
ELSEVIER
VOLUME 27, NUMBER 2
APRIL–JUNE 2011
ISSN 0169-2070
international journal of forecasting
International Institute of Forecasters

International Journal of Forecasting 30 (2014) 235–245

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Asymmetric loss in the Greenbook and the Survey of Professional Forecasters

Yiyao Wang [a], Tae-Hwy Lee [b,*]

[a] Booth School of Business, The University of Chicago, Chicago, IL 60637, USA
[b] Department of Economics, University of California, Riverside, CA 92521, USA

ARTICLE INFO

Keywords:
Greenbook
SPF
Asymmetric loss
Forecast rationality
Real output growth forecasts
Inflation rate forecasts
Real time data
Revised data

ABSTRACT

This paper examines the forecast rationality of the Greenbook and the Survey of Professional Forecasters (SPF) under asymmetric loss functions, using the method proposed by Elliott, Komunjer, and Timmermann (2005) with a rolling window strategy. Over rolling periods, the degree and direction of the asymmetry in forecast loss functions are time-varying. While rationality under symmetric loss is often rejected, forecast rationality under asymmetric loss fails to be rejected over nearly all rolling periods. Besides, real output growth is consistently under-predicted in the 1990s, and the inflation rate is consistently over-predicted in the 1980s and 1990s. In general, inflation forecasts, especially for long horizons, exhibit greater levels of loss asymmetry in magnitude and frequency. The loss asymmetry of real output growth forecasts is more pronounced when the last revised vintage data are used than when the real-time vintage is used. All of these results hold for both the Greenbook and SPF forecasts. The results are also similar with the use of different sets of instrumental variables for estimating the asymmetric loss and testing for forecast rationality.

1. Introduction

Some forecasting research takes the forecast producer's perspective, and deals with issues about constructing optimal forecasts under a given loss function. Other research takes the forecast user's perspective and deals with issues about testing the forecast rationality of a forecast producer. Since various forecasts of economic variables are made available to the public (e.g., government forecasts such as the Greenbook forecasts, private-sector forecasts such as those from the Survey of Professional Forecasters (SPF), and forecasts of international organizations such as IMF and OECD), the forecast user's perspective of testing forecast rationality becomes an important research topic.

Forecast rationality has been often tested under a given symmetric loss function. For example, under symmetric squared error loss, the rational forecast is unbiased, the single-period horizon forecast errors are serially uncorrelated, and the unconditional variance of the forecast error is a non-decreasing function of the forecast horizon (cf. Diebold & Lopez, 1996; Granger & Newbold, 1986; Patton & Timmermann, 2007a). These properties make it convenient to test for forecast rationality. Romer and Romer (2000) support the unbiasedness of Greenbook inflation forecasts by applying the Mincer and Zarnowitz (1969) test, while Capistran (2008) rejects the rationality of Greenbook inflation forecasts in some sub-periods. Rossi (2012) further discusses the time-varying nature of the unbiasedness result. Patton and Timmermann (2012) develop a new testing strategy based on the properties of optimal multi-horizon forecasts under the symmetric squared error loss, and reject the rationality of Greenbook multi-horizon forecasts for quarter-over-quarter changes in GDP,

* Corresponding author.
E-mail addresses: yiyaowang@chicagobooth.edu (Y. Wang), taelee@ucr.edu (T.-H. Lee).

0169-2070/$ – see front matter © 2013 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijforecast.2013.07.017

the GDP deflator and CPI. Croushore (2010) also rejects the forecast rationality of SPF inflation forecasts under the symmetric squared error loss.

Forecast rationality can be tested under a specific asymmetric loss function. If a forecast can only be rationalized by assuming asymmetric loss, the rejection of forecast rationality under symmetric loss will probably be caused by a false assumption of symmetric loss, rather than a lack of forecast rationality. One way to deal with asymmetric loss is to develop new properties to test for forecast rationality, cf. Patton and Timmermann (2007b). Another way is to assume a particular asymmetric loss function. For example, Patton and Timmermann (2007a) use the Linex loss function of Varian (1975) with a regime switching model to develop properties which test for rationality.

