Deep Vector Autoregression for Macroeconomic Data

Patrick Altmeyer (p.altmeyer@tudelft.nl), Delft University of Technology
Marc Agusti, European Central Bank
Ignacio Vidal-Quadras Costa, European Central Bank

Can we leverage the power of deep learning in VAR models?

We propose **Deep VAR**: an approach towards VAR that leverages the power of deep learning in order to model non-linear relationships.

Key contributions

- Simple methodology close in spirit to conventional benchmark.
- Significant improvement in predictive performance.
- Open source [R package]() to facilitate reproducibility.

Less is more

A simple methodology

We developed our idea under the following premise: **maximise performance** of an existing and trusted framework under **minimal intervention**. We relax the assumption of linearity through one simple step: estimate each system equation through a recurrent neural network.

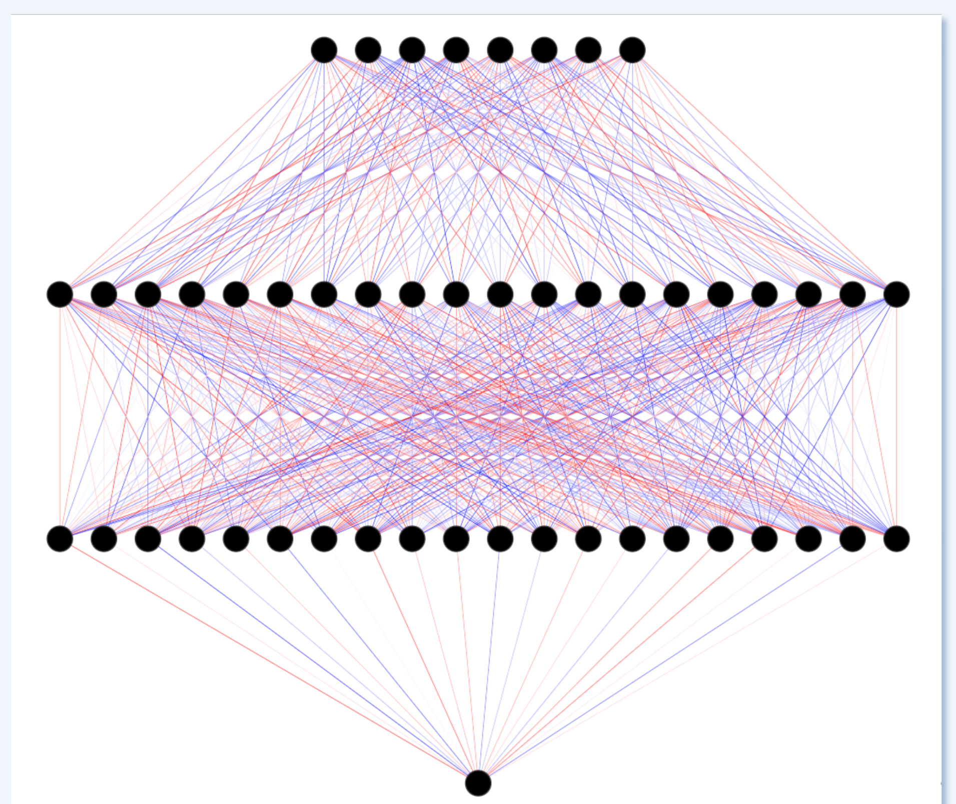

Figure 1: Neural Network Architecture. Inputs are lags of all variables. Output is variable of interest in time *t*.

Significant performance gains

Empirical evidence

To evaluate our proposed methodology empirically we use the **FRED-MD data base to collect a sample of monthly US data** on output (IP), unemployment (UR), inflation (CPI) and interest rates (FFR). Our sample spans the period from January 1959 through March 2021.

