University of Glasgow

Adam Smith Business School

WORKING PAPER SERIES

Evaluating Monetary Policy using Deviation Errors

Dimitris Korobilis and Leif Andres Thorsrud

Paper No. 2025-08
May 2025

Evaluating Monetary Policy using Deviation Errors

Dimitris Korobilis[1,3] and Leif Anders Thorsrud[*] [2,3]

[1]University of Glasgow
[2]BI Norwegian Business School
[3]Centre for Applied Macroeconomics and Commodity Prices

This version: May 19, 2025

Abstract

From the perspective of flexible inflation targeting using a simple targeting rule, this paper introduces the Monetary Policy Deviation Error (MPDE) as a novel metric for assessing central bank performance and deliberations. The MPDE captures potentially time-varying shifts in the trade-off between stabilizing inflation and supporting real economic activity. Specifically, it quantifies the gap between the intended trade-off envisioned by policymakers and the trade-off realized through actual monetary policy outcomes. Under an optimal and unbiased monetary policy strategy, the MPDE should average to zero. Nonzero deviations indicate misalignment between the central bank's stated objectives and the trade-offs actually achieved, suggesting that an alternative interest rate path would have better aligned outcomes with intentions. Applying the MPDE to evaluate the monetary policy strategies of Norges Bank and the Reserve Bank of New Zealand, we find posterior evidence supporting optimal policy alignment in the case of New Zealand.

JEL-codes: C22, E52, E58

Keywords: Monetary policy, Forecast targeting, Evaluation, Time-varying parameters

Comments from Richard Dennis, Øistein Røisland, and Tommy Sveen helped improve the quality of this paper. This work is part of the research activities at the Centre for Applied Macroeconomics and Commodity Prices (CAMP) at the BI Norwegian Business School. We thank the Reserve Bank of New Zealand for providing use with their historical macroeconomic predictions.

[*]Corresponding Author: Centre for Applied Macroeconomics and Commodity Prices, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, Norway. Email: leif.a.thorsrud@bi.no

1 Introduction

Central banks play a crucial role in modern economies, typically operating with substantial independence under politically defined mandates, such as the dual mandate of the U.S. Federal Reserve, balancing stable inflation and minimal economic fluctuations. Rigorous performance evaluations of central banks are therefore essential in democratic societies Svensson (2012). Regular external reviews, such as those in Norway and Sweden, exemplify this practice (e.g., Norges Bank Watch (2024)).

Against this background, we propose a formal evaluation metric for assessing central bank deliberations under flexible inflation targeting. Beginning with a stylized loss function, where a benchmark target rule implies opposite signs for inflation and output-gap forecasts, we leverage predicted and realized values to estimate potentially time-varying policy trade-offs. The Monetary Policy Deviation Error (MPDE) quantifies the average difference between intended and realized trade-offs, and should ideally be zero after controlling for prediction errors. Nonzero MPDE values suggest misalignment, indicating that a different interest rate path would better achieve the intended policy trade-off.

Applying this method to Norges Bank and the Reserve Bank of New Zealand, both flexible inflation-targeting institutions, we find limited evidence supporting a zero MPDE across the full sample. However, focusing on comparable periods, results from New Zealand are consistent with optimal policy alignment.

Our paper closely relates to Svensson (2012) and Argov et al. (2018). Svensson (2012) emphasizes evaluating monetary policy beyond simple inflation-target adherence, advocating metrics based on forecast targeting. We extend this approach by explicitly utilizing both ex-ante policy intentions and ex-post outcomes. Argov et al. (2018) similarly build on forecast targeting but emphasize ex-post policy efficiency via counterfactual analyses; our focus is on systematic alignment of monetary policy intentions with realized outcomes.

A closely related strand evaluates monetary policy through instrument-rule frameworks. Notable examples include Clarida et al. (2000), documenting substantial temporal variations in estimated policy rules, and Orphanides (2001), who emphasizes real-time considerations in policy evaluation. Further relevant contributions include Fendel et al. (2011), who use ex-ante data for Taylor-rule estimation, and studies by Boivin (2006), Kim and Nelson (2006), and Anderl and Caporale (2024), examining time-varying parameters in policy rules.

The paper proceeds by briefly outlining the theoretical foundations (Section 2), defining and discussing MPDE estimation (Section 3), presenting empirical findings for two case studies (Section 4), and concluding (Section 5).

