



Research Division
Federal Reserve Bank of St. Louis
Working Paper Series



Asymmetry and Federal Reserve Forecasts

Julieta Caunedo
Riccardo DiCecio
Ivana Komunjer
and
Michael T. Owyang

Working Paper 2013-012B
<http://research.stlouisfed.org/wp/2013/2013-012.pdf>

Revised November 2015

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to Federal Reserve Bank of St. Louis Working Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors.

Asymmetry and Federal Reserve Forecasts*

Julieta Caunedo[†] Riccardo DiCecio[‡] Ivana Komunjer[§] Michael T. Owyang[¶]

November 9, 2015

Abstract

Forecasts are a central component of policy making; the Federal Reserve's forecasts are published in a document called the Greenbook. Previous studies of the Greenbook's inflation forecasts have found them to be rationalizable but asymmetric if considering particular subperiods, e.g., before and after the Volcker appointment. In these papers, forecasts are analyzed in isolation, assuming policymakers value them independently. We analyze the Greenbook forecasts in a framework in which the forecast errors are allowed to interact. We find that allowing the losses to interact makes the unemployment forecasts virtually symmetric, the output forecasts symmetric prior to the Volcker appointment, and the inflation forecasts symmetric after the onset of the Great Moderation.

Keywords: forecast rationality, loss function, Taylor rule, Greenbook forecasts, break tests.
JEL codes: C53; E47; E52.

*The paper was previously circulated as "Federal Reserve Forecasts: Asymmetry and State-Dependence". The authors thank Mengshi Lu and Chris Otrok for comments and suggestions. Diana A. Cooke, Hannah G. Shell, and Kate Vermann provided research assistance. Views expressed here are the authors' alone and do not necessarily reflect the opinions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

[†]Department of Economics, Cornell University. Email: julieta.caunedo@cornell.edu.

[‡]Research Division, Federal Reserve Bank of St. Louis. Email: dicecio@stls.frb.org.

[§]Department of Economics, University of California, San Diego. Email: komunjer@ucsd.edu.

[¶]Research Division, Federal Reserve Bank of St. Louis. Email: owyang@stls.frb.org.

1 Introduction

The forecasts that the Fed produces to conduct monetary policy are contained in the Greenbook (or, more recently, the Tealbook). Made publicly available at a 5-year lag, Greenbook forecasts have been used to study various aspects of Fed behavior, such as whether the Fed’s forecasts are rationalizable or whether the Fed has an informational advantage over the private sector.¹ The Fed’s forecasting behavior is important for our understanding of monetary policy design. For example, previous studies have argued that observed shifts in Fed’s forecasting behavior relate to changes in monetary policy objectives (see Orphanides, 2002; Capistrán, 2008).

Across the literature, forecasting behavior is modeled as a loss minimization problem in which the losses for each variable of interest are assumed symmetric and independent of other variables. If the forecaster has symmetric preferences, she experiences equal loss for forecast errors of the same size regardless of the sign. In addition, underpredicting inflation and overpredicting output growth could yield higher loss than overpredicting inflation and overpredicting output growth.

In this paper, we study the Fed’s forecasting behavior in an environment in which losses for inflation, unemployment rate, and output growth are asymmetric (see Elliott et al., 2005, 2008) and allowed to interact with each other (see Komunjer and Owyang, 2012). Ellison and Sargent (2012) argue that FOMC members forecasts might be biased (and less efficient under quadratic loss) because robust policymaking under model uncertainty may require policymakers to make “worst case scenario” forecasts. If the risks are asymmetric (i.e., inflation above target and unemployment above target are more costly than their counterparts), it makes sense that the forecasts could also be asymmetric. We also formalize a multiple break test for loss asymmetry parameters.

Consistent with Capistrán (2008), we find that the forecast loss function changes at the time of Volcker appointment as Chairman of the Fed. Unlike previous studies, we show that unemployment forecasts are virtually symmetric, output growth forecasts are symmetric previous to Volcker’s appointment; and that inflation forecast asymmetry in the post Volcker era is mostly accounted for by the disinflation period.

¹For example, Swanson (2004) finds that the Fed’s forecast errors are, on average, biased and/or correlated with the information available at the time the forecasts are made. Standard tests—e.g., those employing Theil-Mincer-Zarnowitz regressions—reject that the Fed forecasts are rationalizable (e.g., Jansen and Kishan, 1996); others (e.g., Sims, 2002; Romer and Romer, 2000) have determined that Fed forecasts are not only rationalizable but encompass private sector forecasts. Other recent papers that examine the Greenbook forecasts are Wang and Lee (2012) and Rossi and Sekhposyan (2011).

We argue these results hint to an alternative interpretation of the historical analysis of monetary policy. Orphanides (2001, 2002, 2004) argues that starting with Chairman Volcker, the Fed came to terms with the difficulties in predicting real variables. Thus, the Fed became less active in trying to stabilize the perceived output gap. Our results are broadly consistent with this previous finding, but can be viewed from a slightly different angle based on the preferences of the forecaster. We conclude that the Fed forecasts changed from emphasizing two-sided accuracy in output growth to emphasizing two-sided accuracy in inflation.

The balance of the paper is outlined as follows. Section 2 reviews the multivariate non-separable asymmetric loss function and its properties. It also provides an illustrative example of symmetric/asymmetric and separable/non-separable loss. Section 3 describes the Greenbook forecast data, the realization data, and the instruments. This section also describes the estimation of the loss function parameters, the J -test used to evaluate forecast rationality, and the break test for the asymmetry parameters. Section 4 describes the results of the forecast rationality test and the estimates of the asymmetry parameters in the forecast loss function. Section 5 discusses the implications of the previous section's findings for analyzing monetary policy. Section 6 concludes.

2 State-dependent Asymmetric Loss

The Federal Reserve's loss function associated with particular economic outcomes may change over time, depending, say, on the preferences of the Chairman. Furthermore, the overall loss may be higher when overpredicting both output and unemployment than the sum of the losses from overpredicting each of the two variables separately. We can capture this asymmetry by making three changes to the loss function underlying the Fed's forecasts: Assuming away symmetry for positive and negative errors of the same magnitude, separability across variables, and time/state invariance.²

Most models of forecaster behavior assume that predictions are generated using a loss function that increases with squared forecast errors. A simple squared loss function is directionally symmetric, separable, and time (and state) invariant. Symmetry means that positive forecast errors are

²Altering the loss function does not necessarily change the posterior distribution for the forecasts. Choosing a loss function can be thought of as selecting the "point value" of the forecast used in, for example, policymaking. As an example, if the loss function is the mean squared error, the point value of the forecast is the mean; if the loss function is the mean absolute error, the point value of the forecast is the median.

equally as costly as negative forecast errors of the same magnitude. Elliott et al. (2005) showed that the symmetry assumption can be relaxed by including a (vector of) parameter(s) that governs the degree of directional asymmetry. Forecasts generated from directionally asymmetric loss functions will be biased, as the forecaster prefers to either systematically underpredict or overpredict.

Separable loss implies that the costs of forecast errors for one variable do not depend on the forecast errors for others. This feature is clearly undesirable if we want to allow the Fed’s objectives in forecasting, say, inflation to depend on its forecasts of output. Komunjer and Owyang (2012) generalized the asymmetric loss introduced by Elliott et al. (2005) to allow for interactions between forecast errors for different variables.³ For example, underpredicting inflation may be less costly if output growth ends up higher than expected and more costly if output growth ends up lower than expected.

Finally, some studies have argued that forecasters’ behavior may change over time.⁴ While Capistrán (2008) focused on how the loss function varies across Fed chairmen, it may also be possible that the cost of some forecast errors depends jointly on the state of the economy.

2.1 The Loss Function

We adopt the multivariate, non-separable, asymmetric loss framework of Komunjer and Owyang (2012) that nests both separability and symmetry. The forecaster (in this case, the Fed) attempts to predict future values of an n vector of macroeconomic variables, \mathbf{y}_t . Define $\mathbf{f}_{t+s,t}$ as the s -period-ahead forecast of \mathbf{y}_{t+s} computed using information available at time t . The forecaster’s information set \mathcal{F}_t may include lagged values of \mathbf{y}_t in addition to other covariates used to predict \mathbf{y}_{t+s} . Hereafter, we focus on one-period-ahead forecasts and set $s = 1$. Define \mathbf{e}_{t+1} as the n vector of forecast errors for \mathbf{y}_{t+1} , i.e., $\mathbf{e}_{t+1} \equiv (\mathbf{y}_{t+1} - \mathbf{f}_{t+1,t})$. The forecaster’s problem is to construct a prediction which, conditional on information known at time t , minimizes the expected loss $E[L_p(\tau, \mathbf{e}_{t+1}) | \mathcal{F}_t]$ with:

$$L_p(\tau, \mathbf{e}) \equiv \left(\|\mathbf{e}\|_p + \tau' \mathbf{e} \right) \|\mathbf{e}\|_p^{p-1}, \quad (1)$$

³Using individual Blue Chip forecasts, Komunjer and Owyang (2012) found that the forecasters’ loss was increased with unexpectedly worse joint economic outcomes, i.e., lower-than-expected output growth, looser-than-expected monetary policy, and higher-than-expected inflation.

