

Ratcliff, Ryan D.

Working Paper

A Comment on "Measuring Monetary Policy in the Euro Area Using SVARs with Residual Restrictions"

I4R Discussion Paper Series, No. 160

Provided in Cooperation with:

The Institute for Replication (I4R)

Suggested Citation: Ratcliff, Ryan D. (2024) : A Comment on "Measuring Monetary Policy in the Euro Area Using SVARs with Residual Restrictions", I4R Discussion Paper Series, No. 160, Institute for Replication (I4R), s.l.

This Version is available at:

<https://hdl.handle.net/10419/303192>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

No. 160

I4R DISCUSSION PAPER SERIES

A Comment on “Measuring Monetary Policy in the Euro Area Using SVARs with Residual Restrictions”

Ryan D. Ratcliff

September 2024

I4R DISCUSSION PAPER SERIES

I4R DP No. 160

A Comment on “Measuring Monetary Policy in the Euro Area Using SVARs with Residual Restrictions”

Ryan D. Ratcliff¹

¹University of San Diego/USA

SEPTEMBER 2024

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](#), and [RWI – Leibniz Institute for Economic Research](#), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

A comment on “Measuring Monetary Policy in the Euro Area Using SVARs with Residual Restrictions”*

Ryan D. Ratcliff

August 30, 2024

Abstract

[Badinger and Schiman \(2023\)](#) use a narrative high-frequency analysis of news and financial markets to develop a small set of restrictions on the structural shocks of a VAR of the Euro area. Their approach does not uniquely identify a structural representation, so their results are based on the distribution of a randomly generated set of parameters that satisfies the restrictions. Their method generates impulse responses that are consistent with macroeconomic theory, but that differ from previous studies that use alternative high-frequency identification strategies. They use this difference to argue that, unlike previous studies, their method is able to separate monetary policy surprises from confounding central bank information shocks – an important new contribution to the literature. I conducted two replication studies of their work on behalf of the Institute for Replication (I4R). First, I used the code provided in their replication package to replicate all of their main results, aside from the small variations expected in replicating a Monte Carlo study. Second, I attempted to use their original data to recreate their results using a different statistical software (Eviews 13). I was unable to replicate their results for two reasons. First, my program is unable to exactly replicate the custom prior they used to generate their reduced-form results. Second, my models routinely generated nonstationary VARs that nevertheless satisfied the identification restrictions. This differs from the author’s results, but is not surprising given the ambiguous stationarity of the underlying macro variables.

KEYWORDS: Structural VAR, Residual Sign Restrictions, Replication

JEL CODES: C32, E43, E44, E52, E58

*Authors: Ratcliff: University of San Diego. E-mail: ratcliff@san-diego.edu. The author received no financial support for this work, and has no conflicts of interest.

1 Introduction:

This report is a replication study of [Badinger and Schiman \(2023\)](#) prepared for the Institute for Replication ([Brodeur et al. 2024](#)).

Using a narrative high-frequency analysis of news, financial markets, and ECB communications, [Badinger and Schiman \(2023\)](#) develop a new, parsimonious identification strategy based on a small set of sign and magnitude restrictions on the shocks in a structural vector autoregression (SVAR). The sample period is monthly Euro zone data from January 1999 to December 2019 (252 months), using data on interest rates (EONIA), industrial production, the harmonized CPI targeted by the ECB, corporate bond spreads, narrow money supply M1, and the euro per dollar exchange rate. EONIA and the spread variable are percentages; all other variables are in natural logs in their data.

The authors begin with a Bayesian estimation of the reduced form VAR (equation 1), using “an independent Normal-Wishart prior with Minnesota-style shrinkage...” ([Badinger and Schiman 2023](#), p. 14).

$$\mathbf{y}_t = \mathbf{c} + \sum_{i=1}^{12} \mathbf{A}_i \mathbf{y}_{t-i} + \mathbf{u}_t \quad (1)$$

where \mathbf{y}_t is a 6x1 vector of the variables described above. They sample reduced-form matrices \mathbf{A}_i from the posterior distribution, collectively stacked into \mathbf{A} . They then seek a structural matrix \mathbf{B} that is consistent with their restrictions:

$$\mathbf{B}\mathbf{y}_t = \mathbf{B}\mathbf{c} + \sum_{i=1}^{12} \mathbf{B}\mathbf{A}_i \mathbf{y}_{t-i} + \mathbf{w}_t \quad (2)$$

where $\mathbf{w}_t = \mathbf{B}\mathbf{u}_t$ represents the independent structural shocks, so $\mathbf{w}_t \sim N(\mathbf{0}, \mathbf{I})$.