2 Monetary policy trade-offs and the deviation error

The theoretical foundation for our evaluation metric is based on standard models of optimal monetary policy under flexible inflation targeting, as discussed in, e.g., Galí (2015) and Walsh (2017). The central bank's objective is to minimize the intertemporal loss function

$$\mathcal{L}_t = (1-\delta)E_t \sum_{\tau=0}^{\infty} \delta^\tau L_{t+\tau}, \tag{1}$$

where $0 < \delta \leq 1$, typically set to one (Svensson, 2003), E_t denotes expectations at time t, and L_t is the period loss:

$$L_t = \frac{1}{2}\left[(\pi_t - \pi^*)^2 + \lambda y_t^2\right], \tag{2}$$

where π_t is inflation, π^* the target, y_t the output-gap, and λ the central bank's weight on output stabilization.[1]

Under (timeless perspective) commitment, the targeting rule in the classical New Keynesian model becomes a function of the inflation-gap, the change in the output-gap, the Phillips curve slope parameter and λ. In slightly more elaborate models, targeting rules quickly become much more complex (Svensson, 2003; Woodford, 2012). For this reason, central banks often use simpler practical rules, replacing the change in output-gap with its level, as advocated by Norges Bank (Qvigstad, 2005; Røisland and Sveen, 2018; Monetary Policy Report 1, 2024). Moreover, assuming linear transmission, additive shocks, and quadratic loss, certainty equivalence allows focus on conditional mean forecasts (Svensson, 2003). This gives the practical forecast targeting rule:

$$(\pi_{t+\tau,t} - \pi^*) + \alpha y_{t+\tau,t} = 0, \tag{3}$$

where the forecasts are target variables and α depends on λ and structural parameters.

Building on Svensson (2012), who assessed policy efficiency using forecast averages for a constant and given α, we propose estimating α directly for each case using forecasts and realizations. Our metric, the MPDE, is the average deviation between intended and realized trade-offs. Persistent discrepancies between intentions and outcomes suggest systematic policy errors.

Simultaneous shocks to inflation and output can bias estimates. We discuss controlling for these using prediction errors. If policy targets expand (e.g., financial stability), the simple rule may not suffice, but tracking α remains informative for deviations from baseline optimal trade-offs (Svensson, 2012).

[1] Some central banks target employment rather than output. Using Okun's law (e.g., in Norway, $u \approx 0.3y$), an equal weight on inflation and employment implies $\lambda \approx 0.1$.

3 Estimating the Monetary Policy Deviation Error (MPDE)

Let $j \in \{i, r\}$ represent intended (forecasted) and realized outcomes, respectively. Using the targeting rule (3), we model the inflation-output gap relationship via local linear regressions:

$$\pi_{j,t+\tau} - \pi^* = \beta_{j,t} y_{j,t+\tau} + \varepsilon_{j,t}, \qquad \varepsilon_{j,t} \sim \mathcal{N}(0, h_{j,t}), \tag{4}$$

where $h_{j,t}$ captures heterogeneity in error variances, and the forecast horizon τ is discussed in Section 4. The MPDE is then the empirical estimate obtained by projecting $\hat{\beta}_{r,t} - \hat{\beta}_{i,t}$ onto a constant. Ideally, with adequate data and appropriate volatility control, a well-calibrated policy should yield an MPDE close to zero.[2]

We assume volatility h_t follows a stochastic process:

$$\log h_t = \log h_{t-1} + \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2). \tag{5}$$

where we for notation simplicity drop the j subscript. (5) captures persistent and potentially abrupt changes in the volatility of the inflation-output-gap relationship not accounted for by changes in the central bank deliberations.

The baseline specification assumes the slope coefficient β_t evolves as a random walk:

$$\beta_t = \beta_{t-1} + v_t, \qquad v_t \sim \mathcal{N}(0, q^2), \tag{6}$$

allowing flexible time variation. To enhance interpretability and address potential biases, we also consider an extended specification with observable instruments:

$$\beta_t = \beta_{t-1} + Z_t \gamma + v_t, \tag{7}$$

where Z_t is a $1 \times m$ vector of instruments capturing external shocks or forecast errors influencing policy trade-offs.