Significantly reduced in-sample error …

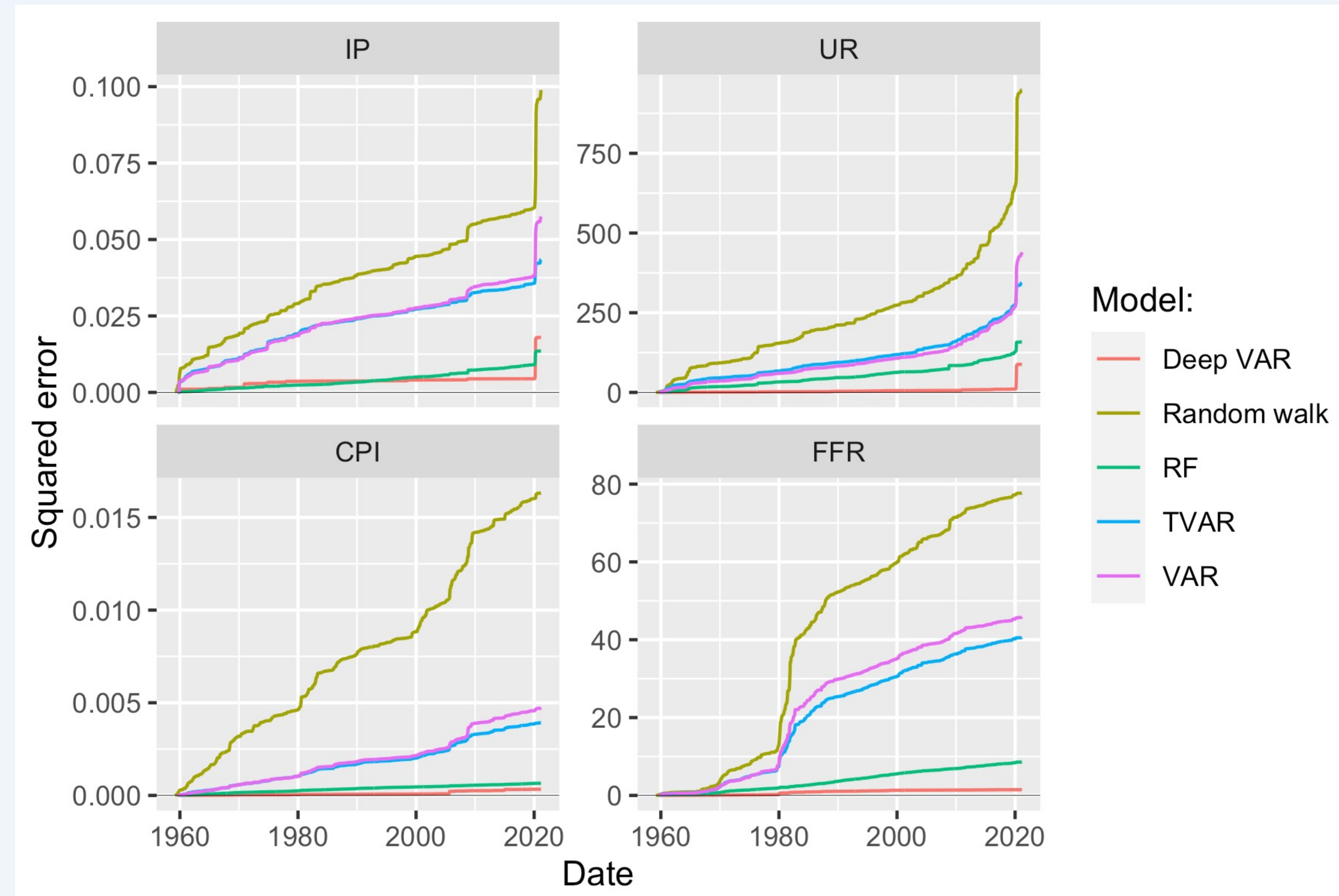

Figure 2: Comparison of cumulative loss over the entire sample period for Deep VAR and benchmarks.

… out-of-sample error …

Variable	DVAR	VAR	Ratio (DVAR / VAR)
IP	0.00494	0.01484	0.33253
UR	0.94542	1.65170	0.57240
CPI	0.00231	0.00342	0.67642
FFR	0.17494	0.23974	0.72972

Table 1: Test root mean squared error (RMSE) for the two models across variables.

… and multi-step-ahead prediction error.