Their narrative approach identifies two dates when the monetary policy shock should be positive (i.e. a contractionary surprise), and two when it should be negative. In addition, their narrative analysis suggests a magnitude restriction for November 2011 where the monetary policy shock should account for at least 50% of the total unexplained change in the EONIA interest rate. Since these restrictions are not sufficient to uniquely identify \mathbf{B} , their approach is to draw a random matrix, and do a QR decomposition where the orthogonal factor is a candidate for \mathbf{B} . They

then test whether the structural shocks implied by that candidate matrix meet the identifying assumptions. They sample the posterior distribution and draw candidate **B** matrices until they have found 5000 combinations of **A** and **B** that satisfy their restrictions.

Their first preliminary result is to show that the impulse response function (IRF) of the variables to a positive (restrictive) shock to monetary policy are consistent with mainstream macroeconomic theory. They focus on the IRF implied by the median of their 5000 successful draws, and judge statistical significance by plotting a 68% credible set around that median using the appropriate percentiles of the distribution of their successful draws. EONIA has a short-lived positive response to a contractionary monetary surprise, while industrial production and consumer prices have negative responses, but only after a year or more. The 68% credible sets for the remaining variables never exclude zero, implying there is not a significant response.

The authors' key contribution is to contrast these results with results from other recent SVARs of the Euro zone, arguing that their method removes an important confounding effect. The authors replicate a high-frequency approach from previous studies that uses the estimated monetary policy shock in a new VAR as an instrument. Consistent with these previous studies, they show this instrumental approach does not find a significant response to the estimated monetary policy shock for industrial production or consumer prices – a counterintuitive result. The previous studies argue that this non-result comes from the identified shocks being a mix of both monetary policy surprises as well as information shocks where the central bank receives superior information on the state of the economy and reacts to it in a way that looks like a surprise to the modeler. Badinger and Schiman argue that the significant responses from their identification strategy indicate that their approach removes the confounding effects of these information shocks.

The paper proceeds from there with several robustness checks, and compares their structural shocks to other interest rate surprise indicators. While these results are interesting in the context of the original study, they are less relevant to my replication efforts.

Section 2 of my comment reports the results of our direct computational repro-

duction using the author’s MATLAB code and data. Since their method involves randomly drawing candidate parameters and then checking whether the sign and magnitude restrictions are satisfied, there is a small amount of variation for each run of the program, but overall I was able to reproduce all of the main results of their paper using their code.

Section 3 describes my efforts at a recreation of their study. Using the same data, I attempted to reproduce their results by recreating their code in Eviews 13. Here, due to a limitation in the Eviews software I was unable to exactly replicate the prior they used in calculating the posterior distribution for the \mathbf{A} matrix. Approximating their approach as closely as I could in Eviews, I found that my estimated models were often nonstationary, in contrast to the authors’ results.

2 Computational Reproducibility

I used the MATLAB-based replication package provided by the authors [here](#). The code was complete and ran without errors, replicating all of the main exhibits in the original paper. Since their method involves random draws for both the posterior distribution of \mathbf{A} as well as candidate structural matrices \mathbf{B} , I found trivial differences in the results: for example, my impulse response function for EONIA’s response to a monetary policy shock peaked one month earlier than IRF shown in Figure 2 in the published paper. Such differences are not surprising, and at no point contradict their published results.

To aid in the replication efforts, I also supplemented their MATLAB code to produce some diagnostic information on the success rates (how often a candidate draw satisfied the sign and/or magnitude restrictions). These results are presented in Table 1 below.

Overall, 6.3 million draws were required to reach 5000 combinations of \mathbf{A} and \mathbf{B} that satisfied the sign and magnitude restrictions. With no learning from an unsuccessful draw, this brute force approach generates a successful draw about 0.08% of the time. The magnitude restrictions are satisfied much more often than the sign restrictions, which is consistent with their observation at the end of their conclusion that “...magnitude restrictions appear to be crucial for the identification

of SVARs with residual sign restrictions.” (Badinger and Schiman 2023, p. 27) ¹

3 Recreation Reproducibility: Same Data in Eviews 13

My second replication effort was a mix of computational and recreation reproducibility: using the same data, I attempted to reproduce their results by recreating their MATLAB code in Eviews 13.² Eviews is widely used for time series analysis, and has an extensive set of tools for Bayesian VAR modelling. These procedures are extensively documented, but the proprietary code is not accessible by users. As a first step, I verified that taking the results of a random draw from MATLAB (i.e. **A** and/or **B**), the Eviews code was able to exactly replicate the downstream results in MATLAB. The Eviews code had similar success rates as reported in Table 1.