We estimate the models using Bayesian Markov Chain Monte Carlo (MCMC), drawing sequentially from conditional posterior distributions for the parameters $\beta_{1:T}$, $h_{1:T}$, q^2, σ_η^2, and γ (if applicable). Regularization of β_t dynamics employs a hierarchical horseshoe prior on q^2 (Makalic and Schmidt, 2016):

$$q^2 \mid \lambda^2, \tau^2 \sim \mathcal{IG}\left(\frac{1}{2}, \frac{1}{\lambda^2 \tau^2}\right), \quad \lambda^2 \sim \mathcal{IG}\left(\frac{1}{2}, 1\right), \quad \tau^2 \sim \mathcal{IG}\left(\frac{1}{2}, 1\right), \tag{8}$$

where λ^2 and τ^2 are local and global shrinkage parameters, respectively. This prior requires no manual hyperparameter selection and offers strong theoretical guarantees: it adapts to sparse

[2] The intermediate estimates $\hat{\beta}_{j,t}$ can also yield insights on structural parameters under stronger assumptions. For instance, prior studies such as Ilbas (2012), Givens and Salemi (2015), Dennis (2004), Lakdawala (2016), and Owyang and Ramey (2004) provide diverse estimates of λ reflecting different central bank preferences over time.

signals and shrinks small state innovations aggressively toward zero, while allowing large changes to remain unpenalized. This property is ideal for economic time series where structural breaks are infrequent but impactful. For the variance σ_η^2 in the log-volatility process, we use a weakly informative inverse-gamma prior $\sigma_\eta^2 \sim \mathcal{IG}(a_0, b_0)$ with small values $a_0 = b_0 = 0.01$ to ensure robustness. The initial state β_0 is assigned a diffuse Gaussian prior $\beta_0 \sim \mathcal{N}(0, 10^2)$. If the instrumented specification is used, we also place a standard Gaussian prior on the coefficients $\gamma \sim \mathcal{N}(0, c^2 I_m)$, with $c = 5$ by default.

Posterior simulations proceeds by alternating draws from the following blocks: 1) Draw $\beta_{1:T}$ via Forward Filtering Backward Sampling (FFBS); 2) Draw $h_{1:T}$ using mixture-of-normals approximations for the log-variance model (using, e.g., the algorithm by Kim et al. (1998)); 3) Draw q^2, λ^2, τ^2 from the conditional horseshoe hierarchy; 4) Draw σ_η^2 and γ from their conditional posteriors.

Each MCMC draw produces a full trajectory of the time-varying slope β_t, from which posterior summaries (e.g., medians and credible bands) are obtained. In total we consider 50000 draws and discard 5000 as initial burn-in steps. Finally, the MPDE is computed at each iteration by projecting the difference $\hat{\beta}_{r,t} - \hat{\beta}_{i,t}$ onto a constant and summarizing the distribution of the resulting posterior draws.

4 Two examples

To illustrate the proposed monetary policy evaluation approach, we apply it to Norges Bank and the Reserve Bank of New Zealand, both early adopters of flexible inflation targeting.

For Norway, the analysis benefits from Norges Bank's extensive publication history of inflation and output-gap forecasts. Using data from the bank's Monetary Policy reports, we estimate the intended trade-off ($\hat{\beta}_{i,t}$) from predictions averaged across one to eight-step forecasting horizons, reflecting monetary policy lag effects. Realized trade-offs ($\hat{\beta}_{r,t}$) are similarly computed using actual inflation (from Statistics Norway) and Norges Bank's historical output-gap estimates. Norges Bank's inflation target shifted from 2.5% to 2% in 2018.[3]

Estimations utilize two specifications, (6) and (7). In the latter, Norges Bank's prediction errors, aligning with traditional state-space models using forecast errors for updating state estimates, and structural shocks from a FAVAR model (Appendix B) serve as instruments.

The case of New Zealand is somewhat more complicated. The Reserve Bank of New Zealand's inflation forecasts, published since 1995, are anchored within a 1%-3% target band, but commonly

[3]The historical output gap estimates are published as real-time vintages; we use the latest vintage as the actual outcome. Norges Bank applies similar averaging logic in its own deliberations (Monetary Policy Report 1, 2024, pp. 55–59). We also tested alternative averaging horizons (e.g., first-year forecasts) and found our qualitative results unchanged.

