.07 **	-0.01	0	0.03
h=4	0.96	1	0.97	0.97	0.9 ***	0.99	0.01	-0.09	0.05 *	0	0	0.03
h=5	0.95	0.99	0.98	1.03	0.88 ***	0.98	0.04	-0.08	0.03	-0.02	0	0.03
h=6	0.95	0.99	1.01	1.02	0.84 ***	0.96	0.05	-0.04	0.02	-0.02	0.04	0.07
h=7	0.97	1	1.02	0.99	0.84 ***	0.94	0.03	-0.02	0.02	-0.01	0.04	0.07
h=8	0.99	1	1	1.01	0.81 ***	0.94	0.02	-0.06	0.03	-0.02	0.06	0.11
Australia												
h=1	1.03	0.99	1.01	0.94	0.97	0.93 *	0.02	0.1 ***	-0.01	0.13	0.01	0.07 ***
h=2	0.97	1.04 *	1.02	1.03	0.92	0.92 *	0.04	0.07 **	-0.04	-0.03	0.01	0.11 ***
h=3	0.96	1.02	1.08 **	1.04	0.92	0.92 *	0.03	0.08 **	-0.07 **	-0.04	-0.01	0.14 ***
h=4	0.97	1	1.11 ***	1.09	0.93	0.92	-0.01	0.06	-0.08 ***	-0.04	-0.03	0.13 **
h=5	1.02	0.99	1.08 ***	1.08	0.93	0.91 *	-0.03	0.05	-0.06 ***	-0.07	-0.01	0.1 **
h=6	1.08	1	1.04	1.17	0.94	0.9 **	-0.08	0.06	-0.02	-0.12 **	0.01	0.12 **
h=7	1.11	0.99	1	1.21	0.96	0.92	-0.09	0.09 ***	0.02	-0.15 ***	0.02	0.08
h=8	1.11	0.98 *	0.97	1.31 *	0.96	0.94	-0.1	0.11 **	0.04 **	-0.19 ***	0.04	0.08 **
UK												
h=1	1.05	0.99	1.19 *	0.87	0.94	0.9 *	-0.02	0.06 ***	-0.11	0.03	0.05	0.04
h=2	1.08	1.06	1.22	0.78	0.97	0.89	-0.02	0.09 ***	-0.07	0.11	0.04	0.09
h=3	1.18	0.97	1.2	0.75	0.94	0.86	-0.06	0.11 ***	-0.08	0.18	0.06	0.13
h=4	1.19	1.04	1.17	0.75	0.91	0.87	-0.05	0.11 ***	-0.03	0.23	0.08	0.14
h=5	1.22	0.97	1.14	0.78	0.89	0.89	-0.08	0.12 ***	0.07	0.22	0.11	0.12
h=6	1.22	0.97	1.09	0.8	0.88	0.91	-0.1	0.11 ***	0.08	0.2	0.11	0.12
h=7	1.19	0.99	1.02	0.83	0.87	0.93	-0.1	0.11 ***	0.14	0.18	0.13	0.1
h=8	1.13	0.97	0.97	0.86	0.87	0.95	-0.08	0.12 ***	0.18	0.16	0.13	0.04
US												
h=1	0.97	0.96	0.86 ***	0.93	0.93 *	0.97	0.06 ***	0.01	0.07 ***	0.06 **	0.03	0
h=2	0.91 *	0.98	0.85 ***	1.1	0.87 *	1.04	0.1 ***	-0.03	0.06 ***	0.05 ***	0.06	-0.03
h=3	0.87 *	0.99	0.85 ***	1.03	0.82	1.07	0.14 ***	-0.04 ***	0.05 **	0.05 ***	0.1	-0.06
h=4	0.86	0.99	0.86 **	0.97	0.8	1.07	0.14 **	-0.02 **	0.07 **	0.05 ***	0.11	-0.06
h=5	0.86	1	0.9 **	1.03	0.76	1.08	0.16 **	-0.03 **	0.05 ***	0.06 ***	0.12	-0.06
h=6	0.86	1	0.94 ***	1.04	0.75	1.1	0.16 **	-0.02 ***	0.04 ***	0.05 **	0.13	-0.06 *
h=7	0.86	1.01 *	0.95 **	1.02	0.75	1.09	0.14 **	-0.02 **	0.04 ***	0.06 ***	0.13	-0.07
h=8	0.86	1	0.97	0.96	0.76	1.07	0.15 **	-0.01	0.05 ***	0.07 ***	0.11	-0.05

Note: Left block: ratio of power-prior to ACP-baseline RMSE by variable, horizon, and country (values below one favor the power prior). Right block: difference in average log predictive score (power prior minus ACP baseline; positive values favor the power prior). *, **, and *** indicate rejection of equal predictive accuracy at the 10%, 5%, and 1% significance levels, respectively. Significance is based on two-sided tests of the mean loss or score differential using HAC standard errors.

