# Learning from crises: Forecasting with time-varying parameter VARs with observable adaptation

Nicolás Hardy[1,2] Dimitris Korobilis[1,3]

[1]Adam Smith Business School, University of Glasgow

[2]Universidad Diego Portales

[3]CAMP, BI Norwegian Business School

European Central Bank, 2026

# Introduction and Motivation

# Why TVP-VARs promise more than they deliver

- **TVP-VARs** are extremely useful generalizations of VARs:
  - Any nonlinear model can be approximated by a TVP linear model (Granger, 2008).
  - Allow for capturing changing economic environments (Primiceri, 2005; Cogley & Sargent, 2005).
- But they are **not as successful at forecasting** as simpler alternatives: univariate UC-SV (Stock & Watson, 2007), constant-parameter VARs with SV, SVO, or Student-$t$ errors.
- The reason is simple. A TVP-VAR is a state-space model:

$$y_t = \beta_t x_t + \varepsilon_t, \qquad \beta_t = \beta_{t-1} + \eta_t$$

  but the state vector can be very large because $x_t$ (lags of all endogenous variables) is also large.
- Stacking over all $T$ observations gives a regression with $T$ observations and $Tk$ parameters:
  - A three-variable VAR(2): $k = 21 \implies$ moderate, but already challenging.
  - A ten-variable VAR(12): $k = 1{,}210 \implies$ massively overparameterized.
- So TVP-VARs **rely heavily on prior hyperparameter choices**, which may go very wrong even when these choices are data-based and not fully subjective.

# Tight priors flatten TVPs — subsample OLS detects more variation

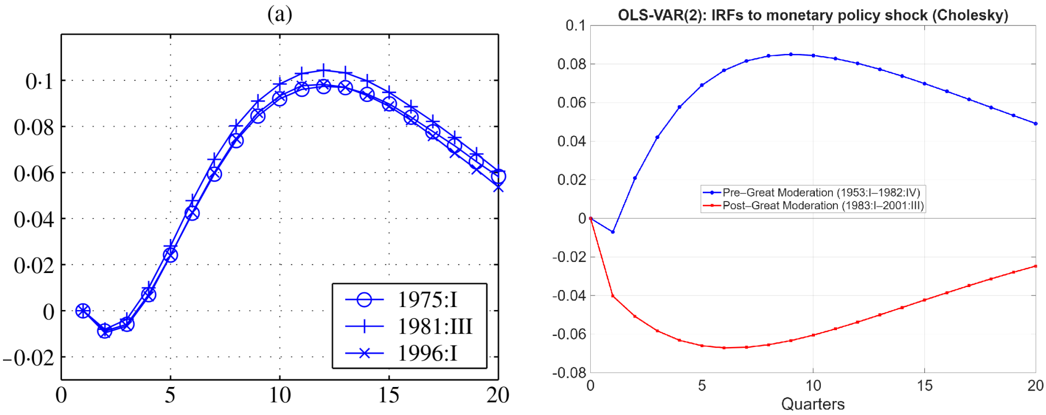



Figure: IRF of unemployment to a monetary policy shock. Left: TVP-VAR (Primiceri, 2005). Right: OLS-VAR(2) estimated on pre- and post-Great Moderation subsamples using the same data. Primiceri's "data-estimated" prior is way too tight killing any benefits from specifying TVPs!

# Beyond the random walk: why we need external information

- TVPs may also forecast poorly not just because of prior choice, but because of the **assumed evolution**. The random walk is persistent and smooth — poorly suited to **abrupt shocks**.
- Consider COVID-19: at monthly frequency, the disruption lasts roughly **three observations**; at quarterly frequency, **a single observation**. Yet its depth is unprecedented.
  - How can a persistent, gradual RW evolution capture such a sharp, transient break?
- Our best bet: **bring external data information** into the parameter evolution.
  - Uncertainty indices, financial conditions, consumer expectations — these series move *with* abrupt breaks.
  - They have valuable information about crises such as COVID, the GFC, and various geopolitical episodes.
- **Key insight:** Rather than asking the random walk to discover breaks from the likelihood alone, we can *tell* the parameters where to look via observable conditioning variables $Z_t$.