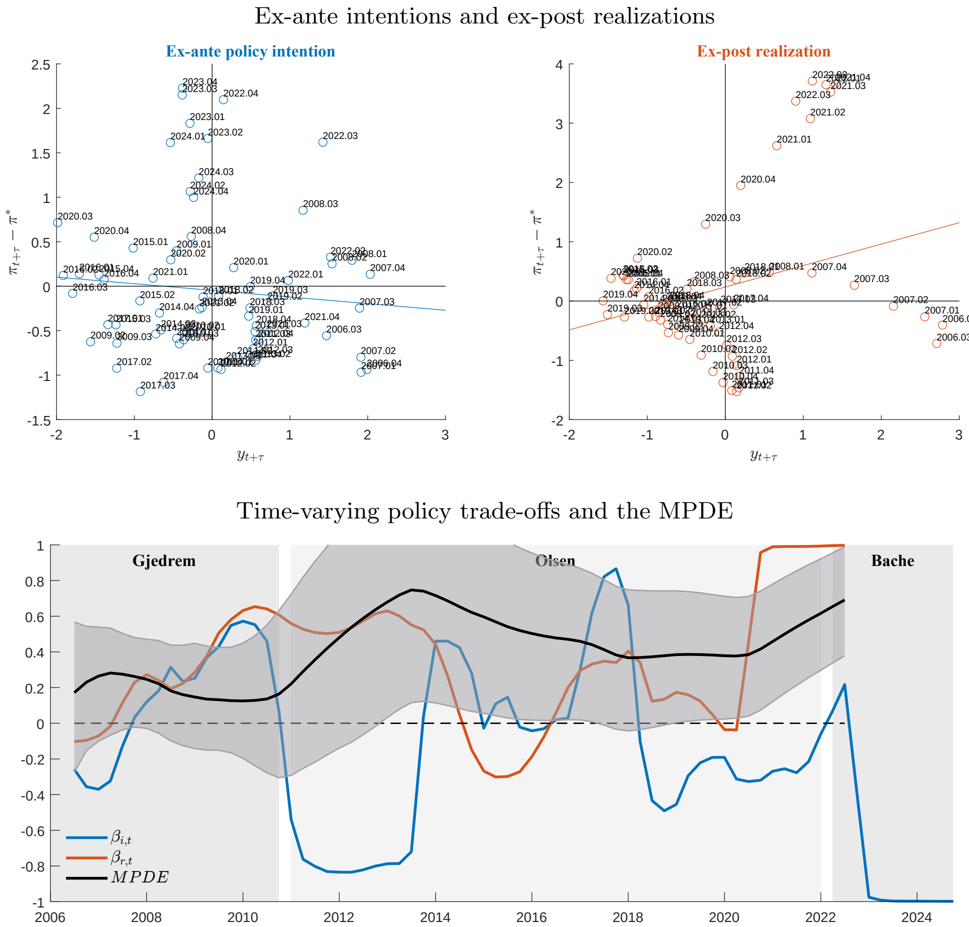

Figure 1. The upper panels show scatter plots of Norges Bank's average one- to eight-quarter-ahead inflation and output-gap forecasts and outcomes. The lower panel displays the estimated time-varying slope from (4) and (6), scaled using the hyperbolic tangent function, $\tanh(\hat{\beta}_t)$, for readability. The black line and grey band indicate the median MPDE and its 95% posterior interval. Shaded vertical areas denote different Governor terms.

analyzed around a 2% midpoint (McDermott and Williams, 2018). Moreover, as official output-gap forecasts are unavailable, we infer them from GDP trends, adopting Hamilton's eight-quarter GDP difference approach (Hamilton, 2018). Due to these complications, the Norwegian analysis remains our primary empirical investigation.

The upper-left panel in Figure 1 shows a scatter plot of Norges Bank's published inflation and output-gap forecasts from 2006 to the present. The plot indicates no clear relationship, although the slope is slightly negative. To examine potential changes over time, we estimate the intended trade-off parameter ($\hat{\beta}_{i,t}$) using models (4) and (6).