I was unable to recreate the authors’ results in Eviews 13: the SVAR estimates from my code included many parameter values that led to non-stationary models, explosive impulse responses, etc. Since the Eviews code replicates their intermediate results, the different estimated models must come from how our respective programs handle the random draws from the posterior distribution. The authors use a unique, custom-coded “independent Normal-Wishart prior with Minnesota-style shrinkage...” that I was unable to exactly replicate (Badinger and Schiman 2023, p. 14). Eviews 13’s standard procedures offer two possible approaches that might replicate their prior. The first is a basic Litterman/Minnesota prior that can match the shrinkage hyperparameters used by the authors, but that involves assuming a known covariance matrix. The other option is an independent Normal-Wishart prior that does not use these shrinkage hyperparameters, and involves an iterative Gibbs sampler of the *conditional* posterior distribution (Eviews 13 User Guide 2022). I ran the model using both available prior options, but neither one generated a stationary VAR.

At one level, this can simply be a limitation of the software: I was unable to recreate their results because Eviews does not offer enough flexibility with the prior distribution to recreate the authors’ approach. But it does raise the question of

¹The authors discuss some complementary evidence on how the magnitude restriction improves identification in Online Appendix D3.

²The Eviews code, etc. is available at <https://github.com/RDRatcliff/BVAR-Replication>

how much the authors' results are sensitive to this custom prior and their particular shrinkage hyperparameters. They do address this briefly in their online Appendix D3, where they report that using a conjugate prior with a slightly different shrinkage setup leaves their results largely unchanged.

The differences in our results could also be something more fundamental: my estimated VARs were often nonstationary in the levels specification used by the authors because the underlying data is in fact nonstationary. Figure 1 show a plot of the EONIA interest rate series over the life of the sample.

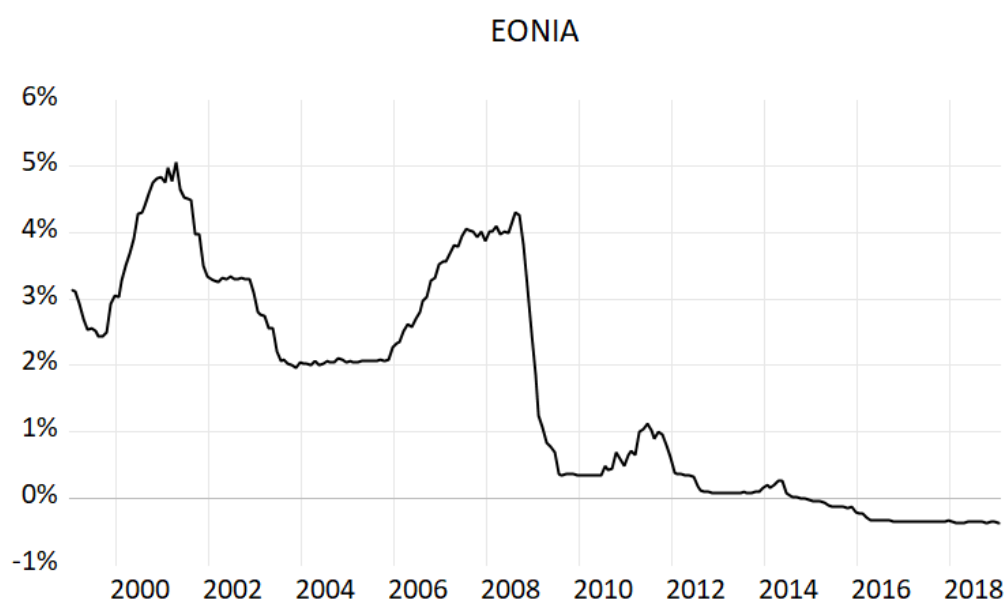


Figure 1: EONIA, 1999-2019

For the sample period used, it is implausible that EONIA is stationary around some constant mean. Table 2 reports the results of a standard battery of Ng-Perron and KPSS univariate unit root tests for all the variables in the VAR. EONIA, HCPI, and M1 are nonstationary, failing to reject the null hypothesis of a unit root in the NP tests, while rejecting the KPSS null hypothesis of stationarity at 1%. The other variables give mixed results across the two tests. Looking at the variables collectively, the CIPS panel unit root test statistic for cross-sectionally dependent data was -1.94 for the six series combined, failing to reject the presence of a unit root.³ Additionally, of the three reduced form VARs I estimated with the data, the

³The 10% critical value is -2.2 (Pesaran 2007)

characteristic polynomial from the BVAR with the independent Normal-Wishart prior had roots outside the unit circle (i.e. is nonstationary as estimated) and the OLS and Minnesota BVAR both had roots that were greater than 0.999.