# Observable adaptation: an old idea with a new implementation

> *"An obvious question that will arise is 'why not consider TVPs with nonlinear models?' In fact the up-dating situation just considered would generate such models. These models would be particularly difficult to use, for forecasting for example, and to interpret."*
>
> — Clive Granger (2008), *Studies in Nonlinear Dynamics & Econometrics*

- The idea is old: define $\beta_t = f(z_t)$ where $f$ is a nonlinear function and $z_t$ is external conditioning information. These are **functional coefficient** time series models (Cai, Fan & Yao, 2000, *JASA*).
- An emerging literature approximates complex $f(\cdot)$ using ML methods: regression trees (Hauzenberger et al., 2022; Goulet Coulombe, 2024), neural networks (Gu et al., 2021), and Gaussian processes (Fox & Dunson, 2015).
- In this paper we argue that a **linear, deterministic** $f(\cdot)$ is the best choice, **conditional on purging the random walk variation**:

$$\boxed{\beta_t = \beta_{t-1} + \gamma Z_{t-1}}$$

- Equivalently: $\beta_t = \beta + \gamma C_t$, $C_t = \sum_{s=1}^{t-1} Z_s$, inherits RW persistence while being fully driven by observables.

# What are the benefits of this approach?

Over the next slides we will try to convince you that this scheme:

1. **Combines the best of both worlds.** Retains slow RW evolution that adapts well to most periods, *plus* a component that can signal crises via $Z_t$.
2. **Enables linear estimation.** No expensive state-space methods needed — estimation reduces to standard linear regression $\Rightarrow$ computational tractability and numerical stability.
3. **Offers a parsimonious replacement for the state error.** In a Bayesian setting, $\gamma Z_t$ is a random term: $\gamma$ (random) provides the stochasticity, while $Z_t$ (deterministic, observed) provides all time-varying patterns not captured by $\beta_{t-1}$. This is more parsimonious than estimating a full state covariance $Q$.
4. **Reduces sensitivity to priors.** TVPs require estimation of a large $Q$. E.g., turning the Bańbura, Giannone & Reichlin (2010) large VAR into a TVP-VAR $\Rightarrow$ $Q$ of dimension $250{,}000 \times 250{,}000$! No information in the likelihood to estimate such a matrix, so the prior dominates everything.
5. **Provides a factor decomposition of TVPs with observed factors.** The term $\gamma Z_t$ decomposes the state innovation into a low-dimensional observable basis — related to the factor decompositions in Canova & Ciccarelli (2004, *JoE*; 2009, *IER*), but with observed rather than latent factors.

# The AVP-VAR: Model and Estimation

# Our starting point: TVP-VAR Framework

Start from the $p$-lag TVP-VAR (Primiceri, 2005):

$$y_t = B_{0t} + \sum_{i=1}^{p} B_{it}\, y_{t-i} + A_t H_t \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, I_n)$$

- $\Omega_t = A_t H_t H_t A_t'$: VAR covariance matrix of reduced-form errors.
- $A_t$: lower triangular $n \times n$ matrix of contemporaneous relations.
- $H_t$: diagonal $n \times n$ matrix of time-varying standard deviations.

$\Rightarrow$ **Ordering dependence:** The lower-triangular structure of $A_t$ makes the decomposition sensitive to the ordering of variables in $y_t$ (Bognanni (2018); Arias et al. (2023, 2024); Chan (2024)).

# Factor Stochastic Volatility

We adopt a factor stochastic volatility (FSV) decomposition of the reduced-form innovations, extending Korobilis (2022), Stock & Watson (2005).

$$y_t = B_{0t} + \sum_{i=1}^{p} B_{it}\, y_{t-i} + \Lambda_t f_t + \Sigma_t^{1/2} \varepsilon_t$$

- $f_t \sim \mathcal{N}(0, I_r)$: vector of latent common factors.
- $\Lambda_t$: $n \times r$ matrix of time-varying loadings.
- $\Sigma_t^{1/2}$: $n \times n$ diagonal matrix of time-varying idiosyncratic standard deviations.
- $\Lambda_t f_t$ captures the common time-varying covariance structure across equations.
- $\Omega_t = \Lambda_t \Lambda_t' + \Sigma_t$ (common + idiosyncratic)

$\Rightarrow$ **Order invariant**: the estimated covariance structure $\Lambda_t \Lambda_t'$ remains the same regardless of how we permute the elements of $y_t$.

# The AVP-VAR in a nutshell

The standard assumption: $\theta_t = \theta_{t-1} + \eta_t$, where $\eta_t$ is the unobserved state innovation.

**Core idea: Replace state innovation with observable exogenous drivers $Z_t$**

$$\theta_t = \theta_{t-1} + \overbrace{\cancel{\eta_t}}^{\text{unobserved}} \implies \theta_t = \theta_{t-1} + \underbrace{\gamma Z_{t-1}}_{\text{observable}} \implies \boxed{\theta_t = \theta + \gamma \underbrace{\textstyle\sum_{s=1}^{t-1} Z_s}_{C_t}}$$

- $Z_t \in \mathbb{R}^m$: observable drivers (uncertainty, financial conditions, expectations, ...).
- $C_t = \sum_{s=1}^{t-1} Z_s$: cumulative drivers — inherits the persistence of a RW without stochastic innovations.
- $\gamma \in \mathbb{R}^{n(k+r)\times m}$: maps drivers to coefficient drift.
- $\theta$: baseline (time-invariant) coefficients.