The lower panel (blue line) in Figure 1 displays the estimated $\hat{\beta}_{i,t}$, revealing considerable variability. Notably, following the 2008 financial crisis, the parameter temporarily turned positive, suggesting an emphasis on financial stability. More recently, the slope parameter has turned significantly negative, consistent with an interpretation where employment and output stability

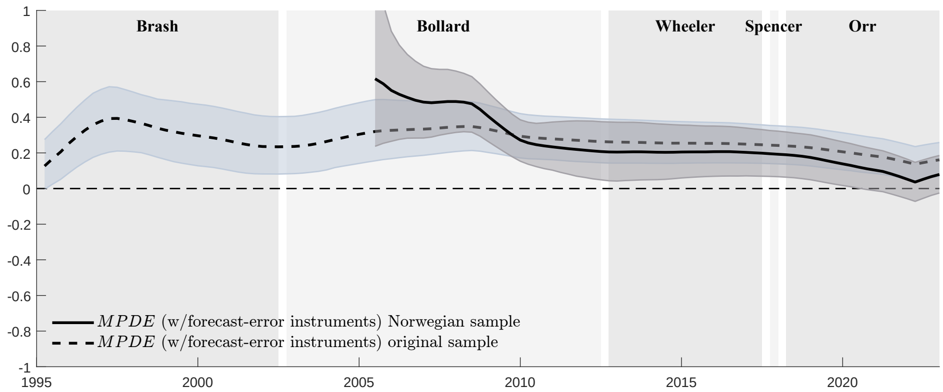

Figure 2. The graph reports New Zealand's MPDE, with $\hat{\beta}_{r,t}$ and $\hat{\beta}_{i,t}$ estimated via (4) and (7). Solid and dashed black lines indicate median estimates; shaded areas show 95% posterior intervals. Results are based on two recursive samples: starting in 1995 (as in Figure A.3) and in 2006 (to match the Norwegian case). Colored vertical bands mark different Governor periods.

are prioritized more strongly.

To assess the deviation error, we start by analyzing the realized trade-offs ($\hat{\beta}_{r,t}$) shown by the upper-right panel and the lower red line in Figure 1. While there is noticeable correlation between intended and realized trade-offs, discrepancies exist, yielding a positive MPDE as data accumulates, shown by the black solid line. Initially uncertain, the MPDE clearly becomes positive with a larger sample.

Figure A.1 (Appendix A) presents results incorporating external instruments (prediction errors and structural shocks). Instrumentation slightly reduces the MPDE magnitude but does not alter qualitative conclusions. Posterior analysis (right panel histogram) indicates minimal evidence of shocks influencing intended policy ($\hat{\beta}_{i,t}$), while realized outcomes ($\hat{\beta}_{r,t}$) are significantly affected by such shocks.

New Zealand's results, summarized in Figures A.2 and A.3 (Appendix A), largely mirror Norway's findings despite a longer sample dating back to the mid-1990s. The intended trade-off ($\beta_{i,t}$) is consistently negative, but discrepancies between intended and realized trade-offs again result in a generally positive MPDE. However, unlike Norway, New Zealand's MPDE exhibits a declining trend starting in the mid-2000s, converging toward zero, as demonstrated in Figure 2. This trend suggests increasing alignment between intended and realized policy outcomes over recent periods.

5 Conclusion

This study introduces the *Monetary Policy Deviation Error (MPDE)*, a metric designed to evaluate central bank performance under flexible inflation targeting and dual mandates. The MPDE quantifies the discrepancy between intended policy decisions - based on forecasts and established targeting rules - and the realized outcomes of the targeted variables. A nonzero MPDE indicates systematic policy errors, making it an intuitive tool for policy assessment.

We compute the MPDE using recent advancements in Bayesian estimation of time-varying parameter models. This methodology offers efficient, robust algorithms that facilitate probabilistic inference, integration of prior knowledge, and inclusion of instrumental variables affecting policy trade-offs differently ex-ante and ex-post. Additionally, these parameter estimates enable analysis of deviations from benchmark targeting rules and the stability of policy deliberations over time.

Applying the MPDE to Norges Bank and the Reserve Bank of New Zealand reveals considerable variability in policy deliberations over time. Empirically, we find stronger posterior evidence supporting a zero MPDE for New Zealand, suggesting alignment between intended and realized policy outcomes, whereas the evidence is weaker for Norway.

Although intentionally straightforward and robust, the MPDE approach can readily be extended to incorporate more complex policy rules or additional target variables if desired.