Forecast rationality can also be tested without the assumption of any particular loss function. Elliott, Komunjer, and Timmermann (2005, EKT henceforth) propose a method for estimating the loss function parameter from a large class of asymmetric loss functions, and develop a test for rationality under the estimated loss. This novel method sets few restrictions on the data generating process, and estimates the asymmetric loss parameter using public information only. EKT discover that, when allowing asymmetric loss, forecast rationality is rarely rejected in IMF and OECD forecasts of budget deficits for G7 countries. Following EKT, some extensions and the consequent empirical work have been added. Patton and Timmermann (2007b) introduce a more flexible asymmetric loss function which depends not only on forecast errors but also on the realized target value. Under such kinds of loss functions, Greenbook real output forecasts can be rationalized perfectly. Elliott, Komunjer, and Timmermann (2008) apply the EKT method to SPF forecasts of nominal and real output growth and find that "only a modest degree of asymmetry is required for the survey expectations to be consistent with rationality". Komunjer and Owyang (2012) extend the EKT framework with a multivariate nonseparable asymmetric loss function to test the forecast rationality of multivariate forecasts jointly.

Moreover, several papers have provided evidence that the degree of asymmetry estimated by EKT may not be constant over time. Capistran (2008) separates the full data period of the Greenbook inflation forecasts into two subperiods, pre-Volcker and post-Volcker (Paul Volcker was the FRB chairman from 1979 to 1987), and conducts the EKT method separately for each period. He finds a significant difference in the direction of asymmetry between the two periods, and suggests that "the cost of having inflation above an implicit time-varying target was larger than the cost of having inflation below it for the post-Volcker period, and that the opposite was true for the pre-Volcker era". Patton and Timmermann (2007b) show that the level of loss asymmetry in Greenbook real output growth forecasts at times of recession is much higher than that in a high GDP period.

Following the above literature, this paper further studies the possible time-varying nature of the loss function asymmetry and forecast rationality of Greenbook and SPF by applying the EKT method. First, we adopt a rolling window strategy and find that the potential level of loss asymmetry is time-varying. In rolling periods, rationality under symmetric loss is often rejected, but rationality under asymmetric loss is rarely rejected. This confirms the finding of EKT (2005) and Capistran (2008) that the asymmetry in loss functions reflects forecasters' cautious risk attitudes, rather than forecast irrationality. Second, both real output growth forecasts and inflation rate forecasts are included in our analysis. We find that real output growth forecasts in the 1990s are produced with a loss function that punishes over-prediction more than under-prediction, which leads to consistent under-prediction, while the inflation forecasts in the 1980s and 1990s are produced with an asymmetric loss in the opposite direction. In general, inflation forecasts, especially for long horizons, embrace a higher level of asymmetry in magnitude and frequency. Third, we use different vintages for the realized data, and discover that the loss asymmetry of real output growth forecasts is more pronounced when the last revised vintage data are used than when the real-time vintage is used. Fourth, both the Greenbook and SPF forecasts are included in our analysis of forecast rationality under asymmetric loss. The results suggest that there is a clear similarity between the time-varying loss asymmetry patterns of the two forecasts, which may be valuable information, because the SPF is available five years before the publication of the Greenbook forecasts. Fifth, these results are also similar to different sets of instrumental variables for the estimation of the asymmetric loss, and for tests of forecast rationality.

The rest of the paper is organized as follows. Section 2 describes Greenbook and SPF forecast values, as well as revised and real-time vintages for realized values. Section 3 provides a brief review of the EKT (2005) method. Section 4 presents a list of the main empirical findings. Finally, Section 5 concludes.

2. Data

Let f_{t+h} be the h-step-ahead forecast of y_{t+h}. We obtain the forecast value f_{t+h} from either the Greenbook or the SPF. Both the forecasts and realized values are at a quarterly frequency at the source. The realized values y_{t+h} of real output growth or the inflation rate are obtained from either the real-time data vintages or the most recently revised data vintage.