**Key insight:** All time variation is now driven by $\gamma Z_t$. The state equation collapses into the measurement equation — estimation reduces to a single linear regression. **No need for filtering.**

- Compact notation: $x_t = [1, y'_{t-1}, \ldots, y'_{t-p}]'$, $X_t = I_n \otimes [x'_t, f'_t]$, and $\theta_t = [\beta'_t, \lambda'_t]'$. Then: $y_t = X_t\theta_t + v_t$.

# From TVP-VAR to static regression

Substituting $\theta_t = \theta + \gamma C_t$ into the measurement equation $y_t = X_t\theta_t + v_t$:

$$y_t = \underbrace{X_t\,\theta}_{\text{baseline}} + \underbrace{X_t\,\gamma\,C_t}_{\text{observable time-variation}} + v_t$$

Using the vec rule $X_t\gamma C_t = (C_t' \otimes X_t)\mathrm{vec}(\gamma)$ and stacking over $t = 1, \ldots, T$:

## Static linear regression

$$Y = \tilde{X}(1_T \otimes \theta) + \tilde{W}\,\mathrm{vec}(\gamma) + V$$

- $\tilde{W} = [C_1' \otimes X_1, \ldots, C_T' \otimes X_T]'$: **precomputed** from data.
- Define augmented regressors and coefficients:
  $\tilde{x}_t = [x_t',\ (x_t \otimes C_t)']'$, $\quad \tilde{\beta}_i = [\beta_i',\ \mathrm{vec}(\gamma_{\beta,i})']'$, $\quad \tilde{f}_t = [f_t',\ (f_t \otimes C_t)']'$, $\quad \tilde{\lambda}_i = [\lambda_i',\ \mathrm{vec}(\gamma_{\lambda,i})']'$
- Each equation ($i = 1, \ldots, n$): $\boxed{y_{i,t} = \tilde{x}_t'\tilde{\beta}_i + \tilde{f}_t'\tilde{\lambda}_i + v_{i,t}}$
- **Horseshoe prior** on $\tilde{\beta}_i$ and $\tilde{\lambda}_i$: automatically selects *which* drivers matter and *for which* coefficients — with minimal tuning (Carvalho et al., 2010).

# Estimation: Gibbs Sampler

The model is estimated equation by equation ($i = 1, \ldots, n$), iterating over:

1. **VAR coefficients** $\tilde{\boldsymbol{\beta}}_i$: conjugate Normal posterior under horseshoe prior. Recover time-varying paths via $\boldsymbol{\beta}_{i,t} = \boldsymbol{\beta}_i + \boldsymbol{\gamma}_{\beta,i}\, C_t$.
2. **Factor loadings** $\tilde{\boldsymbol{\lambda}}_i$: analogous conjugate Normal step. Recover $\boldsymbol{\lambda}_{i,t} = \boldsymbol{\lambda}_i + \boldsymbol{\gamma}_{\lambda,i}\, C_t$.
3. **Factors** $\boldsymbol{f}_t$: draw from $\mathcal{N}(\bar{\boldsymbol{G}}_t \boldsymbol{\Lambda}_t' \boldsymbol{\Sigma}_t^{-1} \tilde{\boldsymbol{y}}_t,\ \bar{\boldsymbol{G}}_t)$ with $\bar{\boldsymbol{G}}_t^{-1} = \boldsymbol{I}_r + \boldsymbol{\Lambda}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\Lambda}_t$.
4. **Stochastic volatilities** $\log \sigma_{i,t}$: KSC mixture approximation + precision-based sampler (Kim et al., 1998; Chan, 2012).
5. **Horseshoe hyperparameters**: local $\psi_{i\ell}$ and global $\tau_i$ via slice sampling (Makalic & Schmidt, 2015).

**Remarks:** Steps 1–2 are simple linear regressions (no Kalman filter). Steps 3–5 are standard. The **order-invariant** factor stochastic volatility structure avoids the Cholesky ordering dependence in Primiceri (2005).