Declaration of generative AI and AI-assisted technologies in the writing process: During the preparation of this work the author(s) used ChatGPT 4.5 in order to shorten and sharpen the original text. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

Anderl, C. and G. M. Caporale (2024). Time-varying parameters in monetary policy rules: a GMM approach. *Journal of Economic Studies 51*(9), 148–176.

Argov, E., A. Binyamini, E. Borenstein, and I. Rozenshtrom (2018). Model-based ex post evaluation of monetary policy. *42th issue of the International Journal of Central Banking*.

Arias, J. E., J. F. Rubio-Ramirez, and D. F. Waggoner (2018). Inference Based on Structural Vector Autoregressions Identified With Sign and Zero Restrictions: Theory and Applications. *Econometrica 86*(2), 685–720.

Bai, J. and S. Ng (2013). Principal components estimation and identification of static factors. *Journal of Econometrics 176*(1), 18 – 29.

Bernanke, B., J. Boivin, and P. S. Eliasz (2005). Measuring the effects of monetary policy: A Factor-augmented Vector Autoregressive (FAVAR) approach. *The Quarterly Journal of Economics 120*(1), 387–422.

Boivin, J. (2006). Has US Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data. *Journal of Money, Credit and Banking*, 1149–1173.

Clarida, R., J. Gali, and M. Gertler (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory. *The Quarterly Journal of Economics 115*(1), 147–180.

Dennis, R. (2004). Inferring Policy Objectives from Economic Outcomes. *Oxford Bulletin of Economics and Statistics 66*(s1), 735–764.

Fendel, R., M. Frenkel, and J.-C. Rülke (2011). 'Ex-ante' Taylor rules – Newly discovered evidence from the G7 countries. *Journal of Macroeconomics 33*(2), 224–232.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications.* Princeton University Press.

Givens, G. E. and M. K. Salemi (2015). Inferring monetary policy objectives with a partially observed state. *Journal of Economic Dynamics and Control 52*, 190–208.

Hamilton, J. D. (2018). Why you should never use the Hodrick-Prescott filter. *Review of Economics and Statistics 100*(5), 831–843.

Ilbas, P. (2012). Revealing the preferences of the US Federal Reserve. *Journal of Applied Econometrics 27*(3), 440–473.

Kim, C.-J. and C. R. Nelson (2006). Estimation of a forward-looking monetary policy rule: A time-varying parameter model using ex post data. *Journal of Monetary Economics 53*(8), 1949–1966.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economic Studies 65*(3), 361–93.

Lakdawala, A. (2016). Changes in Federal Reserve Preferences. *Journal of Economic Dynamics and Control 70*, 124–143.

Makalic, E. and D. F. Schmidt (2016). A simple sampler for the horseshoe estimator. *IEEE Signal Processing Letters 23*(1), 179–182.

McDermott, J. and R. Williams (2018). Inflation Targeting in New Zealand: An Experience in Evolution. In *RBA Annual Conference Volume*. Reserve Bank of Australia.

Monetary Policy Report 1 (2024). Norges Bank. Technical report.

Norges Bank Watch (2024). Elisabeth Holvik and Leif Anders Thorsrud. Technical report, Centre for Monetary Economics.

Orphanides, A. (2001). Monetary policy rules based on real-time data. *American Economic Review 91*(4), 964–985.

Owyang, M. T. and G. Ramey (2004). Regime switching and monetary policy measurement. *Journal of Monetary Economics 51*(8), 1577–1597.

Qvigstad, J. F. (2005). When does an interest rate path "look good"? criteria for an appropriate future interest rate path—a practician's approach. Staff Memo No. 2005/6, Norges Bank.

Røisland, Ø. and T. Sveen (2018). Monetary policy under inflation targeting. *Occasional Papers* (53), 23–29.

Svensson, L. E. (2003). The inflation forecast and the loss function. *Central Banking, Monetary Theory and Practice: Essays in Honour of Charles Goodhart 1*, 135–152.

Svensson, L. E. O. (2012). Evaluating monetary policy. In E. Koenig, R. Leeson, and G. A. Kahn (Eds.), *From the Great Moderation to the Great Deviation: A Round-Trip Journey Based on the Work of John B. Taylor*, Chapter 10, pp. 245–274. Hoover Institution Press.

Walsh, C. E. (2017). *Monetary theory and policy.* MIT press.