⁴For example, Rossi and Sekhposyan (2011) and Croushore (2012) find that the results of rationality tests can vary depending on the subsample used.

where \mathbf{e} is an n vector of forecast errors, and $\|\mathbf{e}\|_p$ denotes its p -norm.^{5,6}

The loss function (1) has $n + 1$ parameters: a shape parameter $p \in [1, \infty)$ and an asymmetry parameter n vector τ that belongs to the unit ball in \mathbb{R}^n equipped with the norm $\|\cdot\|_q$, where $1/q + 1/p = 1$, i.e., $\tau \in \{\mathbf{u} \in \mathbb{R}^n : \|\mathbf{u}\|_q \leq 1\}$. When the shape parameter $p = 1$, the loss function collapses to a multivariate tick loss sometimes used in quantile estimation.

The components (τ_1, \dots, τ_n) of the n vector τ govern the degree and direction of asymmetry associated with each of the forecasted variables. For $\tau = \mathbf{0}$, there is no asymmetry: $L_p(\tau, \mathbf{e})$ places equal weights on forecast errors in either direction. Values of τ_i ($i = 1, \dots, n$) greater than zero indicate greater loss for positive forecast errors (underprediction) in the i th component of \mathbf{y}_{t+1} . This means that, conditional on the information at time t , the distribution of forecast errors for that component will have a nonzero mean. The larger $|\tau_i|$ is, the greater the loss and the more the distribution of the corresponding forecast error will be biased away from zero.

In order to capture potential state-dependence, we allow the forecaster's asymmetry parameter τ to change. Specifically, we focus on the case in which τ can switch between two different values, τ_1 and τ_2 , at known switching times. Letting P be the sample size over which the forecast errors are observed, starting at $t = R$ and ending at $t = (R + P - 1) \equiv T$, we denote by t_1, t_2, \dots, t_S the switching times so that

$$\tau = \begin{cases} \tau_1, & \text{for } t \in [R, t_1 - 1] \cup [t_2, t_3 - 1] \cup \dots \cup [t_S, T] \\ \tau_2, & \text{for } t \in [t_1, t_2 - 1] \cup [t_3, t_4 - 1] \cup \dots \cup [t_{S-1}, t_S - 1] \end{cases} \quad (2)$$

if S is even. Similarly, if S is odd, then the last time intervals during which τ_1 and τ_2 are observed become $[t_{S-1}, t_S - 1]$ and $[t_S, T]$, respectively. Figure 1 plots a simple example in which $S = 6$ and τ is scalar.

{Figure 1: about here }

The forecasting is then performed as follows: The forecaster uses data from 1 to R to compute

⁵Komunjer and Owyang (2012) show that (i) $L_p(\tau, \cdot)$ is continuous and nonnegative on \mathbb{R}^n ; (ii) $L_p(\tau, \mathbf{e}) = 0$ if and only if $\mathbf{e} = \mathbf{0}$ and $\lim_{\|\mathbf{e}\|_p \rightarrow \infty} L_p(\tau, \mathbf{e}) = \infty$; (iii) $L_p(\tau, \cdot)$ is convex on \mathbb{R}^n .

⁶In the univariate case, this flexible loss family includes: (i) the squared loss function, $L_2(0, e) = e^2$ and (ii) the absolute deviation loss function, $L_1(0, e) = |e|$, as well as their asymmetric counterparts obtained when $\tau \neq 0$, which are called (iii) the quad-quad loss, $L_2(\tau, e)$ and (iv) the lin-lin loss, $L_1(\tau, e)$.

the first forecast of \mathbf{y}_{R+1} , $\hat{\mathbf{f}}_{R+1,R}$; the estimation window is then rolled on, and data from 2 to $R+1$ is used to compute $\hat{\mathbf{f}}_{R+2,R+1}$. Up until t_1 , the forecaster's degree of asymmetry equals τ_1 . This means that the forecasts ($\hat{\mathbf{f}}_{R+1,R}, \dots, \hat{\mathbf{f}}_{t_1,t_1-1}$) are computed by minimizing the conditional expectations of $L_p(\tau_1, \mathbf{y}_{t+1} - \mathbf{f}_{t+1,t})$ ($t = R, \dots, t_1 - 1$) given the relevant information. From time t_1 , i.e., starting with the forecast $\hat{\mathbf{f}}_{t_1+1,t_1}$, the forecaster's degree of asymmetry changes to τ_2 and remains so up to time t_2 . Hence, the forecasts ($\hat{\mathbf{f}}_{t_1+1,t_1}, \dots, \hat{\mathbf{f}}_{t_2,t_2-1}$) minimize $E [L_p(\tau_2, \mathbf{y}_{t+1} - \mathbf{f}_{t+1,t}) | \mathcal{F}_t]$, with $t = t_1, \dots, t_2 - 1$. The exercise continues until the end of the forecasting period, at $t = T$.

The forecaster's problem outlined above does not depend on the specific model used to generate the forecasts. As argued above, one can think of the loss function as affecting the forecaster's point estimates of the coefficients of whatever model chosen. This means that the model that the forecaster uses does not affect the econometrician's estimates of the forecaster's loss function parameters. Also, these estimates do not depend on the scheme—rolling or recursive—that the forecaster uses to construct the time series of forecasts.

2.2 Implications for Rationality Testing

Differences in the loss function also have implications for testing rationality. Past studies have sought to test the rational expectations hypothesis by evaluating private sector forecasts. Using the earliest methods (e.g., the Theil-Mincer-Zarnowitz regressions), a forecast was deemed rational if the forecast errors were mean zero and uncorrelated with information available at the time the forecast was made.⁷ Many, if not most, of these studies found evidence against rationality. These methods, however, implicitly assume that the forecaster's loss function is symmetric (quadratic, in most cases).

More recently, Elliott et al. (2005, 2008) showed that private sector forecasts can be rationalizable under asymmetric loss. Unlike squared or absolute error loss, asymmetric loss assigns different penalties depending on whether the realization was above or below the forecast. As an example, Elliott et al. (2005) found that private sector forecasts could be rationalizable if forecasters were attaching more loss to overpredicting output growth than underpredicting.

⁷See Theil (1958), Mincer and Zarnowitz (1969), Figlewski and Wachtel (1981), Mishkin (1981), Zarnowitz (1985), and Keane and Runkle (1990).

2.3 An Illustrative Example

{Figure 2: about here }

The consequences of non-separable asymmetric loss can be made more apparent by describing the bivariate case, $n = 2$. For simplicity of exposition, also assume $p = 2$. In this case, the loss function (1) can be rewritten as

$$L_2(\tau, \mathbf{e}) = e_1^2 + e_2^2 + (\tau_1 e_1 + \tau_2 e_2) (e_1^2 + e_2^2)^{1/2}. \quad (3)$$

The shape of the iso-loss curves—representing combinations of forecast errors corresponding to constant loss—is determined by the parameters τ_1 and τ_2 . Figure 2 shows the iso-loss curves for several different parameterizations. When $\tau_1 = \tau_2 = 0$, the loss $L_2(\tau, \mathbf{e})$ is symmetric and the iso-loss curves are perfectly circular. If either $\tau_1 \neq 0$ or $\tau_2 \neq 0$, the iso-loss curves are warped in the direction of the asymmetry. The directions in which the iso-loss curve is the closest to the origin are the ones in which the loss increases the most.

For general values of τ_1 and τ_2 , the loss (3) is non-separable in the forecast errors e_1 and e_2 . This nonseparability is an important feature as it allows the forecast errors of one variable to affect the forecasts in the other variable. In particular, even if loss is directionally symmetric in the first variable (i.e., $\tau_1 = 0$), asymmetry in the second variable (i.e., $\tau_2 \neq 0$) can produce bias in the forecasts of y_1 . This effect is revealed in the third term of (3) when $\tau_1 = 0$. The loss induced by e_1 is $e_1^2 + (\tau_2 e_2) (e_1^2 + e_2^2)^{1/2}$, which depends on both the magnitude and direction of e_2 .⁸ This dependence suggests an alternative explanation of the biased forecasts previously documented in the literature: It is not that the forecaster’s loss is asymmetric in the the first variable; rather, the loss is symmetric in e_1 but its magnitude depends on the error committed in forecasting the second variable. In this sense, our loss (3) captures the aforementioned effect that underpredicting inflation is less costly when output is higher than expected and more costly when output is lower than expected.