The Greenbook forecasts are produced by the Board of Governors of the Federal Reserve Board before each meeting of the Federal Open Market Committee, starting from the fourth quarter of 1965. As per Sims (2002), the forecast project is broken down into forecasts of several sub-variables for different Fed experts to model and forecast. These sub-variable forecasts are then combined by a primary macroeconomic model known as FRB/US, in order to generate original Greenbook forecasts of several economic variables for further judgmental adjustment and feedback. Since the Greenbook may have a substantial influence on policy making, it is made available to the public with a lag of five years. Although the FOMC meetings occur more than four times a year, with a varying time schedule, we use quarterly Greenbook forecasts, to be compatible with the

SPF frequency. In our analysis, forecasts of both real output growth in the GNP/GDP index and the inflation rate in the GNP/GDP price index are included.[1] Greenbook covers a variety of horizons, up to 8 or 9 quarters, but we only consider horizons of $h = 1, 4$, representing short- and long-horizon forecasts. The data period is different for each horizon: the one-quarter-ahead forecasts are available from 1968Q3 to 2005Q4, and the four-quarter-ahead forecasts are available from 1974Q2 to 2005Q4. The longest possible data period for each horizon is used in order to take advantage of all of the information.

In addition to the Greenbook forecast, we also use real output growth forecasts and inflation rate forecasts with horizons of $h = 1, 4$ from the Survey of Professional Forecasters (SPF), a well-known survey forecast which is currently organized by the Federal Reserve Bank of Philadelphia. The SPF is a set of many forecasts, mostly made by professional forecasters from business companies or Wall Street. In the SPF, one-quarter-ahead forecasts are available from 1968Q4 to 2012Q1 at the time of writing this paper, and four-quarter-ahead forecasts are available from 1974Q4 to 2012Q1.[2] Notably, the forecasts for real output growth actually started in 1981Q3, before which the forecast values were computed from the forecasts of nominal GNP and the GNP price deflator index. Unlike the Greenbook, recent SPF forecasts have been available without a lag, forming a valuable source of forecasts in the absence of Greenbook forecasts. Because the SPF consists of forecasts from many professional forecasters, there is a certain degree of dispersion among the different individual forecasters, a topic which was researched by Capistran and Timmermann (2009) and Patton and Timmermann (2010). In this paper, the median response of the forecast survey is used as a consensus of the SPF forecasts.

Both "real-time data" and "revised data" are used in this paper as the realized value y_{t+h}. "Real-time data" is the name for the second revision of statistical data, which was systematically proposed and discussed by Croushore and Stark (2001) and is now provided by the Real-Time Data Research Center of Federal Reserve Bank of Philadelphia. Data on the current quarter's growth rate are computed in the following quarter's vintage. The second revision data are comparatively complete, with a lower statistical error than initial revision data, and are also closer to what the forecasters are forecasting than later revisions, because "this series does not include the rebenchmarking and definitional changes that occur in the annual and quinquennial revisions" (Romer & Romer, 2000). On the other side, "revised data" are computed with the most recent vintage (2011Q4 vintage in this paper), which should represent the economic status of past periods well from the standpoint of today. Indexes of real GNP/GDP and GNP/GDP prices are made available on the website of the Federal Reserve Bank of Philadelphia. Applying the transformation of $y_t = 400 \times \ln(x_t/x_{t-1})$, with x being either the real GNP/GDP or GNP/GDP price index, we obtain data on real output growth and the inflation rate, both real-time and revised, which are compatible in format with the forecast values.

[1] Both variables are annualized percentage values of the quarter-over-quarter growth rate. The forecasts were for GNP from 1965 to 1991 and for GDP from 1992 onward.

[2] However, the realized value is only available up to and including 2011Q3, so the last few forecasts of SPF are not used in our empirical study—the data periods used in this paper should be 1968Q4 to 2011Q2 (SPF 1-step-ahead) and 1974Q4 to 2010Q3 (SPF 4-step-ahead).

3. The EKT method

The EKT method is based on the linear GMM framework. Here, we review the method briefly. Let $f_{t+h} = \theta' W_t$ be the h-step-ahead linear forecast of y_{t+h} conditional on the information set $\mathcal{F}_t$ at time t, where θ is an unknown k-vector of parameters and W_t is a k-vector of variables that are $\mathcal{F}_t$-measurable. A generalization to a nonlinear forecast $f_{t+h} = \theta(W_t)$ and a nonlinear GMM is straightforward. EKT use a flexible class of loss function

$$L(\alpha, p, \theta) = [\alpha + (1 - 2\alpha) \cdot 1(\varepsilon_{t+h} < 0)] \cdot |\varepsilon_{t+h}|^p, \quad (1)$$