These results on nonstationarity in no way suggests that the authors' results are invalid. Instead, it argues that future work in this space should include robustness/specification checks around the stationarity of the estimated models, as well as a wider investigation of alternative priors.

4 Conclusion

Overall, my efforts to recreate the results of [Badinger and Schiman \(2023\)](#) met with mixed success. I was able to recreate all of the major results of the paper using the authors' provided replication package, except for small variations that one would expect from different random draws from the same distribution. However, when I attempted to recreate their study in Eviews 13 using their original data, I was unable to replicate even their basic results. My code was able to replicate all of the steps of their original program, except that Eviews does not allow the flexibility to implement their customized approach to the Bayesian prior.

This single difference led to substantially different outcomes: almost all of my VARs were nonstationary. This is hardly surprising, as empirical macroeconomics has struggled with the potential nonstationarity of our data and its modelling implications for decades. If nonstationarity was a problem for me and not for the authors, and the only difference between my code and theirs is the handling of the Bayesian prior, then it may be the case that this prior favors stationary models in a way that is not immediately obvious.

To be clear, I do not feel that any of the proceeding discussion invalidates the authors' results, and there is absolutely nothing wrong with the way the authors approached specifying their prior. Instead, this seemingly innocuous choice of prior may have a more significant impact on the results than one would expect. In the same way that the authors suggest that magnitude restrictions are an important part of implementing their identification approach going forward, my results would suggest that future work using this approach should examine the issue of prior

choice and its implications for the stationarity of the model in more detail.

References

- Badinger, H. and Schiman, S.: 2023, Measuring monetary policy in the euro area using svars with residual restrictions, *American Economic Journal: Macroeconomics* **15**(2), 279–305.
- Brodeur, A., Mikola, D. and Cook, N.: 2024, Mass reproducibility and replicability: A new hope, *Institute for Replication(I4R) Discussion Paper 107* .
- Eviews 13 User Guide*: 2022, S&P Global.
- Kwiatkowski, D., Phillips, P. C., Schmidt, P. and Shin, Y.: 1992, Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?, *Journal of econometrics* **54**(1-3), 159–178.
- Ng, S. and Perron, P.: 2001, Lag length selection and the construction of unit root tests with good size and power, *Econometrica* **69**(6), 1519–1554.
- Pesaran, M. H.: 2007, A simple panel unit root test in the presence of cross-section dependence, *Journal of applied econometrics* **22**(2), 265–312.

5 Tables

Table 1: MATLAB Success Rates

	Count	Percent
Total Draws	6.3m	
Total Post. Draws	5,580	
Met Sign Restr.	90,092	1.42%
Met Mag. Restr.	1.2m	19.7%
Met Both	5000	0.08%

Notes: After each draw of the A matrices from the prior, up to 100,000 candidate B matrices were drawn, stopping early on a successful draw for B (both the sign and magnitude restrictions are met).

Table 2: Univariate Unit Root Tests

Test	EONIA	Ind. Prod	HCPI	CB Spread	M1	EUR/USD
H_0 : Unit Root						
MZa	-2.30	-3.69	1.26	-12.0**	1.56	-5.66
MZt	-0.68	-1.24	3.60	-2.44**	3.28	-16.8*
MSB	0.38	0.34	2.87	0.20**	2.10	0.30
MPT	9.38	6.70	553	2.10**	319	4.39*
H_0 : Stationary						
KPSS	1.52***	0.18	1.93***	0.32	1.93***	0.70***

Notes: The first 4 tests are collectively from [Ng and Perron \(2001\)](#), based on H_A of a stationary model with a constant. This is inappropriate for many of the series with an obvious trend (e.g. Ind. Prod, HCPI) but best fits the constant, no trend specification of the authors' original VAR. The last test is the [Kwiatkowski et al. \(1992\)](#) test, where the null hypothesis is that there is no unit root. Both tests have non-standard distributions. Significance is reported as ***[1%] **[5%] *[10%] level.