We can also examine the properties of the forecasts that would be produced from various

⁸Note, however, that if $\tau_1 = 0$, the direction of e_1 does not enter the loss.

parameterizations of the loss function. Suppose that \mathbf{y}_t is generated from a VAR(1):

$$\mathbf{y}_t = \mathbf{c} + \mathbf{A}\mathbf{y}_{t-1} + \varepsilon_t, \quad (4)$$

where ε_t is iid multivariate normal with zero mean and covariance matrix Σ .⁹ Based on the loss function, the one-period-ahead forecast is $\hat{\mathbf{f}}_{t+1,t} = \hat{\mathbf{c}} + \hat{\mathbf{A}}\mathbf{y}_t$, where

$$(\hat{\mathbf{c}}, \hat{\mathbf{A}}) \equiv \arg \min_{(\mathbf{c}, \mathbf{A})} P^{-1} \sum_{t=1}^P L_2(\tau, \mathbf{y}_{t+1} - \mathbf{c} - \mathbf{A}\mathbf{y}_t), \quad (5)$$

which minimizes the expected value of the loss conditional on the data and a correctly-specified VAR(1). We then construct $P = 250$ periods of bivariate forecasts for given sets of asymmetry parameters using the same generated data.

{*Figure 3: about here*}

The joint distribution of the resulting time series of forecast errors is shown in Figure 3. The directional bias of the forecast errors is evident from the plots.

3 Data and Estimation

3.1 Data

Our dataset contains three elements: the forecasts, the realizations, and the instruments used to test rationality. The sample period is September 1966 to December 2007. The forecast data are the one-quarter-ahead forecasts of the output growth rate, the inflation rate, and the unemployment rate taken from the Greenbook. The Greenbook forecasts are publicly available at a 5-year lag and vary in frequency over the sample period.¹⁰ At the beginning of the sample, the Greenbook

⁹To generate the data, we simulate $T = R + P - 1$ periods of data from the VAR(1) after discarding the first 1000 periods to remove any initial values effects. The forecaster uses a rolling window of size $R = 100$ to construct $P = 250$ one-period-ahead forecasts.

¹⁰Greenbook forecasts are prepared by the Board of Governors staff and distributed a few days before each FOMC meeting. Forecasts are constructed using both large scale econometric models and subjective assessments. In particular, the path of interest rate assumed in these forecast is subjective. Over the course of the forecast exercise, the suggested path may be revised if the econometric models or the staff suggest that such path is unlikely to be realized. The forecasts do not necessarily reflect board opinions, as they are calculated before the Federal Reserve Board makes a decisions on policy.

forecasts were available monthly; after 1981, the Greenbook was produced only for Federal Open Market Committee (FOMC) meetings. Thus, prior to 1981, we have twelve monthly vectors of observations per year; after 1981, we have eight or nine observations per year at irregular intervals. The maximum forecast horizon varies across Greenbooks in particular for the earlier dates in our sample.

The realization data is a matter of some controversy, based on one's beliefs about the veracity of data revisions. One could believe, for example, that the intent of the forecaster is to predict the value released in real time. That is, the forecaster tries to predict the value of, say, output growth for the first quarter of 1979 which is released in April 1979. On the other hand, one could argue that the initial release of the data are poor estimates and that the revisions that occur over time are closer to the truth. If the forecaster's intent is to predict this latter value, one should use the most recent vintage of the data. Similar arguments can be made for any intermediate vintage under the assumption that significant amounts of data revisions are unpredictable and outside the scope of most agents' forecasting problems. Two common approaches are taken in the literature. The first one is to use as realizations the one-year-later revision of the data. The second approach is to use the latest vintage. We report results using the latest available vintage as the realization.

In addition to changes in frequency, the Greenbook changes the forecasted output growth variable in 1992 from GNP growth to GDP growth. In order to remain consistent, when the Greenbook changes the forecasted variable, we change the realization accordingly.

Finally, the instruments used to test rationality are one lag of the forecasted series, available at the time the forecast is released.¹¹ Ideally, we would like to instrument with the information available at the time the forecast is created, but unfortunately, we cannot judge when exactly the forecast is created. We know only the time at which the Greenbook was released. We assume that any data released in the previous month was available to the forecaster. Because forecasters would not have revisions available at the time they made the forecasts, we use the previous month's vintage of the forecasted variables in our instrument set.

¹¹In principle, one could use many lags. However, Komunjer and Owyang (2012) noted that this could lead to the common "many instruments problem" and result in size distortions of the rationality test.

3.2 Estimation and Rationality Testing

Komunjer and Owyang (2012) show that the asymmetry parameters in the loss function can be estimated from the forecast errors using GMM. Using Monte Carlo evidence, they argue that the shape parameter requires a very long time series for inference—much longer than we have for the Greenbook forecasts; hence, they suggest calibrating p . We report results for the shape parameter calibrated to 2, i.e., $p = 2$. The estimation chooses the value of the asymmetry parameter that maximizes the GMM objective function derived from the first-order condition for the non-separable asymmetric loss.

The resulting value of the GMM objective function is used to construct the J -statistic for use in the rationality test. The test employs the standard test for overidentification to determine whether the forecasts are rationalizable, conditional on the instrument set (see Komunjer and Owyang, 2012, for details).

3.3 Testing for State-Dependence

To test for state dependence, we construct a Wald test of the no-change hypothesis,

$$H_0 : \tau_1 = \tau_2,$$

against the switching alternative,

$$H_1 : \tau = \begin{cases} \tau_1, & \text{for } t \in [R, t_1 - 1] \cup [t_2, t_3 - 1] \cup \dots \cup [t_S, T] \\ \tau_2, & \text{for } t \in [t_1, t_2 - 1] \cup [t_3, t_4 - 1] \cup \dots \cup [t_{S-1}, t_S - 1] \end{cases}$$

when the number of possible states, S , is even and with an analogous definition when S is odd. To describe the proposed test, split the sample of observed forecast errors into two parts,

$$D_1 \equiv [R, t_1 - 1] \cup [t_2, t_3 - 1] \cup \dots \cup [t_S, T],$$

$$D_2 \equiv [t_1, t_2 - 1] \cup [t_3, t_4 - 1] \cup \dots \cup [t_{S-1}, t_S - 1].$$

if S is even, and use an analogous definition (that swaps the last two intervals) if S is odd. The idea is then to use a partial-sample GMM estimator that uses the data in D_1 to estimate τ_1 and

the data in D_2 to estimate τ_2 . Call $\hat{\tau}_{1P}$ and $\hat{\tau}_{2P}$ the partial-sample GMM estimators of τ_1 and τ_2 obtained using the data in D_1 and D_2 , respectively. The Wald statistic for testing H_0 against H_1 then takes the form

$$W_P \equiv P(\hat{\tau}_{1P} - \hat{\tau}_{2P})' \left(\frac{1}{\pi} \left(\hat{\mathbf{B}}_P' \hat{\mathbf{S}}_{1P}^{-1} \hat{\mathbf{B}}_P \right)^{-1} + \frac{1}{1-\pi} \left(\hat{\mathbf{B}}_P' \hat{\mathbf{S}}_{2P}^{-1} \hat{\mathbf{B}}_P \right)^{-1} \right)^{-1} (\hat{\tau}_{1P} - \hat{\tau}_{2P}),$$

where P is the size of the entire forecasting sample, $\pi \in [0, 1]$ is the limit of the ratio $\frac{P_1}{P_1+P_2}$ when the sizes P_1 and P_2 of the partial samples D_1 and D_2 go to infinity, the estimators $\hat{\mathbf{S}}_{1P}$ and $\hat{\mathbf{S}}_{2P}$ are constructed using the data in D_1 and D_2 , respectively, while the estimator $\hat{\mathbf{B}}_P$ is constructed using the entire sample, i.e.,

$$\begin{aligned} \hat{\mathbf{B}}_P &\equiv \frac{1}{T} \sum_{t=R}^T \left[\|\hat{\mathbf{e}}_{t+1}\|_p^{p-1} (\mathbf{I}_n \otimes \mathbf{x}_t) + (p-1) \|\hat{\mathbf{e}}_{t+1}\|_p^{-1} (\nu_p(\hat{\mathbf{e}}_{t+1}) \otimes \mathbf{x}_t) \hat{\mathbf{e}}'_{1,t+1} \right], \\ \hat{\mathbf{S}}_{1P} &\equiv \frac{1}{\pi T} \sum_{t \in D_1} \left[\left(\hat{\mathbf{M}}_{1,t+1} \hat{\mathbf{M}}'_{1,t+1} \right) \otimes (\mathbf{x}_t \mathbf{x}'_t) \right], \\ \hat{\mathbf{M}}_{1,t+1} &\equiv \left(p \nu_p(\hat{\mathbf{e}}_{t+1}) + \hat{\tau}_{1P} \|\hat{\mathbf{e}}_{t+1}\|_p^{p-1} + (p-1) \hat{\tau}'_{1P} \hat{\mathbf{e}}_{t+1} \|\hat{\mathbf{e}}_{t+1}\|_p^{-1} \nu_p(\hat{\mathbf{e}}_{t+1}) \right), \end{aligned}$$

with analogous definitions for $\hat{\mathbf{S}}_{2P}$ and $\hat{\mathbf{M}}_{2,t+1}$. Note that we can use the entire sample to construct an estimator for \mathbf{B} because the latter does not depend on τ .¹² Under the null of no structural breaks, the partial-sample GMM estimator $(\hat{\tau}_{1P}, \hat{\tau}_{2P})$ has a known distribution, and the Wald statistic is distributed as $\chi^2(n)$, where n is the number of forecasted variables.