where $\alpha \in (0, 1)$ and p is a positive integer. We present the results with $p = 2$. $1(\cdot)$ is the indicator function, which equals 1 when its argument is true and 0 otherwise, and $\varepsilon_{t+h} = y_{t+h} - f_{t+h}$ is the forecast error. The parameter α indicates the level of loss asymmetry. When $\alpha < 0.5$, forecasters tend to punish over-prediction more and create a bias towards under-prediction. On the other hand, when $\alpha > 0.5$, forecasters tend to punish under-prediction more, and create a bias toward over-prediction. For a given (α_0, p_0), the forecast $f^*_{t+h} = \theta^{*\prime} W_t$ is rational if $\theta^* = \arg\min_{\theta \in \Theta} E\{L(\alpha_0, p_0, \theta)\}$ solves the following first order condition (FOC):

$$E\left(W_t \cdot \left[1\left(y_{t+h} - f^*_{t+h} < 0\right) - \alpha_0\right] \times \left|y_{t+h} - f^*_{t+h}\right|^{p_0 - 1}\right) = 0. \quad (2)$$

Given the forecast $\hat{f}_{t+h} = \hat{\theta}_t' W_t$ provided by its producer (such as FRB or SPF), the forecast user wishes to estimate the α of the producer's loss function for a given value of p_0. In order to back out α, the FOC should hold if and only if $\alpha = \alpha_0$. EKT prove that $\theta^*(\alpha_0)$ is a continuous differentiable one-to-one mapping from $(0, 1)$ to Θ. This indicates that a different level of loss asymmetry, α, will yield a different forecast, and a different forecast will reveal a different level of loss asymmetry. W_t is the information set that the forecaster knows at time t. $k = \dim(W_t)$ is typically very large, and not all of the information in W_t is accessible to the users.

Instruments for W_t: To estimate α_0, we only need $m = \dim(V_t)$ (≥ 1) instruments V_t, satisfying the condition

$$A(\alpha_0) \equiv E\left(V_t \cdot \left[1\left(y_{t+h} - f^*_{t+h} < 0\right) - \alpha_0\right] \times \left|y_{t+h} - f^*_{t+h}\right|^{p_0 - 1}\right) = 0. \quad (3)$$

We have considered several sets of instrumental variables V_t, including the following sets: $V_t = 1$ consisting of a constant, $V_t = (1\ y_{t-1})'$, $V_t = \left(1\ \hat{f}_{t+h}\right)'$, $V_t = \left(1\ \hat{\varepsilon}_{t-1}\right)'$, $V_t = \left(1\ y_{t-1}\ \hat{f}_{t+h}\right)'$, $V_t = \left(1\ y_{t-1}\ \hat{\varepsilon}_{t-1}\right)'$, and $V_t = \left(1\ y_{t-1}\right.$

$y_{t-1}^2)'$. As $V_t = W_t$ is optimal, $V_t = f_{t+h}^* = \theta^{*\prime} W_t$ would be a good instrument if θ^* were known. In reality, θ^* and f_{t+h}^* are not known, but it is possible that $V_t = \hat{f}_{t+h} = \hat{\theta}_t' W_t$ could still be a good instrument, as $\hat{\theta}_t$ is measurable with respect to W_t (and so is $\hat{f}_{t+h}$). Hence, $V_t = \left(1\, \hat{f}_{t+h}\right)'$ and $V_t = \left(1\, y_{t-1}\, \hat{f}_{t+h}\right)'$ are valid instruments, satisfying the moment condition in Eq. (3). If a rational forecast $f_{t+h}(\theta, W_t)$ is nonlinear in W_t, the above FOC can be modified to replace W_t with the gradient of $f_{t+h}(\theta, W_t)$ with respect to θ, evaluated at (θ^*, W_t), as was noted by EKT (p. 1110). For example, if $f_{t+h}(\theta, W_t) = \theta_0 + \theta_1 y_{t-1} + \theta_2 y_{t-1}^2$, then the instrument can be taken as $V_t = \left(1\, y_{t-1}\, y_{t-1}^2\right)'$. Overall, these different choices of the instruments obtain similar results. The instruments with $\hat{f}_{t+h}$ yield somewhat unstable estimates of α, with some large changes in $\hat{\alpha}_{T,\tau}$ over the rolling windows, especially when $h = 4$, which may be due to the small sample size $T = 40$ for the estimation windows. We present the empirical results with $V_t = (1 y_{t-1})'$ in the next section, with brief remarks on the results from using the other sets of instrumental variables (which are available in a supplementary appendix on the authors' websites).