Table 10: Relative RMSE and log-score difference: power prior vs. panel VAR

	Relative RMSE						LPS difference					
	Y	Dp	DU	DR	CY	S	Y	Dp	DU	DR	CY	S
Euro Area												
h=1	0.8 **	1.02	1.24 **	0.99	0.93	1.02	0.27 *	-0.03	-0.34 ***	-0.13 **	0.02	-0.09 **
h=2	0.83 **	1.03	1.3 **	0.89	0.89	0.9	0.3 *	-0.07 ***	-0.42 ***	-0.09 *	0.03	0.01
h=3	0.87	1.11	1.17	0.88	0.85	0.79	0.25	-0.07 ***	-0.35 ***	-0.1 **	0.08	0.1
h=4	0.89	0.97	1.08	0.88	0.86	0.76	0.25	-0.03 *	-0.25 **	-0.1 **	0.04	0.11
h=5	0.91	0.99	1.01	0.92	0.83 *	0.77	0.28	-0.02	-0.15	-0.1 ***	0.05	0.12
h=6	0.87 *	1.05 *	0.93	0.94	0.84	0.81	0.31	-0.02	-0.09	-0.11 ***	0.01	0.11
h=7	0.89	0.97	0.85 *	0.95	0.86	0.86	0.29	0.02	-0.09	-0.13 ***	-0.02	0.08
h=8	0.91	0.99	0.87 *	0.95	0.88	0.9	0.27	0	-0.08	-0.14 ***	-0.05	0.06
Canada												
h=1	0.89	0.99	0.89	1.01	0.89 *	0.93	0.12	-0.02	0.14 ***	0.07	0.06	0.08 **
h=2	0.83 **	0.99	0.93	0.93	0.85	0.9	0.14	-0.05	0.12 *	0.05 *	0.08	0.1 *
h=3	0.87	0.97	1.03	0.89 **	0.74 *	0.86 **	0.09	-0.05	0.09	0.06 ***	0.13	0.13 **
h=4	0.93	0.99	0.93	0.89 ***	0.71 **	0.82 ***	0.06	-0.05	0.12 **	0.07 ***	0.18	0.21 ***
h=5	0.88	0.99	1.04	1.06	0.67 **	0.82 ***	0.07	-0.04	0.09 *	0.04 **	0.22	0.24 ***
h=6	0.92	1	1	0.85 ***	0.63 ***	0.8 ***	0.03	-0.01	0.11 ***	0.07 ***	0.26 **	0.26 ***
h=7	0.94	1	1.03	0.92	0.6 ***	0.8 ***	0.01	0	0.1 **	0.08 ***	0.29 ***	0.26 ***
h=8	0.94	1	0.94	0.98	0.56 ***	0.81 ***	0	0	0.14 ***	0.06 ***	0.33 ***	0.25 ***
Australia												
h=1	0.91	0.75 **	0.61 ***	0.92	0.9 **	0.87 **	0.12 ***	0.25 ***	0.12 **	0.15 ***	0.14 ***	0.19 ***
h=2	0.86	0.87	0.67 ***	1.1	0.91	0.89 *	0.13 **	0.16 ***	0.07	0.05	0.12 ***	0.14 *
h=3	0.85	0.89	0.89	1.01	0.9	0.89	0.14 *	0.16 ***	0.01	0.04	0.13 **	0.14
h=4	0.81	1	0.91	1.04	0.91	0.91	0.14 *	0.12	0.03	0.01	0.11	0.08
h=5	0.87	0.93 **	0.86	1.02	0.82	0.95	0.09	0.15 *	0.03	0.01	0.12	0.02
h=6	0.89	0.95 ***	0.87	0.94	0.87	0.99	0.05	0.15 *	0.05	-0.05	0.1	0.05
h=7	0.91	0.94 *	0.73 **	1.21	0.85	0.96	0.04	0.19 ***	0.08 *	-0.05	0.11 *	0.07
h=8	0.93	0.97	0.78 ***	1.32 *	0.83 *	0.93	0.05	0.2 ***	0.08 **	-0.07 *	0.09 *	0.11
UK												
h=1	0.79 ***	0.8	0.81 *	0.78 *	0.93	0.74 ***	0.18 ***	0.21 ***	-0.19 **	0.18 ***	0.16 ***	0.22 ***
h=2	0.72 ***	0.85 ***	0.87	0.72 *	1	0.7 ***	0.15 ***	0.23 ***	-0.15	0.22 ***	0.1 **	0.21 ***
h=3	0.79 ***	0.86 **	0.82	0.7	0.94	0.69 ***	0.1 **	0.26 ***	-0.1	0.25	0.13 **	0.18
h=4	0.81 **	0.85 ***	0.81	0.71	0.9	0.69 **	0.07	0.26 ***	-0.05	0.25	0.15 **	0.14
h=5	0.83 **	0.93 *	0.84	0.72	0.9	0.68 **	0.08	0.28 ***	-0.03	0.26	0.13 *	0.11
h=6	0.85 **	1.06	0.81 *	0.75	0.9	0.71	0.08	0.28 ***	0.01	0.26	0.14 *	0.11
h=7	0.87 *	0.89 **	0.78 **	0.76	0.87	0.74	0.09	0.29 ***	0.06	0.28	0.15 **	0.09
h=8	0.84 **	0.86 **	0.75 **	0.79	0.86 *	0.77	0.11	0.3 ***	0.05	0.27	0.18 **	0.07
US												
h=1	0.86 **	0.87 *	0.8 ***	0.89	0.96	0.88	0.29 ***	0.18 ***	0.29 ***	0.28 ***	0.28 ***	0.22 ***
h=2	0.81 *	0.92	0.71 ***	0.89	0.87	0.96	0.22 ***	0.09 ***	0.23 ***	0.26 ***	0.27 ***	0.09
h=3	0.85	0.96	0.75 ***	0.76 **	0.86	1	0.17 *	0.1 ***	0.21 ***	0.27 ***	0.22 ***	0.06
h=4	0.88	0.9 ***	0.82 **	0.88	0.84	0.96	0.12	0.12 ***	0.21 ***	0.25 ***	0.19 ***	0.08
h=5	0.89 *	1.01	0.8 **	0.88	0.82	0.93	0.11	0.1 ***	0.21 ***	0.26 ***	0.19 ***	0.07
h=6	0.89	1.04 *	0.75 ***	0.82 *	0.79 **	0.96	0.09	0.11 ***	0.21 ***	0.26 ***	0.2 ***	0.04
h=7	0.91	0.91 ***	0.84 ***	0.76 ***	0.79 **	0.99	0.08	0.14 ***	0.2 ***	0.28 ***	0.2 ***	0
h=8	0.9	0.95 *	0.82 ***	0.82 **	0.78 **	1.01	0.09	0.15 ***	0.21 ***	0.28 ***	0.2 ***	-0.01