4 Empirical Results

In this section, we assess the directional asymmetry in the Fed's forecast loss function, estimated using data from the Greenbook forecast. In evaluating the results, it is important to keep in mind that $\tau > 0$ indicates greater loss for positive forecast errors (underprediction) and that $\tau < 0$ indicates greater loss for negative forecast errors (overprediction). Values of τ that are not statistically different from zero suggest symmetric loss. In each case, the estimated asymmetry parameter can be interpreted as the (smallest) degree of asymmetry most consistent with rationality, assuming the forecast can be rationalized; larger magnitudes of τ suggest more asymmetric preferences.

¹²Technical details for the test can be found in the Appendix.

4.1 Benchmark Full Sample Results

As a benchmark, we first estimate a set of (constant) asymmetry parameters for the full sample of Greenbook forecasts (1966:09 to 2007:12) using both separable loss—which assumes independent losses for all of the variables—and non-separable loss—which assumes that the directional losses interact. We obtain the separable loss function parameters by estimating univariate versions of the loss function, equation (1).

The top row of Table 1 compares the estimates obtained with the separable and non-separable loss functions for the full sample; the bottom row shows the values of the J -statistic used to test rationalizability. Consistent with previous studies of both Fed and private sector forecasts, the Greenbook forecasts are rationalizable if one takes into account potential directional asymmetry. This result obtains regardless of whether we assume that the loss function is separable or non-separable; separability only affects the estimates of the minimum degree of asymmetry required to obtain rationality.¹³ While the Greenbook forecasts for inflation appear unbiased (and, hence, symmetric) in the full sample, the loss function parameter most consistent with rationality for output growth exhibits a substantial degree of asymmetry. The inflation result is also consistent with the full sample result in Capistrán (2008) and is invariant to the assumption of separable versus non-separable loss.

The difference between separable and non-separable loss is highlighted by the results for the unemployment rate. The Greenbook forecasts underpredict the unemployment rate if losses are assumed to be independent—that is, the unemployment forecasts are unconditionally biased downward. However, if the magnitude and direction of the inflation and output growth forecast errors are taken into account, the loss for unemployment is essentially symmetric—the unemployment forecast errors are conditionally unbiased.

This result could imply that policymaking is conducted under the belief of some type of Phillips curve or Okun-type relationship that relates these variables. In either case, once asymmetry in inflation and output growth are accounted for, unemployment forecast errors appear symmetric.¹⁴

¹³One might be concerned that using a flexible loss results in an “automatic” finding of rationality. Komunjer and Owyang (2012), however, found that some private sector forecasters can be found to be unrationalizable even under asymmetric loss.

¹⁴While the parameter remains statistically different from zero, the value is small, implying that the relative differences in the directional losses are very small.

4.2 Changes in the Fed’s Behavior

The empirical literature on the behavior of U.S. monetary policy is rife with instances of instability attributed to any number of different root causes: Changes in the Fed leadership, differences in the Fed’s operating procedure, or variation in the volatility of macroeconomic aggregates could be thought to alter the Fed’s forecast behavior. Capistrán (2008) examines whether the forecast behavior of the Fed has changed over time, focusing primarily on the change in the Fed’s leadership when Paul Volcker became the chairman of the FOMC. His benchmark model considers the Greenbook’s inflation forecasts and splits the sample in 1979. Capistrán finds evidence that the asymmetry parameter for inflation changes across the Volcker split, with higher losses for overprediction in the pre-Volcker sample and higher losses for underprediction in the post-Volcker sample.¹⁵

If the Fed’s forecasting behavior is indeed related to its perception of policy risk, as Ellison and Sargent (2012) seem to suggest, the asymmetries may imply that the Fed perceived the risk of being above the inflation objective differently under the two regimes. Thus, Capistrán’s result suggests that the attitude of policymakers toward inflation changed significantly across the Volcker disinflation, a result consistent with other studies (e.g., Clarida et al. (2000)). However, (Orphanides, 2004, 2001) argue against changes in policymaker preferences, opting instead to attribute the differences in policymaker behavior to variation in the quality of the real-time data. In addition, studies in the forecasting literature suggest that the volatility reduction that occurred during the Great Moderation may have changed the overall forecastability of macroeconomic aggregates.

Given the potential differences in the interpretation of the loss function asymmetry parameter shift occurring at various times, we tested a few different break dates using the fixed-break-date test proposed above. In particular, we tested for breaks at (1) the Volcker appointment, (2) the Great Moderation, and (3) jointly at the Volcker appointment and at the Great Moderation.¹⁶

Table 2 contains the results of our break tests. The top panel shows the Wald statistics for a joint break in the asymmetry parameters of the full vector of variables under non-separable loss at the two single-break dates. The second panel contains the results for tests considering breaks in each individual data series, estimated with separable loss. We perform the second set of ex-

¹⁵Capistrán’s univariate analysis is equivalent to isolating the inflation parameter in a separable loss function. In this sense, our loss function can be thought of as nesting the baseline model in Capistrán.

¹⁶The Volcker appointment is assumed to occur in October 1979 and the Great Moderation break is assumed to occur in January 1984 (see McConnell and Perez-Quiros (2000)).

periments in order to determine if there are breaks in the forecast error series for the individual datasets and we use separable loss to avoid imposing a null hypothesis on the other variables. The third panel contains the Wald statistics for the joint break. In this panel, we are testing for a break at the onset of the Great Moderation, assuming the existence of a break at the Volcker appointment.

We find strong evidence of a break—in particular, at the time of the Volcker appointment. However, we also find evidence of a second break in at least two of the three variables occurring at the time of the Great Moderation, even after the sample is split at the Volcker appointment.

Based on these tests, we propose an alternative cause for the variation in the forecasting behavior. The break results suggest that the Volcker disinflation—the time between the Volcker’s appointment and the beginning of the Great Moderation—was a possibly unique period for forecasting. The change in inflation’s forecastability subsequent to the onset of the Great Moderation (and post Volcker disinflation) was highlighted in Atkeson and Ohanian (2001), who argued that simple random-walk forecasts outperform Phillips curve models during the Great Moderation period, and Stock and Watson (2007), who argue for an integrated moving-average forecast of inflation given the variation in trends over the U.S. post-War period.

Whether the cause is a decline in macroeconomic volatility or an increase in the central bank’s anti-inflation credibility in the wake of the Volcker disinflation, it seems sensible to analyze independently the forecasts before the Volcker appointment and subsequent to the onset of the Great Moderation. In addition to including two breaks, we allow the forecast errors to interact in the loss function, taking into account the possibility of the forecaster’s belief in Phillips curve-type macroeconomic relationships.

Table 3 contains the estimates of the asymmetry parameters for the loss functions with various break dates. The last column of the first two rows of Table 3 repeats Capistrán’s one-break, separable loss, inflation-only experiment with longer subsamples. Results for the Volcker break are broadly consistent with Capistrán’s results for inflation: If the sample is split in 1979:III, the Fed appears to prefer overprediction prior to Volcker and underprediction after Volcker is appointed. The other columns of Table 3 show the parameter for separable and non-separable loss. In both cases, output forecasts are essentially symmetric prior to Volcker but very conservative subsequent to the Volcker appointment. Consistent with the full sample results, unemployment rate forecasts

appear asymmetric when analyzed in isolation but appear symmetric if taken together with other forecasts.

The next two rows of Table 3 show the estimated asymmetry parameters for a single break at the onset of the Great Moderation. Our preferred specification is the two break model that treats the Volcker disinflation as a special event and, thus, consists of the first and fourth rows under the non-separable loss. The salient features of this specification are that (1) inflation forecasts were asymmetric (underpredictive) prior to the Volcker disinflation but essentially symmetric after the volatility reduction; (2) output forecasts were essentially symmetric prior to the Volcker disinflation but underpredictive during the volatility reduction; and (3) unemployment rate forecasts were always symmetric when analyzed in the context of the other forecasts, regardless of the sample.

5 Implications for How We Interpret Monetary Policy

Our results may have additional implications for the historical analysis of monetary policy. Clarida et al. (2000) argued that, prior to Volcker's appointment, the Fed's policy rule put too little emphasis on stabilizing inflation. Subsequently, under Volcker and Greenspan, the Fed responded to expected inflation more aggressively, resulting in the end of the Great Inflation. Orphanides (2004), using real-time data and Greenbook forecasts, finds that the policy rule was essentially stable since the 1960s except for the output response. Thus, Orphanides concludes that the period of high inflation was the result of inaccurate real-time data leading to inaccurate forecasts rather than a policy not aggressive enough in fighting inflation. Capistrán's result tends to support Clarida et al. (2000), claiming that the Fed became more pro-active about maintaining price stability simultaneously with an increased loss for underpredicting inflation.