Backing out the asymmetric parameter: Let $B = E(V_t \cdot |y_{t+h} - f_{t+h}^*|^{p_0-1})$ and $C = E(V_t \cdot 1(y_{t+h} - f_{t+h}^* < 0) \cdot |y_{t+h} - f_{t+h}^*|^{p_0-1})$. Write $A(\alpha) \equiv C - \alpha B$. The orthogonality condition in Eq. (3) can be solved by minimizing $Q(\alpha) = A(\alpha)' S^{-1} A(\alpha)$, which yields $\alpha_0 = (B'S^{-1}C)/(B'S^{-1}B)$. This can be estimated by $\hat{\alpha}_{T,\tau} = (\hat{B}_{T,\tau}' \hat{S}_{T,\tau}^{-1} \hat{C}_{T,\tau})/(\hat{B}_{T,\tau}' \hat{S}_{T,\tau}^{-1} \hat{B}_{T,\tau})$, where $\hat{B}_{T,\tau} = \frac{1}{T}\sum_{t=\tau}^{T+\tau-1} V_t \cdot |y_{t+h} - \hat{f}_{t+h}|^{p_0-1}$, $\hat{C}_{T,\tau} = \frac{1}{T}\sum_{t=\tau}^{T+\tau-1} V_t \cdot 1(y_{t+h} - \hat{f}_{t+h} < 0) \cdot |y_{t+h} - \hat{f}_{t+h}|^{p_0-1}$, and $\hat{S}_{T,\tau}$ is a consistent estimate of S. The estimate $\hat{\alpha}_{T,\tau}$ depends on the estimation of S, which in turn depends on α_0, so we iterate the estimation of α and the Newey and West (1987) estimator of S. EKT establish consistency and asymptotic normality: $T^{\frac{1}{2}}(\hat{\alpha}_{T,\tau} - \alpha_0) \xrightarrow{d} N(0, (B'S^{-1}B)^{-1})$.

Rolling windows: We apply the rolling window strategy to the examination of the possible time-varying behavior of the asymmetric loss parameter and to the analysis of its asymmetry preferences in different periods of time. Let τ denote the beginning of a rolling estimation sample and T denote the size of the rolling estimation sample, with the index t being used to denote the time point at which an h-step-ahead forecast $\hat{f}_{t+h}$ is made. Let $n + h$ be the total number of periods available. We use the first rolling sample $\left\{V_t, y_{t+h}, \hat{f}_{t+h}\right\}_{t=2}^{t=T+1}$ to compute $\hat{\alpha}_{T,1}$, the τth rolling sample $\left\{V_t, y_{t+h}, \hat{f}_{t+h}\right\}_{t=\tau+1}^{t=T+\tau}$ to compute $\hat{\alpha}_{T,\tau}$, and the last rolling sample $\left\{V_t, y_{t+h}, \hat{f}_{t+h}\right\}_{t=n-T+1}^{t=n}$ to compute $\hat{\alpha}_{T,n-T}$. Hence, we obtain a total of $n - T$ asymmetric loss parameter estimates $\hat{\alpha}_{T,\tau}$, where the first index T denotes the size of the rolling window and the second index $\tau = 1, \ldots, n - T$ denotes the time when the rolling window *begins*. In presenting $\left\{\hat{\alpha}_{T,\tau}\right\}$ in Section 4, we use figures whose horizontal axis is $\tau = 1, \ldots, n - T$, with a fixed $T = 40$ (40 quarters in a 10-year window) and with n being determined by the maximum length of the forecasts available in the Greenbook and SPF. In applying the rolling window scheme, we move the period forward one quarter at a time and compute the corresponding estimate of the asymmetric parameter, until we reach the end of the data period.