Note: Left block: ratio of power-prior to panel VAR RMSE by variable, horizon, and country (values below one favor the power prior). Right block: difference in average log predictive score (power prior minus panel VAR; positive values favor the power prior). *, **, and *** indicate rejection of equal predictive accuracy at the 10%, 5%, and 1% significance levels, respectively. Significance is based on two-sided tests of the mean loss or score differential using HAC standard errors.

Table 11: Relative RMSE and log-score difference: power prior vs. hierarchical VAR of Prüser and Blagov (2026)

	Relative RMSE						LPS difference					
	Y	Dp	DU	DR	CY	S	Y	Dp	DU	DR	CY	S
Euro Area												
h=1	0.93	1.19 **	1.22 ***	1.06	1	1.04	0.05	-0.08 ***	-0.33 ***	-0.15 ***	-0.06	-0.1 ***
h=2	0.96	1.09 **	1.18 *	0.98	1.04	0.98	0.02	-0.07 ***	-0.31 ***	-0.15 ***	-0.1 **	-0.04
h=3	1	1.06 *	1.15	0.98	1.04	0.88	0.01	-0.06 ***	-0.27 ***	-0.18 ***	-0.12 **	0.01
h=4	1.02	1.01	1.1	0.96	1.08	0.84	0.02	-0.05 ***	-0.21 **	-0.18 ***	-0.15 **	0.03
h=5	1.03	1	1.03	0.96	1.06	0.83	0.04	-0.06 ***	-0.12	-0.18 ***	-0.15 **	0.04
h=6	1.04	1.02 *	0.98	0.94	1.08	0.86 *	0.05	-0.06 ***	-0.04	-0.19 ***	-0.18 **	0.03
h=7	1.04	1.01	0.92	0.95	1.11	0.89	0.03	-0.04 **	0.02	-0.2 ***	-0.21 ***	0.02
h=8	1.05	1.01	0.91	0.97	1.13	0.91	0.03	-0.06 ***	0.03	-0.2 ***	-0.22 **	0.01
Canada												
h=1	0.96	1.12 ***	0.96	1.08	0.97	0.99	-0.05	-0.12 ***	0.07 **	0.01	0	0.04 *
h=2	0.94	1.06 ***	0.96	1.06	0.98	0.97	-0.02	-0.1 *	0.06 *	-0.01	0	0.06
h=3	0.95	1.02	0.97	1	0.93	0.98	-0.01	-0.11 **	0.07 *	-0.02	0.01	0.05 *
h=4	0.97	1	0.97	0.98	0.93	0.95	0.01	-0.09	0.07	-0.02 *	0	0.07 *
h=5	0.98	1	0.98	0.98	0.9	0.94	0.03	-0.08	0.06	-0.03 *	0.03	0.08 **
h=6	0.99	0.99	0.99	0.98	0.89 *	0.95	0.01	-0.03	0.05	-0.01	0.05	0.06
h=7	0.99	0.99	1.01	0.95	0.89 **	0.95	0	-0.02	0.04	-0.02 *	0.06	0.06 ***
h=8	1.01	0.99	0.97	0.97	0.87 **	0.96	-0.01	-0.04	0.04	-0.02 *	0.08	0.05 **
Australia												
h=1	1.01	1.07	0.98	0.93	0.97	0.95	0.08 ***	0.09 ***	0.05 *	0.11 **	0.04 *	0.08
h=2	0.96	1.11 ***	0.98	1.03	0.87 **	0.94	0.07 *	0.08 **	0.03	0	0.07 **	0.08
h=3	0.95	1.03 *	1	1.06	0.83 **	0.92	0.05	0.07 *	0.03	-0.04	0.09 *	0.09
h=4	0.97	1	1	1.1	0.84 **	0.9	0.02	0.06	0.01	-0.08	0.08	0.13
h=5	1.01	1	1	1.06	0.81 **	0.89	-0.01	0.06	0.02	-0.06	0.11	0.11
h=6	1.06	0.99	0.99	1.15	0.83 *	0.9	-0.03	0.06	0.01	-0.12 ***	0.11	0.1
h=7	1.08	0.98 *	0.98	1.17	0.84 *	0.92	-0.05	0.09 **	0.02	-0.16 ***	0.12	0.08
h=8	1.08	1	0.98	1.27 *	0.85 *	0.93	-0.03	0.09 ***	0	-0.22 ***	0.12	0.08
UK												
h=1	0.99	1.09 **	1.24 **	1.04	0.96	1.03	0.05 ***	0.14 ***	-0.21 **	-0.01	0.05	0.04 *
h=2	0.99	1.07 **	1.24	0.95	0.94	0.98	0.03	0.13 ***	-0.16	0.04	0.06	0.06 *
h=3	1.09	1	1.23	0.93	0.9	0.96	-0.04	0.12 ***	-0.1	0.06	0.12	0.06
h=4	1.11	1.02	1.2	0.92	0.84	0.96	-0.06	0.11 ***	-0.06	0.08	0.18	0.07
h=5	1.14	0.98	1.17	0.92	0.83	0.96	-0.09	0.12 ***	-0.01	0.08	0.21	0.05
h=6	1.15	0.99	1.12	0.91	0.82	0.96	-0.11	0.12 ***	0.07	0.08	0.24	0.05
h=7	1.14	0.99	1.06	0.91	0.82	0.95	-0.12	0.11 ***	0.12	0.09	0.27	0.03
h=8	1.09	0.98	1.02	0.92	0.81	0.96	-0.1	0.12 ***	0.13	0.09	0.32 *	0.02