On one hand, we can think of our results as augmenting Capistrán's result. He argued that the Volcker disinflation changed the preferences from underpredictive to overpredictive. Indeed, we find a change in the forecast behavior for inflation consistent with this directionality; however, we do not find that the inflation forecasts are overpredictive once we account for the forecast errors of the other variables.

Could our result be consistent with Orphanides? If the pre-Volcker Fed cared about output and had a misspecified Phillips curve, the Greenbook inflation forecasts may have been consistently be

low. Thus, prior to the Volcker break, the forecast errors are consistent with a model in which the policymaker believes in a Phillips curve-type trade-off. Subsequent to the onset of the Great Moderation, an apparent Phillips curve-type trade-off still exists but more emphasis appears to be placed on accurate inflation forecasts. In a Taylor-type policy rule, “overly optimistic” downward biased inflation forecasts can result in overly accommodative policy. The forecast biases appear even controlling for the real-time data environment, suggesting a preference-based alternative to Orphanides view from the trenches. As in Orphanides, policymakers can be forward-looking and attempting to implement optimal policy, yet still induce instability based on biases in the forecasts. In our case, however, these biases appear to be recurring, if not intentional.

6 Conclusions

We study the forecasting behavior of the Federal Reserve and attempt to understand its implications for policy design. We find evidence of systematic bias in the Greenbook forecasts of inflation and output growth used to conduct policy. Biases in the unemployment forecasts are mitigated when taken into context with the directionally-biased forecasts in the other variables.

We find an interesting pattern in the forecasts when considering subsamples prior to the Volcker appointment and subsequent to the onset of the Great Moderation. In the former subperiod, inflation forecasts are biased downward but output growth forecasts are essentially unbiased. In the latter period, the opposite is true: Inflation forecast errors are essentially mean zero and output forecasts appear conservative.

These results are broadly consistent with other studies on the change in monetary policy that occurred in the late 1970s, which has been interpreted by Orphanides as the Fed coming to terms with the difficulties in predicting real variables and by Capistrán as a change in the Fed’s forecast behavior for inflation. Our results extend and refine Capistrán’s in a multivariate setting, where the inflation forecasts are not taken in isolation. The change in the policy behavior observed in the literature appears to be simultaneous with a shift in the Fed’s forecast behavior.

References

- Andrews, D. W. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4):821–56.
- Andrews, D. W. & Fair, R. C. (1988). Inference in nonlinear econometric models with structural change. *Review of Economics Studies*, 55(6):615–40.
- Atkeson, A. & Ohanian, L. E. (2001). Are phillips curves useful for forecasting inflation? *Federal Reserve bank of Minneapolis Quarterly Review*, 25(1):2–11.
- Capistrán, C. (2008). Bias in federal reserve inflation forecasts: Is the federal reserve irrational or just cautious? *Journal of Monetary Economics*, 55(8):1415–27.
- Clarida, R. J., Galí, J., & Gertler, M. L. (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory. *Quarterly Journal of Economics*, 115(1):147–80.
- Croushore, D. (2012). Forecast bias in two dimensions. University of Richmond, unpublished manuscript.
- Elliott, G., Komunjer, I., & Timmermann, A. (2005). Estimation and testing of forecast rationality under flexible loss. *Review of Economics Studies*, 72(4):1107–25.
- Elliott, G., Komunjer, I., & Timmermann, A. (2008). Biases in macroeconomic forecasts: Irrationality or asymmetric loss? *Journal of the European Economic Association*, 6(1):122–57.
- Ellison, M. & Sargent, T. J. (2012). A defense of the fomc. *International Economic Review*, 53(4):1047–1065.
- Figlewski, S. & Wachtel, P. (1981). The formation of inflationary expectations. *The Review of Economics and Statistics*, 63(1):1–10.
- Jansen, D. W. & Kishan, R. P. (1996). An evaluation of federal reserve forecasting. *Journal of Macroeconomics*, 18(1):89–109.
- Keane, M. P. & Runkle, D. E. (1990). Testing the rationality of price forecasts: New evidence from panel data. *American Economic Review*, 80(4):714–35.

- Komunjer, I. & Owyang, M. T. (2012). Multivariate forecast evaluation and rationality testing. *Review of Economics and Statistics*, 94(4):1066–80.
- McConnell, M. M. & Perez-Quiros, G. (2000). Output fluctuations in the united states: What has changed since the early 1980's? *American Economic Review*, 90(5):1464–1476.
- Mincer, J. A. & Zarnowitz, V. (1969). The evaluation of economic forecasts. In: *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, J. A. Mincer, ed., Studies in Business Cycles, chapter 1, pages 1–46. National Bureau of Economic Research.
- Mishkin, F. S. (1981). Are market forecasts rational? *American Economic Review*, 71(3):295–306.
- Orphanides, A. (2001). Monetary policy rules based on real-time data. *American Economic Review*, 91(4):964–85.
- Orphanides, A. (2002). Monetary-policy rules and the Great Inflation. *American Economic Review*, 92(2):115–20.
- Orphanides, A. (2004). Monetary policy rules, macroeconomic stability, and inflation: A view from the trenches. *Journal of Money, Credit, and Banking*, 36(2):151–75.
- Romer, C. D. & Romer, D. H. (2000). Federal reserve information and the behavior of interest rates (digest summary). *American Economic Review*, 90(3):429–457.
- Rossi, B. & Sekhposyan, T. (2011). Forecast optimality tests in the presence of instabilities. Economic Research Initiatives at Duke (ERID) Working Paper No. 109.
- Sims, C. A. (2002). The role of models and probabilities in the monetary policy process. *Brookings Papers on Economic Activity*, 2002(2):1–62.
- Stock, J. H. & Watson, M. W. (2007). Why has us inflation become harder to forecast? *Journal of Money, Credit and banking*, 39(s1):3–33.
- Swanson, E. T. (2004). Assessing the rationality of Federal Reserve “Greenbook” forecasts. Unpublished manuscript.
- Theil, H. (1958). *Economic Forecasts and Policy*. North Holland Publishing Co.

- Wang, Y. & Lee, T.-H. (2012). Asymmetric loss in the Greenbook and the Survey of Professional Forecasters. UC Riverside, unpublished manuscript.
- West, K. D. (2006). Forecast evaluation. In: *Handbook of Economic Forecasting*, G. Elliott, C. W. Granger, & A. Timmermann, ed., volume 1, chapter 3, pages 99–134. Elsevier Science, North-Holland.
- Zarnowitz, V. (1985). Rational expectations and macroeconomic forecasts. *Journal of Business and Economic Statistics*, 3(4):293–311.

A Testing for Breaks in Parameter Estimates

We consider the setup in which, at time t , the forecaster is interested in constructing one-step-ahead predictions of an n -vector of interest, \mathbf{y}_{t+1} ($n \geq 1$). Letting $\mathbf{f}_{t+1,t}$ and $\mathbf{e}_{t+1} \equiv \mathbf{y}_{t+1} - \mathbf{f}_{t+1,t}$ denote the one-step-ahead forecast and the corresponding forecast error, we shall assume that the forecasts are constructed to minimize the expected n -variate loss, $L_p(\tau_0, \cdot)$,

$$L_p(\tau_0, \mathbf{e}) \equiv \left(\|\mathbf{e}\|_p + \tau_0' \mathbf{e} \right) \|\mathbf{e}\|_p^{p-1}, \quad (6)$$

with unknown parameter $\tau_0 \in \mathcal{B}_q^n$, where $1/p + 1/q = 1$ and $1 \leq p < \infty$ is known. Specifically, the forecaster is assumed to solve the following minimization problem:

$$\mathbf{f}_{t+1,t}^* = \arg \min_{\{\mathbf{f}_{t+1,t}\}} E [L_p(\tau_0, \mathbf{y}_{t+1} - \mathbf{f}_{t+1,t}) | \mathcal{F}_t], \quad (7)$$

where \mathcal{F}_t denotes the forecaster's information set that is informative for \mathbf{y}_{t+1} (e.g., lagged values of \mathbf{y}_t as well as other covariates), and the loss $L_p(\tau_0, \cdot)$ is defined as above. Komunjer and Owyang (2012) show that (7) is equivalent to the following forecast rationality condition:

$$E \left[p\nu_p(\mathbf{e}_{t+1}^*) + \tau_0 \|\mathbf{e}_{t+1}^*\|_p^{p-1} + (p-1)\tau_0' \mathbf{e}_{t+1}^* \frac{\nu_p(\mathbf{e}_{t+1}^*)}{\|\mathbf{e}_{t+1}^*\|_p} \middle| \mathcal{F}_t \right] = \mathbf{0}, \text{ a.s.} - P, \quad (8)$$

where $\mathbf{e}_{t+1}^* \equiv \mathbf{y}_{t+1} - \mathbf{f}_{t+1,t}^*$ is the optimal forecast error, and for any $\mathbf{u} = (u_1, \dots, u_n)$,

$$\nu_p(\mathbf{u}) \equiv c(\lim \operatorname{sgn}(u_1)|u_1|^{p-1}, \dots, \lim \operatorname{sgn}(u_n)|u_n|^{p-1})'$$

We are concerned with situations in which the forecaster's asymmetry parameter τ is allowed to change in time. We focus on the case in which τ can switch between two different values, τ_1 and τ_2 , at known switching times. Letting P be the sample size over which the forecast errors are observed, starting at $t = R$ and ending at $t = R + P - 1 \equiv T$, we denote by t_1, t_2, \dots, t_S the switching times so that

$$\tau = \begin{cases} \tau_1, & \text{for } t \in [R, t_1 - 1] \cup [t_2, t_3 - 1] \cup \dots \cup [t_S, T] \\ \tau_2, & \text{for } t \in [t_1, t_2 - 1] \cup [t_3, t_4 - 1] \cup \dots \cup [t_{S-1}, t_S - 1] \end{cases} \quad (9)$$

if S is even. Similarly, if S is odd, then the last time intervals during which τ_1 and τ_2 are observed become $[t_{S-1}, t_S - 1]$ and $[t_S, T]$, respectively.