▸ Technical Details

# Competitors and Alternatives

# Competitors I: How else can we use observables $Z_t$?

- **TVP-SS-VAR**: Retain stochastic innovations *alongside* the observable drivers: $\theta_t = \theta_{t-1} + \gamma Z_{t-1} + \eta_t$, with $\sigma^2_\eta$ fixed at $0.01$ to prevent explosive VAR behaviour.
  - **TVP-SS-VAR-FLEX**: same, but with $\sigma^2_\eta = 0.1$ (more flexible, riskier).
- **FAVAR / FAVAR-SV**: Augment $y_t$ with a factor extracted from $Z_t$ as an additional endogenous variable (with or without stochastic volatility).
- **INT-VAR**: Interaction VAR — include $x_t \times Z_t$ terms on the right-hand side. Equivalent to AVP-VAR *without* the cumulative-sum structure — no persistence in coefficient variation.
- **Nonlinear basis functions of $Z_t$**: Apply our AVP-VAR framework but with enriched transformations:
  - **TVP-SS-VAR-P / P3**: Polynomials $(Z, Z^2)$ or $(Z, Z^2, Z^3)$.
  - **TVP-SS-VAR-INT / P+INT**: Pairwise interactions among elements of $Z_t$.

# Competitors II: Baseline VARs and ML approaches

- **Baseline VARs** (no external information in TVPs):
  - **CP-VAR / CP-VAR-SV**: Constant parameters, with or without SV.
  - **VAR-SVO / VAR-SVO-t**: Stochastic volatility with outlier states and/or Student-$t$ innovations (Carriero et al., 2023).
  - **TVP-VAR-EB**: Empirical Bayes prior calibrated on training sample (Primiceri, 2005). Ordering dependent.
  - **TVP-VAR-FB**: Full Bayes with horseshoe prior on state innovations.
- **Full ML approach** (Fischer et al., 2023): **FHHP-TVP-VAR** combines three "effect modifiers" — observed drivers $Z_t$, Markov-switching indicators, and latent RW factors.
  - **FHHP-TVP-VAR-R**: Reverse variable ordering.
  - **FHHP-TVP-VAR-d3**: $d = 3$ latent RW factors.
  - **FHHP-TVP-VAR-Z**: Only observed drivers $Z_t$ (no MS, no RW factors).
- **Key issue for all flexible approaches:** No guarantees of stationary parameter draws! More flexible mappings from $Z_t$ to $\theta_t$ risk explosive VAR dynamics.

# Empirical Results: Euro Area

# Data: Euro Area

**Euro Area quarterly** (1975:Q1–2024:Q3):

- $y_t$: real GDP growth, HICP inflation, short-term interest rate ($\Delta$).
- $Z_t$ ($m = 10$): commodity prices (oil, non-oil), unemployment, productivity growth, global activity measures, exchange rates.

**Forecast evaluation:** Expanding window, iterated forecasts at horizons $h = 1, \ldots, 8$ quarters. Performance measured by MAE ratios relative to an OLS-VAR benchmark (values $< 1$ indicate improvement).

Additional results for U.S. monthly data (1985:M1–2023:M11, $m = 12$ drivers) and FRED-QD quarterly data (1963:Q1–2025:Q2, $m = 9$ drivers) available in the paper and online supplement.

# MAE ratios for GDP growth — Baseline VARs

| Model | $h = 1$ | $h = 2$ | $h = 3$ | $h = 4$ | $h = 5$ | $h = 6$ | $h = 7$ | $h = 8$ | Avg |
|---|---|---|---|---|---|---|---|---|---|
| AVP-VAR | †**0.69** | **0.75** | †**0.77** | †**0.75** | **0.73** | †**0.76** | †**0.82** | **0.79** | †**0.76** |
| CP-VAR | **0.81** | **0.87** | **0.87** | **0.84** | **0.80** | **0.82** | **0.88** | **0.82** | **0.84** |
| CP-VAR-SV | **0.70** | **0.78** | **0.79** | **0.77** | **0.76** | **0.82** | **0.87** | **0.81** | **0.79** |
| VAR-SVO | **0.70** | **0.79** | **0.79** | **0.78** | **0.76** | **0.81** | **0.87** | **0.83** | **0.79** |
| VAR-SVO-t | **0.70** | **0.78** | **0.78** | **0.78** | **0.76** | **0.81** | **0.88** | **0.82** | **0.79** |
| TVP-VAR-EB | **0.74** | **0.81** | **0.82** | **0.81** | **0.79** | **0.82** | **0.88** | **0.83** | **0.81** |
| TVP-VAR-FB | **0.71** | †**0.75** | **0.77** | **0.76** | †**0.72** | **0.77** | **0.84** | †**0.78** | **0.76** |
| FHHP-TVP-VAR-L | **0.94** | 1.11 | 1.18 | 1.18 | 1.16 | 1.24 | 1.34 | 1.25 | 1.17 |

Table: MAE ratios relative to OLS-VAR. Bold = below one. † = best at each horizon.