US												
h=1	0.93	0.98	0.97	1.43 **	0.94	1.23	0.13 ***	0.04	0.11 ***	0.11 ***	0.06 **	-0.03
h=2	0.89 *	1	0.96	1.34 **	0.87	1.25	0.11 ***	-0.04	0.07 ***	0.07 ***	0.07	-0.09
h=3	0.91	0.98	0.96	1.15	0.86	1.18	0.09	-0.03 *	0.04 **	0.06 ***	0.07	-0.09
h=4	0.92	0.97	0.98	1.12	0.84	1.14	0.07	-0.03 **	0.04 *	0.06 ***	0.04	-0.07
h=5	0.94	1	1	1.13	0.84	1.14	0.08	-0.05 ***	0.02	0.05 ***	0.04	-0.08
h=6	0.95	0.99	1.01	1.07	0.82	1.16	0.07	-0.04 ***	0.01	0.05 ***	0.04	-0.1
h=7	0.96	1	1.02	1.01	0.82	1.17	0.06	-0.04 ***	0	0.06 ***	0.03	-0.11
h=8	0.96	1	0.99	0.97	0.82	1.16	0.07	-0.06 ***	0	0.07 ***	0.03	-0.11

Note: Left block: ratio of power-prior to hierarchical VAR RMSE by variable, horizon, and country (values below one favor the power prior). Right block: difference in average log predictive score (power prior minus hierarchical VAR; positive values favor the power prior). *, **, and *** indicate rejection of equal predictive accuracy at the 10%, 5%, and 1% significance levels, respectively. Significance is based on two-sided tests of the mean loss or score differential using HAC standard errors.

B.3 Global VAR Application

This subsection presents an additional forecasting exercise based on a broader country sample and a different evaluation period. Its purpose is to assess whether the favorable forecasting performance of the power prior documented in the main application extends to a larger multicountry setting. We compare the power-prior specification with the hierarchical VAR of Prüser and Blagov (2026), which provides a particularly relevant benchmark because it also exploits cross-country information, but does so by estimating the countries jointly under a hierarchical pooling structure. The comparison therefore illustrates whether targeted

focal-country borrowing can remain competitive with a fully joint multicountry approach.

The exercise is based on the global VAR dataset of Mohaddes and Raissi (2024). We use quarterly data on five macro-financial variables and evaluate forecasts over the period 2000Q1–2017Q4, thereby excluding the Covid period. Each country VAR contains seasonally adjusted log real GDP, CPI inflation, the term spread measured as the long-term minus the short-term interest rate, the log real equity price index, and the log real exchange rate. Although the original dataset covers 33 countries, only 18 have complete observations for all five variables over the relevant sample period.⁸ We estimate quarterly VARs with four lags throughout.

For each focal country, the remaining 17 countries are treated as potential donors. The power-prior specification uses the asymmetric conjugate prior of Chan (2022) and allows the discount factors to vary across donor countries and VAR equations. The international cross section therefore serves as a set of candidate information sources, while estimation and forecasting remain targeted at one focal country at a time. By contrast, the hierarchical specification of Prüser and Blagov (2026) estimates the cross-country system jointly and transmits information through its common pooling structure.