The forecaster then uses data from 1 to R to compute the first forecast of \mathbf{y}_{R+1} , $\hat{\mathbf{f}}_{R+1,R}$; the estimation window is then rolled on, and data from 2 to $R + 1$ is used to compute $\hat{\mathbf{f}}_{R+2,R+1}$. Up until t_1 , the forecaster's degree of asymmetry equals τ_1 . From t_1 , i.e., starting with the forecast $\hat{\mathbf{f}}_{t_1+1,t_1}$, the forecaster's degree of asymmetry is allowed to change to τ_2 up to time t_2 . The exercise continues until the end of the forecasting period, at $t = T$. We shall make the following assumption:

Assumption 1 *Let P_1 and P_2 denote the lengths of the forecasting samples corresponding to regimes 1 and 2, respectively, i.e., $P_1 \equiv (t_1 - R) + (t_3 - t_2) + \dots + (T - t_S)$ and $P_2 \equiv (t_2 - t_1) + (t_4 - t_3) + \dots + (t_S - t_{S-1})$ if S is even, and $P_1 \equiv (t_1 - R) + (t_3 - t_2) + \dots + (t_S - t_{S-1})$ and $P_2 \equiv (t_2 - t_1) + (t_4 - t_3) + \dots + (T - t_S)$ if S is odd. Then, $(R, P_1, P_2) \rightarrow \infty$ and $\frac{P_1}{P_1 + P_2} \rightarrow \pi \in [0, 1]$.*

Put in words, Assumption 1 says the following: Consider the two samples of forecast errors $\hat{\mathbf{e}}_{t+1} \equiv \mathbf{y}_{t+1} - \hat{\mathbf{f}}_{t+1,t}$: those of size P_1 constructed under τ_1 , and those of size P_2 constructed under τ_2 . Then, the first requirement is that both sample sizes P_1 and P_2 go to infinity. This requirement restricts the class of switching processes we can allow for, though the restriction is fairly weak. The number of switches S can be either finite or infinite and either deterministic or random. For example, if S is finite, then Assumption 1 requires that the forecaster starts the forecasting exercise with the parameter τ_1 and ends the latter with τ_2 , thus ensuring that both P_1 and P_2 go to infinity. This situation would exclude situations in which the forecaster's parameter shifts from τ_1 to τ_2 and back somewhere within the forecasting sample, since in this case P_2 remains finite. With P_2 finite, it would be impossible to distinguish this case from the “no-switch” one in which the parameters remains equal to τ_1 throughout the forecasting sample. As already pointed out, the switching process can be either deterministic or random. In the latter case, however, it is necessary that the process be such that the ratio $P_1/(P_1 + P_2)$ converges to some known constant π in $[0, 1]$. The final requirement of Assumption 1 is that the length R of the first estimation window also goes to infinity, which is a standard requirement in the forecasting literature (see, e.g., West, 2006).

For simplicity, we shall in the first approximation ignore the forecast estimation uncertainty and thus assume that the observed forecast errors $\hat{\mathbf{e}}_{t+1}$ do not differ from their optimal counterparts \mathbf{e}_{t+1}^* . To keep the notation simple, we hereafter denote the forecast errors by \mathbf{e}_{t+1} . The objective

of the exercise is to estimate τ_1 and τ_2 from the sequences of observed forecast errors and to test whether or not they are equal. Here, the sequence of times of change t_1, \dots, t_S is assumed to be known to the forecast evaluator.

Estimation and testing is done following the ideas in (Andrews and Fair, 1988; Andrews, 1993). For this, we first need to generalize their setup, which allows only for a single shift (break). We start by splitting the sample of observed forecast errors into two parts,

$$\begin{aligned} D_1 &\equiv [R, t_1 - 1] \cup [t_2, t_3 - 1] \cup \dots \cup [t_S, T], \\ D_2 &\equiv [t_1, t_2 - 1] \cup [t_3, t_4 - 1] \cup \dots \cup [t_{S-1}, t_S - 1], \end{aligned}$$

if S is even, with an analogous definition (that swaps the last two intervals) if S is odd. The idea is then to use a partial-sample GMM (PS-GMM) estimator that uses the data in D_1 to estimate τ_1 and the data in D_2 to estimate τ_2 . This estimator is the building block of the Wald test of the asymmetry parameter constancy. To describe the estimator, suppose the true value of $\theta \equiv (\tau'_1, \tau'_2)'$ is $\theta_0 \equiv (\tau'_{10}, \tau'_{20})'$. Note that the dimension of the parameter θ equals $2n$ and that θ takes values in $\Theta \equiv \mathcal{B}_q^n \times \mathcal{B}_q^n$. For the observations $t \in D_1$, we have the population orthogonality conditions

$$E[\mathbf{g}_p(\tau_{10}; \mathbf{e}_{t+1}, \mathbf{x}_t)] = \mathbf{0}, \quad (10)$$

while for the observations $t \in D_2$ we have the population orthogonality conditions

$$E[\mathbf{g}_p(\tau_{20}; \mathbf{e}_{t+1}, \mathbf{x}_t)] = \mathbf{0}, \quad (11)$$

where \mathbf{x}_t is an \mathcal{F}_t -measurable d -vector of instruments, and $\mathbf{g}_p(\cdot; \mathbf{e}_{t+1}, \mathbf{x}_t) : \mathcal{B}_q^n \rightarrow \mathbb{R}^{nd}$ is the nd -vector-valued moment function given by

$$\mathbf{g}_p(\tau; \mathbf{e}_{t+1}, \mathbf{x}_t) \equiv \left(p\nu_p(\mathbf{e}_{t+1}) + \tau \|\mathbf{e}_{t+1}\|_p^{p-1} + (p-1)\tau' \mathbf{e}_{t+1} \|\mathbf{e}_{t+1}\|_p^{-1} \nu_p(\mathbf{e}_{t+1}) \right) \otimes \mathbf{x}_t. \quad (12)$$

The idea is then to define an estimator

$$\hat{\theta}_P \equiv (\hat{\tau}'_{1P}, \hat{\tau}'_{2P})'$$

of $\theta_0 = (\tau'_{10}, \tau'_{20})'$ that is based on the sample analogues of these orthogonality conditions. Given p and π , $1 \leq p < \infty$ and $\pi \in (0, 1)$, and given the observations $((\mathbf{x}'_R, \hat{\mathbf{e}}'_{R+1})', \dots, (\mathbf{x}'_T, \hat{\mathbf{e}}'_{T+1})')$, the PS-GMM estimator $\hat{\theta}_P$ is defined as a solution to the minimization problem:

$$\min_{\theta \in \Theta} \bar{\mathbf{g}}_p(\theta; \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t)' \mathbf{W} \bar{\mathbf{g}}_p(\theta; \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t), \quad (13)$$

where $\bar{\mathbf{g}}_p(\cdot; \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t) : \Theta \rightarrow \mathbb{R}^{2nd}$ is a $2nd$ -vector of stacked moment functions,

$$\bar{\mathbf{g}}_p(\theta; \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t) \equiv \frac{1}{P} \sum_{t \in D_1} \begin{pmatrix} \mathbf{g}_p(\tau_1; \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t) \\ \mathbf{0} \end{pmatrix} + \frac{1}{P} \sum_{t \in D_2} \begin{pmatrix} \mathbf{0} \\ \mathbf{g}_p(\tau_2; \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t) \end{pmatrix}, \quad (14)$$

$P = P_1 + P_2$ as before, and \mathbf{W} is a positive definite symmetric $2nd \times 2nd$ weighting matrix. As the definition of the moment function $\bar{\mathbf{g}}_p$ in (14) indicates, the estimator $\hat{\theta}_P$ consists of two components: $\hat{\tau}_{1P}$, which uses the data in D_1 to estimate τ_1 , and $\hat{\tau}_{2P}$, which uses the data in D_2 to estimate τ_2 . The estimator of Andrews and Fair (1988) is a special case of the above estimator obtained under a single shift (break).