# MAE ratios for GDP growth — Predictor-Driven VARs

| Model | $h = 1$ | $h = 2$ | $h = 3$ | $h = 4$ | $h = 5$ | $h = 6$ | $h = 7$ | $h = 8$ | Avg |
|---|---|---|---|---|---|---|---|---|---|
| AVP-VAR | **†0.69** | **0.75** | **†0.77** | **†0.75** | **†0.73** | **†0.76** | **†0.82** | **†0.79** | **†0.76** |
| FAVAR | **1.00** | 1.06 | 1.01 | 1.01 | 1.03 | 1.01 | 1.04 | 1.05 | 1.03 |
| FAVAR-SV | **0.69** | **0.75** | **0.78** | **0.76** | **0.73** | **0.78** | **0.86** | **0.80** | **0.77** |
| INT-VAR | **0.69** | **†0.75** | **0.80** | **0.79** | **0.75** | **0.78** | **0.85** | **0.79** | **0.78** |
| TVP-SS-VAR | **0.92** | 1.03 | 1.08 | 1.08 | 1.11 | 1.21 | 1.32 | 1.29 | 1.13 |
| TVP-SS-VAR-FLEX | **0.93** | 1.04 | 1.17 | 1.17 | 1.27 | 1.43 | 1.77 | 1.88 | 1.33 |
| TVP-SS-VAR-P | 1.46 | 1.50 | 1.72 | 1.79 | 1.89 | 1.80 | 2.38 | 2.25 | 1.85 |
| TVP-SS-VAR-INT | 1.11 | 1.15 | 1.24 | 1.24 | 1.26 | 1.35 | 1.52 | 1.49 | 1.29 |
| TVP-SS-VAR-P3 | 3.32 | 3.63 | 4.28 | 4.02 | 4.43 | 4.35 | 5.58 | 4.79 | 4.30 |
| TVP-SS-VAR-P+INT | **0.98** | 1.04 | 1.09 | 1.10 | 1.08 | 1.21 | 1.35 | 1.33 | 1.15 |
| FHHP-TVP-VAR | **0.86** | **0.97** | 1.09 | 1.10 | 1.11 | 1.18 | 1.30 | 1.23 | 1.11 |
| FHHP-TVP-VAR-R | **0.98** | 1.08 | 1.17 | 1.20 | 1.18 | 1.24 | 1.38 | 1.28 | 1.19 |
| FHHP-TVP-VAR-d3 | **0.88** | 1.01 | 1.11 | 1.13 | 1.13 | 1.20 | 1.35 | 1.26 | 1.13 |
| FHHP-TVP-VAR-Z | **0.97** | 1.08 | 1.19 | 1.21 | 1.20 | 1.25 | 1.38 | 1.27 | 1.19 |

Table: MAE ratios relative to OLS-VAR. Bold = below one. † = best at each horizon.

# MAE ratios for Inflation — Baseline VARs

| Model | $h = 1$ | $h = 2$ | $h = 3$ | $h = 4$ | $h = 5$ | $h = 6$ | $h = 7$ | $h = 8$ | Avg |
|---|---|---|---|---|---|---|---|---|---|
| AVP-VAR | 1.15 | 1.02 | **0.93** | †**0.90** | **0.95** | **0.92** | **0.89** | †**0.88** | **0.95** |
| CP-VAR | **0.97** | **0.95** | **0.93** | **0.96** | **0.95** | **0.92** | **0.91** | **0.92** | **0.94** |
| CP-VAR-SV | 1.01 | †**0.94** | **0.91** | **0.91** | **0.93** | †**0.87** | †**0.87** | **0.88** | †**0.91** |
| VAR-SVO | 1.18 | 1.03 | 1.03 | **0.98** | 1.03 | **0.92** | **0.95** | **0.92** | 1.00 |
| VAR-SVO-t | 1.19 | 1.03 | 1.02 | **0.98** | 1.03 | **0.91** | **0.95** | **0.91** | 1.00 |
| TVP-VAR-EB | 1.02 | **0.95** | †**0.90** | **0.93** | **0.94** | **0.90** | **0.89** | **0.92** | **0.93** |
| TVP-VAR-FB | 1.63 | 1.55 | 1.39 | 1.37 | 1.14 | 1.09 | 1.08 | 1.10 | 1.29 |
| FHHP-TVP-VAR-L | †**0.91** | 1.79 | 1.18 | 1.64 | †**0.90** | 1.34 | **0.99** | 1.32 | 1.26 |

Table: MAE ratios relative to OLS-VAR. Bold = below one. † = best at each horizon.