Tables 12 and 13 report the detailed pseudo-out-of-sample results by country, variable, and forecast horizon. Table 12 reports ratios of the power-prior RMSE to the RMSE of the hierarchical VAR. Values below one indicate more accurate point forecasts under the power prior. Table 13 reports differences in average log predictive scores, calculated as the power-prior score minus the hierarchical-VAR score. Positive entries therefore indicate better density forecasts under the power prior.

⁸Austria, Australia, Belgium, Canada, France, Germany, Italy, Japan, South Korea, the Netherlands, Norway, New Zealand, South Africa, Spain, Sweden, Switzerland, the United Kingdom, and the United States.

Table 12: Relative RMSE: power prior vs. hierarchical VAR of Prüser and Blagov (2026)

	austria	austria	bel	can	france	germ	italy	japan	kor	neth	nor	nzld	safrz	spain	swe	switz	uk	usa
GDP																		
h=1	0.69 **	0.82	0.87	0.6 ***	0.92	0.89	1.21	0.99	0.65	0.77 *	0.78 ***	0.62 ***	0.7 **	1.05	0.93	0.83 **	0.88	0.74 *
h=2	0.5 **	0.73	0.74	0.49 **	0.93	0.79	1.25	0.84	0.46	0.77	0.57 ***	0.47 ***	0.54 **	1.07	0.84	0.82 *	0.75	0.59 *
h=3	0.38 **	0.59 **	0.6	0.39 **	0.88	0.71	1.26	0.72	0.31 *	0.79	0.41 ***	0.37 ***	0.43 *	1.06	0.7	0.76	0.63	0.48 *
h=4	0.27 ***	0.54 ***	0.5 *	0.32 **	0.8	0.63	1.24	0.64	0.21 *	0.79	0.34 ***	0.3 ***	0.34 *	1.01	0.56	0.71	0.53 *	0.41 *
h=5	0.19 ***	0.48 ***	0.42 *	0.26 ***	0.71	0.55	1.2	0.56	0.16 **	0.74	0.27 ***	0.25 ***	0.29	0.94	0.49	0.66	0.44 **	0.34 **
h=6	0.15 ***	0.42 **	0.35 **	0.22 ***	0.63	0.49	1.12	0.51	0.12 **	0.7	0.23 ***	0.21 ***	0.25	0.87	0.41	0.62	0.37 ***	0.3 **
h=7	0.11 ***	0.36 **	0.31 **	0.18 ***	0.58	0.45	1.06	0.45 *	0.09 **	0.65	0.21 ***	0.17 ***	0.23	0.81	0.36	0.57	0.31 ***	0.25 **
h=8	0.09 ***	0.33 ***	0.26 **	0.15 ***	0.52	0.4 *	1.02	0.4 *	0.06 ***	0.63	0.19 ***	0.15 ***	0.21	0.75	0.32 *	0.53	0.27 **	0.22 **
Inflation																		
h=1	1.16	1.14	1.08 **	1.12	1.15 *	1.26	1.01	0.95	0.93	1.08	1	1.74 **	0.92	1	1.01	1	0.94	1.02
h=2	1.28 *	1.11	1.06	1.1	1.17	1.23	1.01	0.9	0.83 **	1.04	1.01	1.59 *	0.91	1.02	0.93	1.01	0.97	1
h=3	1.36	1.1	1.11 **	1.18	1.15	1.37 **	0.91	0.92	0.88	1.13	1.03	1.66 *	0.96	1.02	0.92	0.97	0.9	1.05
h=4	1.68	1.19	1.11	1.17	1.13	1.29 *	0.79	0.96	0.82 ***	1.2 **	1.07	1.72	1	1	0.86	0.9	0.85	1.02
h=5	1.89 *	1.24	1.23 **	1.2	1.25 *	1.34 **	0.76 **	0.93	0.84 **	1.23 **	1.06	1.95	1.01	0.94	0.84	0.83	0.83	0.99
h=6	2.04 *	1.17	1.27 **	1.19	1.22	1.23 *	0.68 **	0.9	0.83 **	1.17	1.1	2.04 **	1.03	0.91	0.84	0.82	0.79	0.94
h=7	2.25 *	1.19	1.3 * *	1.1	1.25	1.36	0.62 ***	0.95	0.79 ***	1.28 *	1.12	1.74 *	1.11	0.84	0.88	0.75 *	0.8	0.94
h=8	2.33 *	1.16	1.36 **	1.13	1.31	1.38	0.56 ***	0.95	0.76 ***	1.24 *	1.21	1.72 *	1.13	0.82	0.77	0.77	0.83	0.91
Equity Price																		
h=1	1.4 *	1	1.18 *	1.25 **	1.16 *	1.11	1.15	1.07	1.03	1.13	1.09	1.1	1.01	1.14 *	1.04	0.95	1.36 **	1.15