Woodford, M. (2012). Forecast targeting as a monetary policy strategy: Policy rules in practice. In E. Koenig, R. L. Leeson, and G. A. Kahn (Eds.), *From the Great Moderation to the Great Deviation: A Round-Trip Journey Based on the Work of John B. Taylor*, Chapter 8, pp. 185–233. Hoover Institution Press.

Appendices for online publication

Appendix A Figures and tables

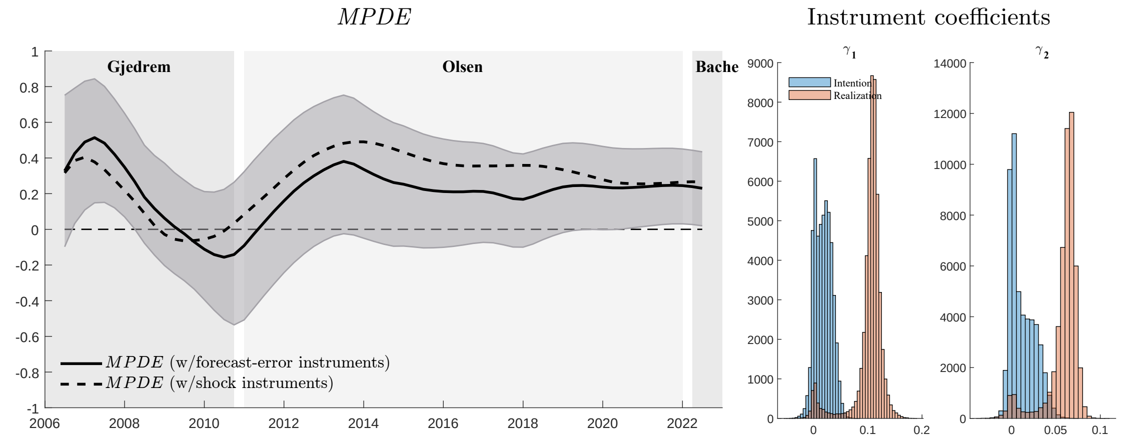

Figure A.1. The graph to the left reports the MPDE in Norway when $\hat{\beta}_{r,t}$ and $\hat{\beta}_{i,t}$ are obtained from (4) and (7). The solid and broken black lines report median estimates, while the gray shaded area illustrate the 95% posterior bands for the specification using forecast-error instruments in (7). The graph to the rights reports the posterior distribution of γ in (7) when the structural shocks are used as instruments.