# MAE ratios for Inflation — Predictor-Driven VARs

| Model | $h = 1$ | $h = 2$ | $h = 3$ | $h = 4$ | $h = 5$ | $h = 6$ | $h = 7$ | $h = 8$ | Avg |
|---|---|---|---|---|---|---|---|---|---|
| AVP-VAR | 1.15 | 1.02 | **0.93** | †**0.90** | **0.95** | **0.92** | **0.89** | **0.88** | **0.95** |
| FAVAR | 1.01 | 1.00 | **0.98** | **1.00** | **1.00** | **0.99** | **0.98** | **1.00** | **0.99** |
| FAVAR-SV | 1.02 | †**0.94** | †**0.89** | **0.92** | **0.93** | **0.87** | †**0.86** | **0.88** | †**0.91** |
| INT-VAR | 1.36 | 1.44 | 1.42 | 1.14 | **0.95** | †**0.85** | **0.88** | †**0.85** | 1.11 |
| TVP-SS-VAR | 1.23 | 1.14 | 1.03 | 1.03 | 1.01 | **0.98** | **0.95** | **0.94** | 1.04 |
| TVP-SS-VAR-FLEX | 1.20 | 1.22 | 1.04 | 1.13 | 1.15 | 1.17 | 1.15 | 1.24 | 1.16 |
| TVP-SS-VAR-P | 1.40 | 1.22 | 1.13 | 1.08 | 1.14 | 1.05 | 1.01 | 1.06 | 1.13 |
| TVP-SS-VAR-INT | 1.27 | 1.14 | 1.07 | 1.09 | **1.00** | **0.96** | **0.91** | **1.00** | 1.05 |
| TVP-SS-VAR-P3 | 1.39 | 1.54 | 1.64 | 1.51 | 1.28 | 1.44 | 1.45 | 1.51 | 1.47 |
| TVP-SS-VAR-P+INT | 1.27 | 1.11 | **0.99** | 1.02 | **0.96** | **0.93** | **0.88** | **0.91** | 1.01 |
| FHHP-TVP-VAR | **0.94** | 1.82 | 1.21 | 1.69 | **0.90** | 1.37 | 1.00 | 1.37 | 1.29 |
| FHHP-TVP-VAR-R | **0.98** | 1.88 | 1.23 | 1.72 | **0.91** | 1.36 | **0.98** | 1.36 | 1.30 |
| FHHP-TVP-VAR-d3 | †**0.93** | 1.81 | 1.18 | 1.69 | †**0.90** | 1.35 | **0.99** | 1.35 | 1.28 |
| FHHP-TVP-VAR-Z | **0.95** | 1.82 | 1.20 | 1.68 | **0.90** | 1.35 | 1.00 | 1.35 | 1.28 |

Table: MAE ratios relative to OLS-VAR. Bold = below one. † = best at each horizon.

# Cumulative forecast gains over time (U.S. Industrial Production)

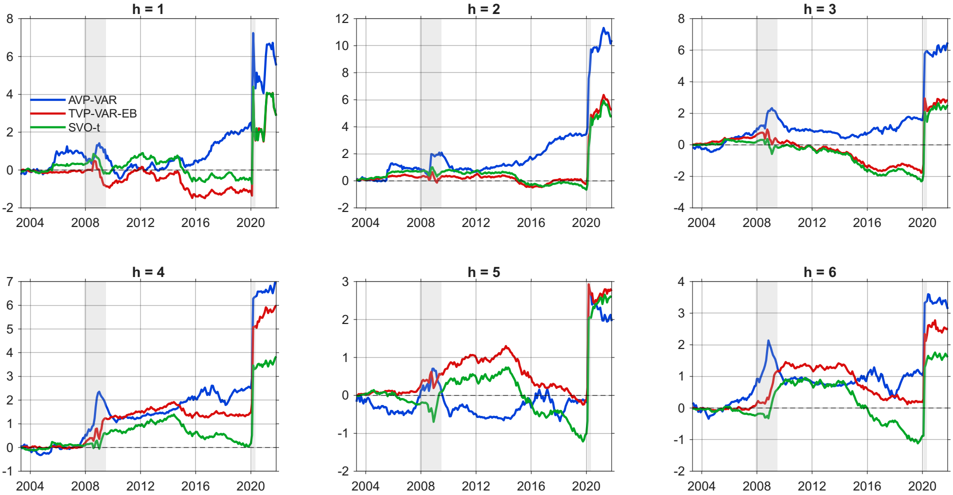
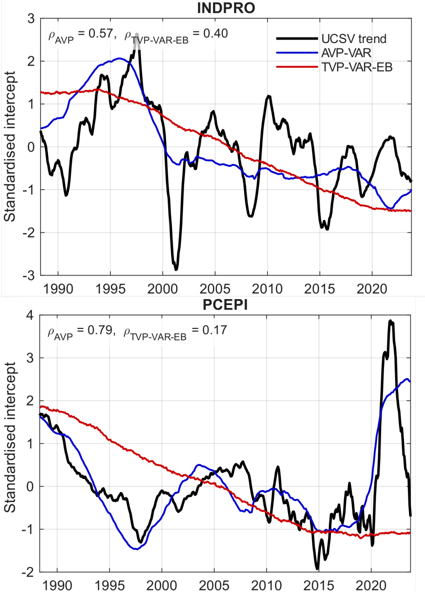



Figure: Cumulative difference in absolute forecast errors relative to a constant-parameter VAR. Values above zero indicate the model outperforms the benchmark.