h=2	1.5	1	1.19 **	1.29	1.37	1.16	1.24	1.1 *	0.98	1.22	1	1.28	1.01	1.25 *	0.97	1.01	1.61 **	1.3
h=3	1.68	1.03	1.2 *	1.42	1.46	1.19	1.3	1.09	0.92	1.3	0.99	1.4	0.98	1.39 **	0.9	1.02	1.69 **	1.44
h=4	1.78	1.09	1.26	1.52	1.71 **	1.3	1.44	1.06	0.81	1.42	0.98	1.5	0.93	1.55 **	0.85	1.05	1.78 *	1.51
h=5	1.86	1.13	1.3	1.51	1.89 **	1.39	1.54	1.06	0.77	1.56 *	0.97	1.62	0.89	1.65 **	0.81	1.08	1.86 *	1.53
h=6	1.94	1.14	1.3	1.48	2.12 **	1.49	1.63	1.08	0.7	1.66 *	0.96	1.71	0.85	1.81 ***	0.75	1.11	1.93	1.52
h=7	1.93	1.15	1.33	1.46	2.27 *	1.57	1.7	1.07	0.55 *	1.78 **	0.93	1.81 *	0.8	1.94 **	0.7	1.08	2.05 *	1.5
h=8	1.93	1.15	1.37	1.38	2.4	1.66	1.78	1.09	0.43 **	1.88 **	0.89	1.87 *	0.76	2.12 **	0.67	1.05	2.19 *	1.45
Exchange Rate																		
h=1	1.07	1.31 **	1.21	1.01	1.18	1.28 *	1.21 **	1.13	0.99	1.03	1	1.01	1.22 **	1.16	1.19 *	1.16 *	1.24	1.01
h=2	1.04	1.35	1.2	0.98	1.22	1.33	1.3 *	1.15	0.91	1.01	0.94	0.97	1.28 *	1.18	1.23 **	1.2	1.34	1.02
h=3	0.97	1.32	1.2	0.95	1.22	1.23	1.29	1.2	0.88	1.01	0.91	0.96	1.34	1.16	1.24 **	1.2	1.35	1.04
h=4	0.91	1.29	1.09	0.92	1.21	1.12	1.24	1.21	0.86	0.99	0.89	0.94	1.39	1.14	1.32 ***	1.17	1.31	1.07
h=5	0.86	1.26	0.93	0.89	1.18	1.05	1.2	1.15	0.85	0.96	0.88	0.93	1.46	1.12	1.41 ***	1.17	1.26	1.1
h=6	0.8	1.27	0.79 *	0.86	1.19	1	1.18	1.14	0.84	0.9	0.87	0.92	1.55	1.13	1.54 ***	1.15	1.23	1.12
h=7	0.78	1.27	0.73 ***	0.8	1.22	1	1.15	1.19	0.8	0.89	0.85	0.91	1.62 *	1.19	1.6 ***	1.14	1.19	1.13
h=8	0.75	1.27	0.69 ***	0.75	1.21	0.99	1.13	1.14	0.75	0.87	0.84	0.89	1.67 *	1.24	1.64 ***	1.08	1.14	1.14
Spread																		
h=1	1.28	1.12	0.93	0.97	1.09	1.05	0.93	2.47	1.46 **	1.01	1.03	1.25	1.07	0.86	1.13	1.05	1.06	1.01
h=2	1.12	1.07	0.84	0.89	0.88	1.04	0.7 *	2.46	1.29	1.01	0.77 *	1.11	1.13	0.76	0.86	0.98	0.98	0.97
h=3	1.22	1	0.82	0.86	0.79	1.14	0.66 **	2.45	1.19	1.01	0.72 **	1.06	1.27	0.78	0.83	0.99	0.91	1.07
h=4	1.25 *	1	0.8	0.95	0.79	1.21	0.64 *	2.5	1.28	1.06	0.7 **	0.99	1.45	0.77	0.87	0.92	0.93	1.16
h=5	1.36 ***	1.02	0.79	1.04	0.81	1.27	1.02 **	2.53	1.45	1.12	0.7 **	1.03	1.57	0.77	0.97	0.87	0.93	1.25
h=6	1.48 ***	1.03	1.02	0.83	1.28	1.54	0.66 *	2.54	1.44	1.17	0.69 ***	1.12	1.62	0.74	1.02	0.84	0.98	1.35
h=7	1.56 ***	1.06	0.84	1.17	0.85	1.19	0.68	2.49	1.5	1.15	0.72 **	1.24	1.68	0.73	1.07	0.87	0.99	1.41
h=8	1.63 ***	1.09	0.86	1.2	0.84	1.11	0.68 *	2.42 *	1.51	1.16	0.76 **	1.31	1.68	0.69	1.11	0.84	0.99	1.42

Note: Ratio of power-prior to hierarchical VAR RMSE by variable, horizon, and country. The VAR includes log real GDP, CPI inflation, the term spread, the log real equity price index, and the log real exchange rate. *, **, and *** indicate rejection of the null of equal forecast accuracy at the 10%, 5%, and 1% significance levels, respectively, based on the Diebold–Mariano test of Diebold and Mariano (2002).