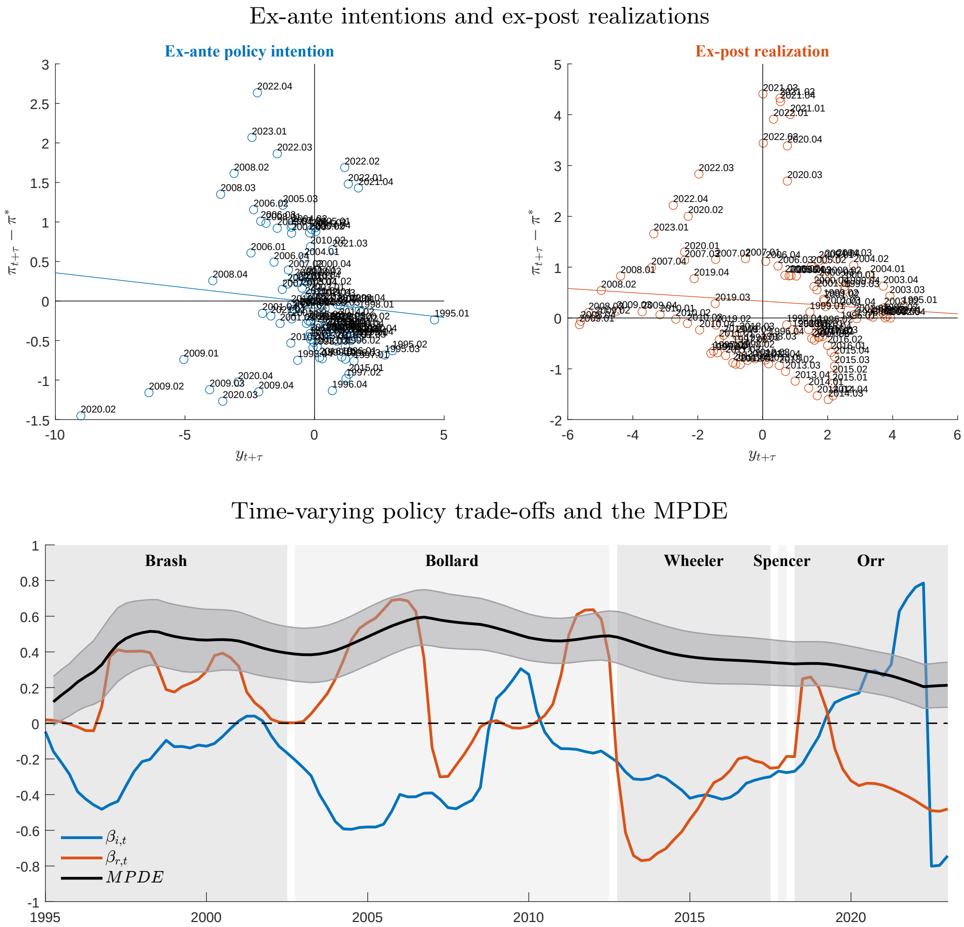

Figure A.2. The upper panels show scatter plots of the Reserve Bank of New Zealand's average one- to eight-quarter-ahead inflation and output-gap forecasts and outcomes. The lower panel displays the estimated time-varying slope from (4) and (6), scaled using the hyperbolic tangent function, $\tanh(\hat{\beta}_t)$, for readability. The black line and grey band indicate the median MPDE and its 95% posterior interval. Shaded vertical areas denote different Governor terms.

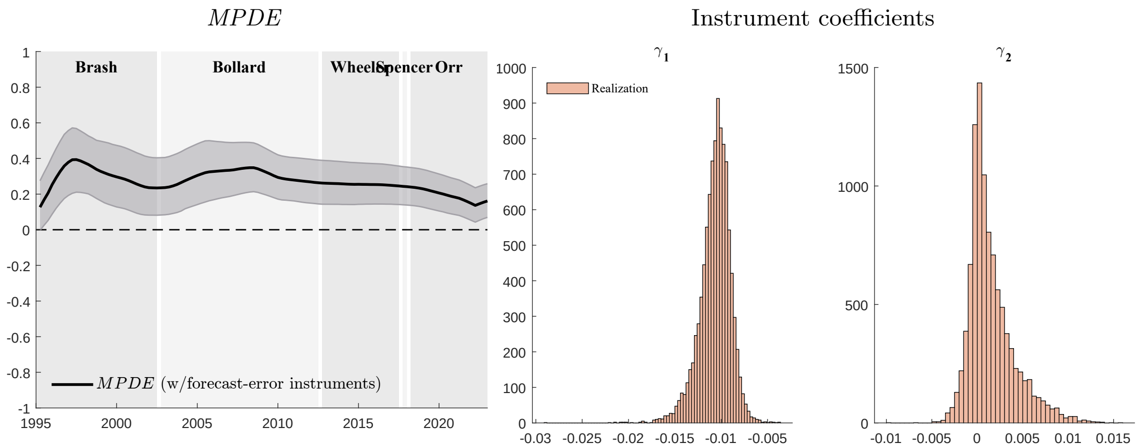

Figure A.3. The graph to the left reports the MPDE in New Zealand when $\hat{\beta}_{r,t}$ and $\hat{\beta}_{i,t}$ are obtained from (4) and (7). The solid black line report the median estimate while the gray shaded area illustrate the 95% posterior bands. The graph to the rights reports the posterior distribution of γ in (7) when the forecast-errors are used as instruments.

Appendix B FAVAR

The Factor Augmented Vector Autoregression (FAVAR) is specified assuming that two latent factors drive the common variation of year-on-year changes in real GDP, investments, consumer prices, oil prices, house prices, GDP-to-credit ratio, the spread (levels), and unemployment (levels).

The model is estimated in two-steps (Bernanke et al., 2005). In the first step we use standard Principal Components to estimate two common components. Factors are identified as activity and inflation factors using the unit identity rotation scheme proposed in Bai and Ng (2013). The identified factors are then included as endogenous variables in a standard VAR framework.

The sample used is 1987Q1-2024:Q3. We allow for four lags in the VAR and the structural shocks are identified using sign restrictions following the procedure in Arias et al. (2018).