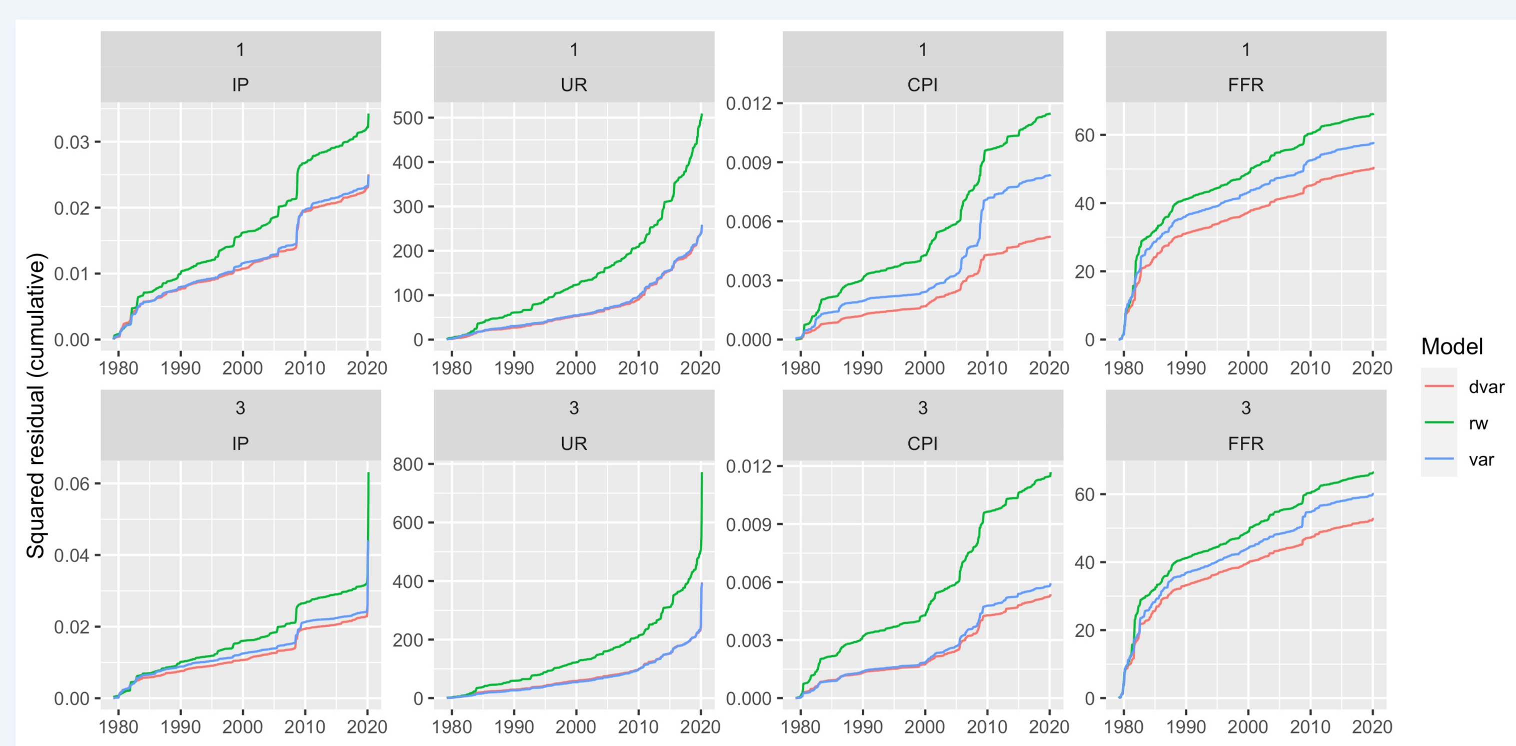

Figure 3: Cumulative rolling-window prediction error for the 1-month and 3-month horizon. TVAR not shown here, since errors blew out by too much.

Have we merely scratched the surface?

Hyperparameter tuning

For a our baseline comparison we keep things simple: for example, we let the conventional VAR guide our search for optimal lag length. A short exercise in hyperparameter tuning demonstrates that the Deep VAR is less prone to overfitting with respect to the number of lags among other things.