We now establish the asymptotic distribution of the PS-GMM estimator $\hat{\theta}_P$ for the case of no structural change. Let $\theta_0 \equiv (\tau'_0, \tau'_0)'$ denote the true value of θ in the case when no structural change occurs. The following quantities shall be of importance in the asymptotic variance of $\hat{\theta}_P$:

$$\begin{aligned} \mathbf{M}_{t+1} &\equiv \left(p\nu_p(\mathbf{e}_{t+1}) + \tau \|\mathbf{e}_{t+1}\|_p^{p-1} + (p-1)\tau' \mathbf{e}_{t+1} \|\mathbf{e}_{t+1}\|_p^{-1} \nu_p(\mathbf{e}_{t+1}) \right) \in \mathbb{R}^n, \\ \mathbf{S} &\equiv E \left[(\mathbf{M}_{t+1} \mathbf{M}'_{t+1}) \otimes (\mathbf{x}_t \mathbf{x}'_t) \right] \in \mathbb{R}^{nd \times nd}, \\ \mathbf{B} &\equiv E \left[\|\mathbf{e}_{t+1}\|_p^{p-1} (\mathbf{I}_n \otimes \mathbf{x}_t) + (p-1) \|\mathbf{e}_{t+1}\|_p^{-1} (\nu_p(\mathbf{e}_{t+1}) \otimes \mathbf{x}_t) \mathbf{e}'_{t+1} \right] \in \mathbb{R}^{nd \times n}, \end{aligned}$$

where $\pi = P_1/P$ as defined before. We shall consider the case where the weighting matrix \mathbf{W} in (13) is optimally chosen as $\mathbf{W} = \mathbf{V}^{-1}$.

Theorem 1 *Let Assumption A1 and additional technical conditions hold. Then, given p , $1 \leq p < \infty$, and as $(R, P_1, P_2) \rightarrow \infty$, we have*

$$\sqrt{P}(\hat{\theta}_P - \theta_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Omega}) \quad \text{where} \quad \mathbf{\Omega} \equiv \begin{pmatrix} \frac{1}{\pi}(\mathbf{B}'\mathbf{S}^{-1}\mathbf{B})^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{1-\pi}(\mathbf{B}'\mathbf{S}^{-1}\mathbf{B})^{-1} \end{pmatrix}.$$

The above distribution is needed to construct a Wald test of the no change hypothesis,

$$H_0 : \tau_1 = \tau_2,$$

against the switching alternative,

$$H_1 : \tau = \begin{cases} \tau_1, & \text{for } t \in [R, t_1 - 1] \cup [t_2, t_3 - 1] \cup \dots \cup [t_S, T] \\ \tau_2, & \text{for } t \in [t_1, t_2 - 1] \cup [t_3, t_4 - 1] \cup \dots \cup [t_{S-1}, t_S - 1] \end{cases}.$$

Using the result of Theorem 1, we get the following:

Corollary 2 *Let the assumptions of Theorem 1 hold. Then, given p , $1 \leq p < \infty$, and as $(R, P_1, P_2) \rightarrow \infty$, we have*

$$\left(\frac{1}{\pi} + \frac{1}{1-\pi} \right)^{-1} P(\hat{\tau}_{1P} - \hat{\tau}_{2P})' \mathbf{B}' \mathbf{S}^{-1} \mathbf{B} (\hat{\tau}_{1P} - \hat{\tau}_{2P}) \xrightarrow{d} \chi_n^2 \quad \text{under } H_0.$$

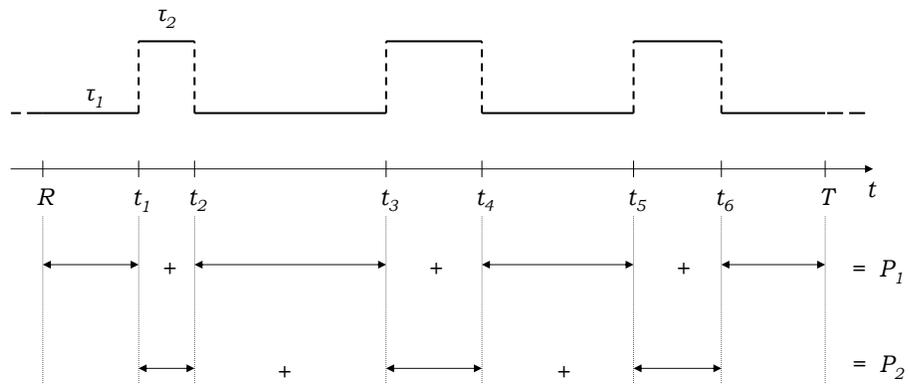
The Wald statistic for testing H_0 against H_1 takes the form

$$W_P \equiv P(\hat{\tau}_{1P} - \hat{\tau}_{2P})' \left(\frac{1}{\pi} \left(\hat{\mathbf{B}}'_P \hat{\mathbf{S}}_{1P}^{-1} \hat{\mathbf{B}}_P \right)^{-1} + \frac{1}{1-\pi} \left(\hat{\mathbf{B}}'_P \hat{\mathbf{S}}_{2P}^{-1} \hat{\mathbf{B}}_P \right)^{-1} \right)^{-1} (\hat{\tau}_{1P} - \hat{\tau}_{2P}),$$

where the estimators $\hat{\mathbf{S}}_{1P}$ and $\hat{\mathbf{S}}_{2P}$ are constructed using the data in D_1 and D_2 , respectively, while the estimator $\hat{\mathbf{B}}_P$ is constructed using the entire sample, i.e.,

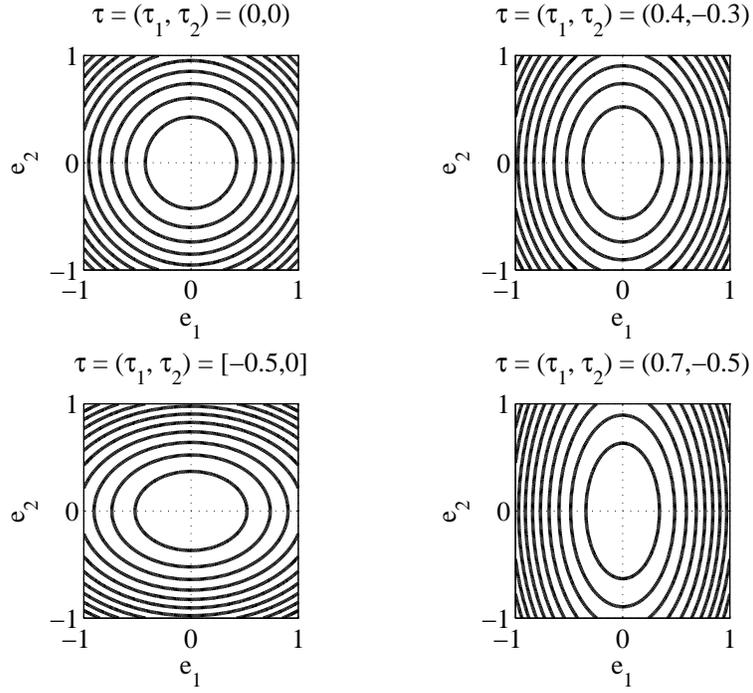
$$\begin{aligned} \hat{\mathbf{B}}_P &\equiv \frac{1}{T} \sum_{t=R}^T \left[\|\hat{\mathbf{e}}_{t+1}\|_p^{p-1} (\mathbf{I}_n \otimes \mathbf{x}_t) + (p-1) \|\hat{\mathbf{e}}_{t+1}\|_p^{-1} (\nu_p(\hat{\mathbf{e}}_{t+1}) \otimes \mathbf{x}_t) \hat{\mathbf{e}}'_{1,t+1} \right], \\ \hat{\mathbf{S}}_{1P} &\equiv \frac{1}{\pi T} \sum_{t \in D_1} \left[\left(\hat{\mathbf{M}}_{1,t+1} \hat{\mathbf{M}}'_{1,t+1} \right) \otimes (\mathbf{x}_t \mathbf{x}'_t) \right], \\ \hat{\mathbf{M}}_{1,t+1} &\equiv \left(p \nu_p(\hat{\mathbf{e}}_{t+1}) + \hat{\tau}_{1P} \|\hat{\mathbf{e}}_{t+1}\|_p^{p-1} + (p-1) \hat{\tau}'_{1P} \hat{\mathbf{e}}_{t+1} \|\hat{\mathbf{e}}_{t+1}\|_p^{-1} \nu_p(\hat{\mathbf{e}}_{t+1}) \right), \end{aligned}$$

with analogous definitions for $\hat{\mathbf{S}}_{2P}$ and $\hat{\mathbf{M}}_{2,t+1}$. Note that we can use the entire sample to construct an estimator for \mathbf{B} because the latter does not depend on τ .