Table 13: Log-score difference: power prior vs. hierarchical VAR of Prüser and Blagov (2026)

	austria	austria	bel	can	france	germ	italy	japan	kor	neth	nor	nzld	safr	spain	swe	switz	uk	usa
GDP																		
h=1	0.16 ***	0.12	0.11 *	0.24 ***	0.02	0.06	-0.04	-0.12	0.19 ***	0.14 ***	0.11 *	0.21 ***	0.14 ***	-0.02	0.04	0.09 **	-0.05	0.11
h=2	0.31 ***	0.16	0.12	0.34 *	-0.02	0.17	-0.12	0.08	0.33 ***	0.15	0.29 ***	0.33 ***	0.24 **	-0.14	0.09	0.1	0.06	0.2
h=3	0.45 ***	0.26	0.19	0.46 **	-0.02	0.18	-0.12	0.03	0.51 ***	0.13	0.42 ***	0.45 ***	0.36 **	-0.1	0.11	0.09	0.2	0.32
h=4	0.62 ***	0.24	0.32	0.58 **	0	0.16	-0.19	0.09	0.68 ***	0.12	0.54 ***	0.58 ***	0.48 **	-0.18	0.12	0.14	0.17	0.45
h=5	0.79 ***	0.29	0.39	0.72 ***	0.06	0.19	-0.29	0.09	0.8 ***	0.07	0.64 ***	0.65 ***	0.56 **	-0.23	0.16	0.21	0.3	0.53
h=6	0.9 ***	0.31	0.48 *	0.82 ***	0.13	0.3	-0.35	0.15	0.9 ***	0.09	0.72 ***	0.75 ***	0.65 **	-0.23	0.3	0.25	0.29	0.63
h=7	1.01 ***	0.38	0.56 **	0.93 ***	0.21	0.33	-0.36	0.25	1.01 ***	0.09	0.78 ***	0.84 ***	0.72 ***	-0.16	0.41	0.3	0.46	0.74 *
h=8	1.12 ***	0.44	0.65 ***	1.03 ***	0.29	0.42	-0.39	0.28	1.12 ***	0.08	0.85 ***	0.93 ***	0.78 **	-0.05	0.59	0.36 *	0.53	0.83 *
Inflation																		
h=1	-0.1	-0.03	-0.07	-0.09 **	-0.04	-0.04	0.05 *	0.02	0.14 ***	-0.05	-0.02	-0.02	0.06	-0.06	0.03	0.02	0.03	-0.01
h=2	-0.11 **	-0.04	-0.03	-0.04	-0.05	-0.02	0.01	0.09	0.17 ***	-0.04	0.07	-0.02	0.07	-0.05	0.04	0.03	0	-0.05
h=3	-0.05	-0.05	-0.11 **	-0.05	-0.06	-0.06	0.04	0.04	0.17 ***	-0.06	0	-0.03	0.03	0	0.06	0.03	0.02	-0.05
h=4	-0.11	-0.08	-0.08	-0.04	-0.07	-0.05	0.09	0.1	0.2 **	-0.08 **	-0.09	-0.03	0	0.02	0.09	0.05	0.03	-0.02
h=5	-0.13	-0.1	-0.13 **	-0.04	-0.09 *	-0.05	0.12 **	0.07	0.2 ***	-0.1 **	-0.03	-0.05	0.01	0.04	0.09	0.1 **	0.02	0.03
h=6	-0.14	-0.09	-0.13 **	-0.04	-0.08	-0.03	0.16 **	0.09	0.21 ***	-0.09 **	-0.06	-0.08	-0.01	0.06	0.08	0.08 **	0.03	0.03
h=7	-0.17	-0.08	-0.14 *	0.01	-0.08	-0.06	0.18 ***	0.04	0.23 ***	-0.11 ***	0	-0.04	-0.06	0.09	0.07	0.11 ***	0.03	0.04
h=8	-0.16	-0.07	-0.15 **	0.03	-0.09	-0.06	0.23 ***	0.09 **	0.23 ***	-0.11 **	0.01	-0.04	-0.08	0.1	0.09	0.11 **	0.03	0.07
Equity Price																		
h=1	-0.1	0.03	-0.07 *	-0.13 **	-0.05	-0.05	-0.03	-0.04 *	-0.01	-0.08	0	-0.05 ***	0.02	-0.02	-0.01	0.01	-0.13 ***	-0.06
h=2	-0.19	0.04	-0.14	-0.13	-0.16 *	-0.1	-0.09	-0.1 **	-0.02	-0.11	0.05	-0.1 **	0.04	-0.07	0.04	0	-0.26 **	-0.23
h=3	-0.24	0	-0.15	-0.19	-0.23	-0.14	-0.11	-0.04	0.06	-0.29	0.03	-0.15 **	0.05	-0.13 *	0.07	-0.01	-0.27 *	-0.27
h=4	-0.28	-0.11	-0.23	-0.23	-0.33 **	-0.19	-0.19	-0.03	0.12	-0.3 *	-0.01	-0.18 ***	0.08	-0.18 **	0.1	0	-0.32	-0.32
h=5	-0.34 *	-0.1	-0.21	-0.23	-0.3 * *	-0.2	-0.25	-0.04	0.13	-0.45 *	0.06	-0.23 ***	0.09	-0.21 **	0.14	-0.01	-0.34 *	-0.31