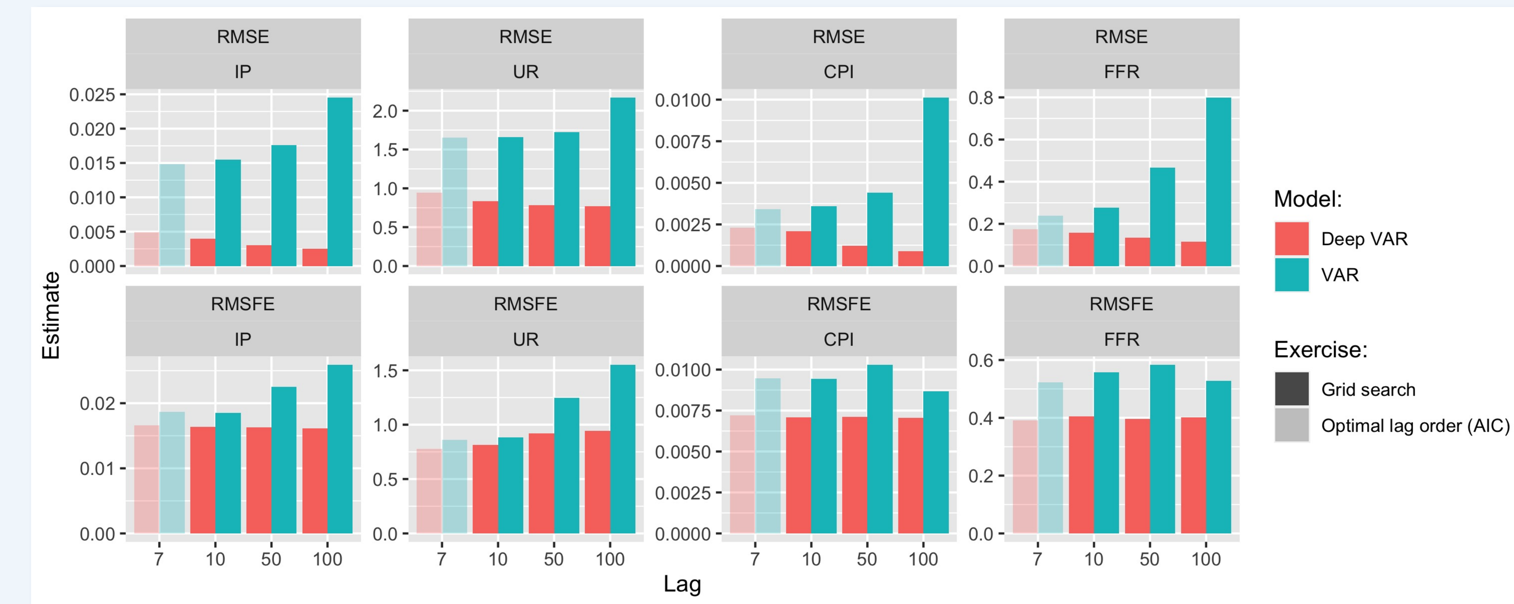

Figure 4:: Pseudo out-of-sample RMSE and RMSFE for VAR and Deep VAR across different lag choices.

Where to go from here

Avenues for future research

- ❑ Verify if outperformance is **robust** through additional data & benchmarks.
- ❑ **Uncertainty quantification:** can we just assume Gaussian residuals? Probably not. Bootstrap? Costly!
- ❑ From deterministic to **Bayesian deep learning**: this enables uncertainty quantification and should aid with **interpretability and robustness**.
  - ➢ Recent work by Daxberger et al. (2021) shows that **Laplace Approximation** could be a promising way forward.
- ❑ Can the existing **inference toolbox** (IRFs, variance decomposition, policy counterfactuals, …) be developed for Deep VAR?
- ❑ **Structural identification** – how to proceed? Verstyuk (2020) works with Cholesky decomposition as in conventional VAR. What about GNN?

References

Daxberger et al. (2021). "Laplace Redux-Effortless Bayesian Deep Learning.". In: Advances in Neural Information Processing Systems 34

Kilian and Luetkepohl (2017). "Structural Vector Autoregressive Analysis.". In: Cambridge University Press.

Verstyuk (2020). "Modeling Multivariate Time Series in Economics: From Auto-Regressions to Recurrent Neural Networks.". In: Available at SSRN 3589337.