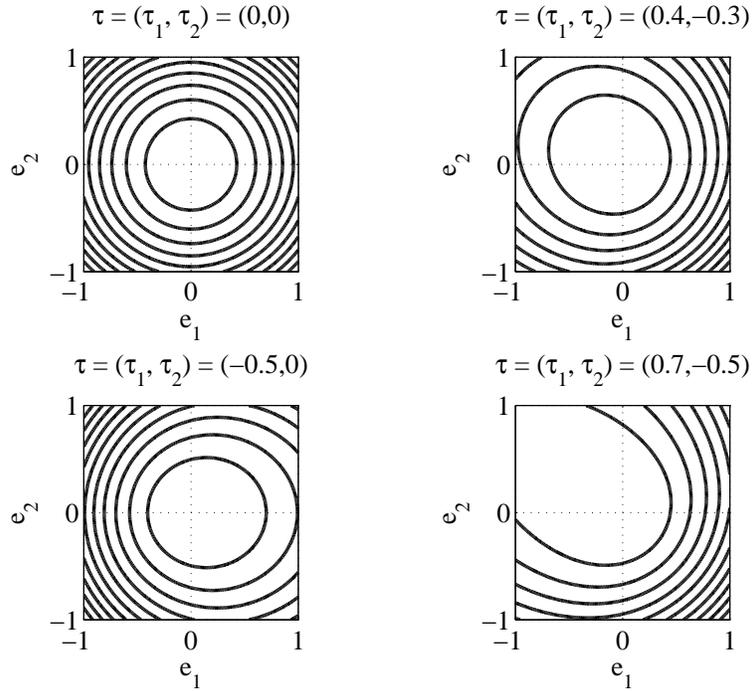


$$P_1 + P_2 = P, \quad R + P = T$$

Figure 1: Switching times, $S = 6$

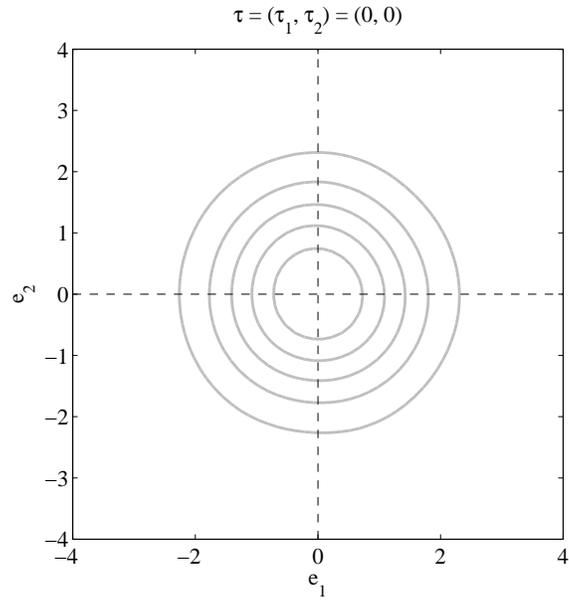


(a) Separable Loss.

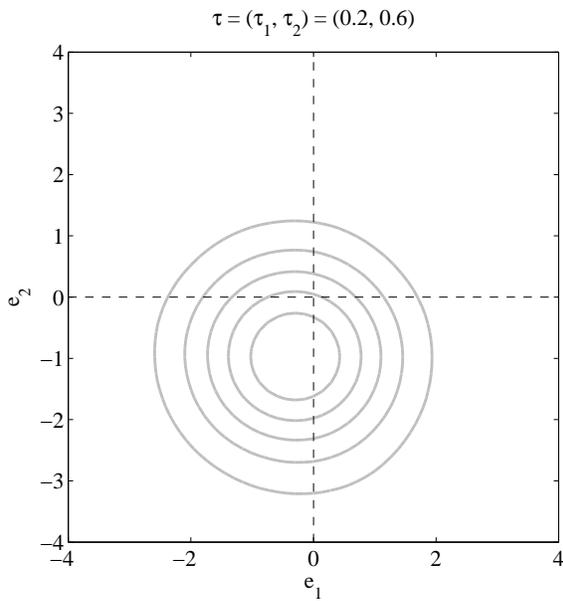


(b) Nonseparable Loss.

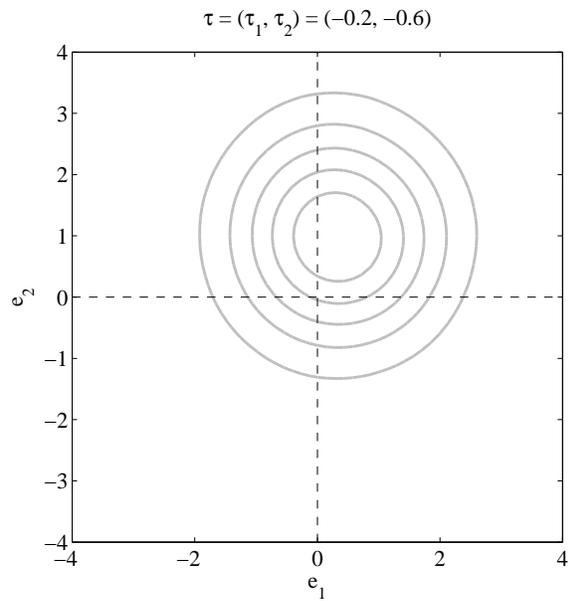
Figure 2: Iso-loss contours.



(a) Symmetric loss.



(b) Higher losses from overestimation.



(c) Higher losses from underestimation.

Figure 3: Contour plots of the distribution of simulated forecast errors (kernel density smoothed).

	Nonseparable Loss			Separable Loss		
	UR	Growth	Inflation	UR	Growth	Inflation
τ_n	0.03*** (0.01)	-0.32*** (0.08)	-0.03 (0.05)	0.28*** (0.1)	-0.34*** (0.08)	0.01 (0.08)
J -stat [†]	9.53 (p-value, 0.14)			9.44 (p-value, 0.15)		

Table 1: Full sample estimates, $p = 2$.

Notes: The table shows the asymmetry parameters for the unemployment rate (UR), output growth, and inflation for the Greenbook forecasts using the full sample of data, 1966:09-2005:12. Standard errors are shown in parentheses. p -values correspond to the 99th percentile (***) , 95th percentile (**), and 90th percentile (*). The J -stat tests the null of rationalizability of the forecasts. p -values of the J -test correspond to a χ^2 distribution with six degrees of freedom.

Nonseparable Loss		
	Volcker	Great Moderation
Joint	23.03 (0.00)	7.80 (0.05)
Separable Loss		
	Volcker	Great Moderation
Output growth	13.99 (0.00)	7.50 (0.01)
Inflation	17.32 (0.00)	1.96 (0.16)
Unempl. rate	0.30 (0.58)	0.05 (0.83)
Second Break		
	Volcker	Great Moderation
Output growth	-	0.00 (0.98)
Inflation	-	30.64 (0.00)
Unempl. rate	-	5.88 (0.02)

Table 2: Loss function subsample analysis: Wald tests.

Notes: The table contains the Wald statistic and the p-values for the tests of breaks in the asymmetry parameters. The top panel tests for a joint break in each variable's asymmetry parameter. The middle panel tests for breaks in each of the three variables assuming separable loss. The bottom panel tests for a one-time break at the Great Moderation date using only data from the post-Volcker period. The Volcker break is October 1979. The Great Moderation break is January 1984.

	Pre-Volcker	Post-Volcker	Pre-Great Moderation	Post-Great Moderation
Nonseparable Loss				
Unemp. Rate	0.03* (0.01)	0.03** (0.01)	0.03** (0.01)	0.03* (0.01)
Output Growth	-0.01 (0.13)	-0.54*** (0.07)	-0.17 (0.11)	-0.52*** (0.08)
Inflation	-0.24*** (0.07)	0.13** (0.06)	-0.08 (0.06)	0.03 (0.08)
J-Stat	4.57 (0.60)	6.25 (0.40)	6.15 (0.41)	5.98 (0.43)
Separable Loss				
Unemp. Rate	0.24* (0.13)	0.33*** (0.12)	0.28** (0.12)	0.24* (0.13)
Output Growth	0.02 (0.13)	-0.55*** (0.07)	-0.17 (0.11)	-0.55*** (0.09)
Inflation	-0.34*** (0.11)	0.28*** (0.10)	-0.09 (0.10)	0.15 (0.13)
J-Stat	4.55 (0.60)	6.39 (0.38)	6.3 (0.39)	5.44 (0.49)

Table 3: Subsample analysis: τ_n estimates, $p = 2$.

Notes: The asymmetry parameter is asymptotically normal. p-values correspond to the 99th percentile (***), 95th percentile (**), and 90th percentile (*). The J-stat tests the null of nonrationalizability of the forecasts. p-values of the J-test correspond to a 2 distribution with sixdegrees of freedom. The asymmetry parameter is asymptotically normal. p-values correspond to the 99th percentile (***), 95th percentile (**), and 90th percentile (*).