h=6	-0.4	-0.08	-0.22	-0.18	-0.36 ***	-0.21	-0.29	-0.04	0.17 **	-0.46 **	0.04	-0.26 **	0.11	-0.25 ***	0.18	-0.01	-0.34 *	-0.33
h=7	-0.43 *	-0.17	-0.2	-0.36 **	-0.19	-0.32	-0.29	-0.02	0.23 ***	-0.5 *	0.06	-0.29 ***	0.13	-0.27 ***	0.21	-0.39 *	-0.39	-0.33
h=8	-0.43	-0.11	-0.31	-0.18	-0.39 *	-0.19	-0.31	-0.04	0.28 ***	-0.46 **	0.11	-0.32 ***	0.16	-0.29 **	0.25	0.04	-0.42 *	-0.3
Exchange Rate																		
h=1	0.02	-0.1 **	-0.06	0.02	-0.06	-0.08	-0.06 *	-0.02	0.03	0	0.02	-0.02	-0.14 **	-0.06	-0.11 **	-0.04	-0.1	-0.01
h=2	-0.01	-0.16	-0.03	0.24	-0.08	-0.09	-0.12 *	-0.02	0.05	0	0.03	-0.05	-0.22 *	-0.07	-0.16 **	-0.03	-0.04	-0.06
h=3	-0.03	-0.14	-0.03	0.06	-0.17	-0.06	-0.11	-0.06	0.01	0.05	0.02	-0.03	-0.16 *	-0.03	-0.19	0.01	-0.08	-0.19
h=4	0.01	-0.16	0.03	0.1	-0.04	-0.01	-0.09	-0.01	0.1	0.04	0.04	0	-0.35 *	-0.04	-0.33 ***	-0.02	-0.15	-0.16
h=5	0.09	-0.1	0.07	0.12	-0.02	0.02	-0.05	-0.01	0.08	0.06	0.05	0	-0.34	0	-0.19 *	0	-0.15	-0.02
h=6	0.14	-0.08	0.12 ***	0.24	0.01	0.04	-0.02	0.01	0.09	0.09	0.07	0.03	-0.43 *	0.01	-0.23 **	0.01	-0.13	0.03
h=7	0.15	-0.08	0.27	0.22	0.06	0.2	0.11	0.08	0.14 **	0.08	0.05	-0.41	0.03	-0.25 **	0.01	0.03	-0.13	0.02
h=8	0.19	-0.07	0.17 ***	0.39	0.03	0.07	0.01	0.01	0.14	0.13 **	0.07	0.07	-0.44 **	0.04	-0.26 *	0.04	-0.1	0.02
Spread																		
h=1	0.02	-0.03	0.04 **	0.2 **	0.04 **	-0.01	0	0.1	0.1 **	0.02	0.04 **	0.06 ***	0.03	0.12 ***	0.05 ***	0.04	0.04 **	0.05
h=2	0.05	-0.03	0.06 **	0.17 ***	0.06 **	0.01	0.16	0.3	0.16 ***	0.02	0.12	0.08	0.01	0.16 ***	0.12 ***	0.06 **	0	0.06
h=3	0.02	0.07	0.15 ***	0.09 **	0.04 **	0.03	0.02	0.3	0.18 ***	0.02	0.07 **	0.05	0.01	0.17 **	0.04 **	0.15 ***	0.03	0.05
h=4	0	-0.02	0.06	0.09 *	0.1 *	-0.06	0.37	-0.12	0.15 *	-0.04	0.16 ***	0.08	-0.07	0.14	0.1 ***	0.06 **	0.03	-0.06
h=5	-0.03	-0.04	0.07	0.06	0.09 **	-0.09	0.38	-0.17	0.13 *	-0.07	0.16 ***	0.03	-0.09	0.12	0.08 **	0.04 *	0.03	-0.12
h=6	-0.04	-0.04	0.06	0.1	0.1 *	-0.17	0.2	-0.12	0.12	-0.09	0.17 ***	0.01	-0.14	0.11	0.07	0.02	0.02	-0.2
h=7	-0.08 ***	-0.09	0.04	-0.01	0.05	-0.1 *	0.44	-0.21	0.11	-0.12 ***	0.15 ***	-0.06	-0.1	0.13 *	0.05	0.03	0.02	-0.24
h=8	-0.09 ***	-0.08	0.03	-0.01	0.05	-0.11 *	0.43	-0.24	0.1	-0.14 ***	0.14 ***	-0.1	-0.12	0.14 **	0.03	0.03	0.01	-0.25

countries and horizons, while density-forecast gains for the term spread are also common. Relative performance is mixed for inflation and exchange rates, and the hierarchical VAR generally performs better for equity prices. The exercise therefore supports the claim that targeted focal-country borrowing can remain competitive in a broader dataset, rather than a claim of uniform dominance over joint hierarchical pooling.