# Understanding the Mechanism

# Why does AVP-VAR forecast well? In-sample dynamics




Time-varying intercepts from a VAR(1) on U.S. monthly data (1985–2023).

**Three key observations:**

1. **Timeliness:** AVP-VAR intercepts adjust *rapidly* during crises (GFC, COVID), closely tracking UC-SV. TVP-VAR adjusts too slowly.
2. **Discipline:** Outside crises, AVP-VAR stays stable (no signal in $Z_t$). TVP-VAR sometimes produces erratic local movements.
3. **Heterogeneity:** Different equations display different amounts of variation — AVP-VAR accommodates this naturally via sparse driver selection.

# Monte Carlo Evidence

# Monte Carlo: Do observable drivers help recover true TVPs?

- We design two DGPs to stress-test the AVP-VAR:
  - **DGP 1**: Near-unit-root parameters with transitory jumps (mimicking COVID-type breaks).
  - **DGP 2**: Regime switching + threshold effects + volatility co-movement (complex non-linear dynamics).
- Compare AVP with *targeted* drivers (informative about break timing, not magnitude) vs. *agnostic* drivers (Gaussian noise) vs. canonical RW TVP.
- **Key findings:**
  1. Under DGP 1, even *agnostic* AVP matches or beats TVP when predictors are persistent ($\rho = 0.95$). Targeted AVP dominates strongly.
  2. Under DGP 2, agnostic AVP underperforms TVP (needs informative signals for complex dynamics). But targeted AVP with $m \geq 40$ drivers dominates TVP across all configurations, with MSPE reductions up to 79%.
  3. The number of drivers $m$ matters more under DGP 2: richer information sets dramatically improve performance.
- **Bottom line:** When the RW evolution is contaminated with jumps, regime shifts, or threshold effects, observable adaptation yields more accurate parameter estimates — provided the drivers carry at least some relevant information.

# Conclusions

# Takeaways

1. **The AVP-VAR** replaces unobserved state innovations with observable drivers:
   - No Kalman filter needed — estimation is a standard linear regression.
   - Interpretable: practitioners know *why* parameters change.
2. **Monte Carlo evidence:** AVP-VAR accurately recovers the true time-varying DGP, especially when the true dynamics contain abrupt breaks or regime shifts.
3. **Empirical results** across three macroeconomic datasets:
   - Competitive with or superior to most benchmarks for output growth.
   - Strong performance for inflation, particularly at medium-to-long horizons.
   - Robust across lag orders — parsimony advantage over richer specifications.
4. **In-sample analysis** reveals the mechanism: conventional TVP-VARs over-regularize, leaving parameters flat. AVP-VAR tracks crisis dynamics while staying disciplined in tranquil periods.
5. **More flexible alternatives** (polynomials, interactions, ML methods) tend to produce explosive VAR dynamics — the linear, deterministic specification is not a limitation but a *feature*.

# Thank you!


**Nicolás Hardy**

nicolas.hardy@udp.cl

**Dimitris Korobilis**

dimitris.korobilis@glasgow.ac.uk

# AVP-VAR: Gibbs Sampler (I)

Define augmented regressors and coefficient vectors: $\tilde{\boldsymbol{x}}_t = [\boldsymbol{x}_t', (\boldsymbol{x}_t \otimes C_t)']' \in \mathbb{R}^{k+km}$, $\tilde{\boldsymbol{\beta}}_i = [\boldsymbol{\beta}_i', \mathrm{vec}(\boldsymbol{\gamma}_{\beta,i})']' \in \mathbb{R}^{k+km}$ and $\tilde{\boldsymbol{f}}_t = [\boldsymbol{f}_t', (\boldsymbol{f}_t \otimes C_t)']' \in \mathbb{R}^{r+rm}$, $\tilde{\boldsymbol{\lambda}}_i = [\boldsymbol{\lambda}_i', \mathrm{vec}(\boldsymbol{\gamma}_{\lambda,i})']' \in \mathbb{R}^{r+rm}$. Working equation by equation $(i = 1, \ldots, n)$

$$\boxed{y_{i,t} = \tilde{\boldsymbol{x}}_t' \tilde{\boldsymbol{\beta}}_i + \tilde{\boldsymbol{f}}_t' \tilde{\boldsymbol{\lambda}}_i + v_{i,t}, \qquad v_{i,t} \sim \mathcal{N}(0, \sigma_{i,t}^2)}$$

**Step 1: Augmented VAR coefficients** $\tilde{\boldsymbol{\beta}}_i, \quad i = 1, \ldots, n.$

Define $\tilde{y}_{i,t} = y_{i,t} - \boldsymbol{f}_t' \boldsymbol{\lambda}_{i,t}$. Place a horseshoe prior: $\tilde{\beta}_{i,\ell} \mid \psi_\ell, \tau_i \sim \mathcal{N}(0, \psi_\ell^2 \tau_i^2)$.

$$\tilde{\boldsymbol{\beta}}_i \mid \cdot \ \sim \ \mathcal{N}(\bar{\boldsymbol{\mu}}_{\beta,i}, \bar{\boldsymbol{V}}_{\beta,i}); \qquad \bar{\boldsymbol{V}}_{\beta,i}^{-1} = \boldsymbol{D}_{\beta,i}^{-1} + \sum_{t=1}^{T} \sigma_{i,t}^{-2} \tilde{\boldsymbol{x}}_t \tilde{\boldsymbol{x}}_t', \qquad \bar{\boldsymbol{\mu}}_{\beta,i} = \bar{\boldsymbol{V}}_{\beta,i} \sum_{t=1}^{T} \sigma_{i,t}^{-2} \tilde{\boldsymbol{x}}_t \tilde{y}_{i,t}$$

where $\boldsymbol{D}_{\beta,i} = \mathrm{diag}(\psi_{i,1}^2 \tau_i^2, \ldots, \psi_{i,k+km}^2 \tau_i^2)$. Recover $\boldsymbol{\beta}_{i,t} = \boldsymbol{\beta}_i + \boldsymbol{\gamma}_{\beta,i} C_t$.

**Step 2: Augmented factor loadings** $\tilde{\boldsymbol{\lambda}}_i, \quad i = 1, \ldots, n.$

Define $\hat{y}_{i,t} = y_{i,t} - \boldsymbol{x}_t' \boldsymbol{\beta}_{i,t}$. Analogous horseshoe posterior:

$$\tilde{\boldsymbol{\lambda}}_i \mid \cdot \ \sim \ \mathcal{N}(\bar{\boldsymbol{\mu}}_{\lambda,i}, \bar{\boldsymbol{V}}_{\lambda,i})$$

Recover $\boldsymbol{\lambda}_{i,t} = \boldsymbol{\lambda}_i + \boldsymbol{\gamma}_{\lambda,i} C_t$.

# AVP-VAR: Gibbs Sampler (II)

**Step 3: Factors** $\boldsymbol{f}_t, \quad t = 1, \dots, T$.

Define $\tilde{\boldsymbol{y}}_t = \boldsymbol{y}_t - (\boldsymbol{I}_n \otimes \boldsymbol{x}_t')\boldsymbol{\beta}_t$. The conditional posterior is:

$$\boldsymbol{f}_t \mid \cdot \sim \mathcal{N}\left(\bar{\boldsymbol{G}}_t \boldsymbol{\Lambda}_t' \boldsymbol{\Sigma}_t^{-1} \tilde{\boldsymbol{y}}_t,\ \bar{\boldsymbol{G}}_t\right), \qquad \bar{\boldsymbol{G}}_t^{-1} = \boldsymbol{I}_r + \boldsymbol{\Lambda}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\Lambda}_t$$

**Step 4: Stochastic volatilities** $\log \sigma_{i,t}, \quad i = 1, \dots, n$.

Let $v_{i,t} = y_{i,t} - \boldsymbol{x}_t' \boldsymbol{\beta}_{i,t} - \boldsymbol{f}_t' \boldsymbol{\lambda}_{i,t}$. Define $y_{i,t}^* = \log(v_{i,t}^2)$.

- Approximate $\log \chi_1^2$ by a 7-component Gaussian mixture (Kim, Shephard & Chib, 1998).
- Draw $(\log \sigma_{i,1}, \dots, \log \sigma_{i,T})'$ via precision-based sampler.
- Update $\omega_i^2 \mid \cdot \sim \text{Inv-Gamma}\left(\frac{a_0+T-1}{2}, \frac{b_0+\sum_{t=2}^T (\Delta \log \sigma_{i,t})^2}{2}\right)$ with $a_0 = 1$, $b_0 = 0.01$.

**Step 5: Horseshoe hyperparameters.**

Local scales $\{\psi_{i\ell}\}$ and global scales $\tau_i$, $\tau_{\lambda,i}$ updated via slice sampling.

# Model Nesting and Special Cases