Rationality test under asymmetric loss: We construct the J-statistic for the rationality test under asymmetric loss,

$$
\begin{aligned}
J_{T,\tau}(\hat{\alpha}_{T,\tau}) &= T \times \hat{Q}_{T,\tau}(\hat{\alpha}_{T,\tau}) \\
&= T \times \hat{A}_{T,\tau}(\hat{\alpha}_{T,\tau})' \hat{S}_{T,\tau}^{-1}(\hat{\alpha}_{T,\tau}) \hat{A}_{T,\tau}(\hat{\alpha}_{T,\tau}),
\end{aligned}
\tag{4}
$$

where $\hat{A}_{T,\tau}(\hat{\alpha}_{T,\tau}) = \hat{C}_{T,\tau} - \hat{\alpha}_{T,\tau}\hat{B}_{T,\tau}$. This J-test for overidentification checks whether the orthogonality condition in Eq. (3) holds for $\alpha = \hat{\alpha}_{T,\tau}$; that is, it checks whether forecast rationality holds for $\alpha = \hat{\alpha}_{T,\tau}$. If the number of orthogonal conditions (the number of instruments) is larger than one, the $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ statistic follows the asymptotic chi-square distribution with $(m - 1)$ degrees of freedom, χ_{m-1}^2. A large value of the $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ statistic would indicate the rejection of the forecast rationality condition in Eq. (3) when $\alpha = \hat{\alpha}_{T,\tau}$.

Rationality test under symmetry: On the condition that the $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ statistic has not rejected the orthogonality condition in Eq. (3), the $J_{T,\tau}(0.5)$ statistic with $\alpha = 0.5$ can be used to test for rationality under loss symmetry. The rejection of this test would point to the rejection of loss symmetry if the rationality has not been rejected by $J_{T,\tau}(\hat{\alpha}_{T,\tau})$. However, the conditional distribution of the restricted statistic $J_{T,\tau}(0.5)$ conditional on $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ is hard to obtain. Hence, the statistic $J_{T,\tau}(0.5)$ is taken as a *joint* test statistic for forecast rationality and loss symmetry. In this case, $J_{T,\tau}(0.5)$ follows the chi-square distribution χ_m^2 asymptotically (with degrees of freedom equal to m, as no parameter has been estimated). In the next section, however, we will interpret $J_{T,\tau}(0.5)$ loosely as a test for loss symmetry. The test for loss symmetry may also be conducted by computing 95% confidence intervals of α_0 using the asymptotic normality result of $\hat{\alpha}_{T,\tau}$.

It may be noted that allowing for a time-varying asymmetry parameter does not necessarily imply that forecast rationality will be satisfied. When the symmetry is in fact not true ($\alpha_0 \neq 0.5$), it is true that the power of the rationality test $J_{T,\tau}(0.5)$ with assumed symmetry would be higher, rejecting forecast rationality more often than the forecast rationality test $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ allowing time-varying asymmetry. However, $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ will still have power against departures from forecast rationality, as it is designed to. On the other hand, when the symmetry is in fact true ($\alpha_0 = 0.5$), the power of the rationality test $J_{T,\tau}(0.5)$ with assumed symmetry could be lower, rejecting forecast rationality less often than the forecast rationality test $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ which allows time-varying asymmetry.[3]

[3] $J_{T,\tau}(0.5)$ can be computed using $\hat{\alpha}_{T,\tau} = 0.5$ in Eq. (4). However, we use $\hat{S}_{T,\tau}(\hat{\alpha}_{T,\tau})$ instead of $\hat{S}_{T,\tau}(0.5)$ because it can improve the finite sample power, as was noted by EKT (p. 1114).

Table 1
The full sample results of estimates and tests.

$\left[\hat{f}_{t+h}, y_{t+h}\right]$	Real output growth		Inflation	
	$h = 1$	$h = 4$	$h = 1$	$h = 4$
[Greenbook, real-time]				
$\hat{\alpha}_{T,\tau}$ $(se(\hat{\alpha}_{T,\tau}))$	0.572 (0.064)	0.584 (0.075)	0.530 (0.078)	0.609 (0.095)
$J_{T,\tau}(0.5)$ (p-value)	1.700 (0.427)	1.287 (0.526)	0.148 (0.929)	1.344 (0.511)
$J_{T,\tau}(\hat{\alpha}_{T,\tau})$ (p-value)	0.429 (0.513)	0.009 (0.925)	0.001 (0.970)	0.001 (0.971)
[Greenbook, revised]				
$\hat{\alpha}_{T,\tau}$ $(se(\hat{\alpha}_{T,\tau}))$	0.438 (0.058)	0.465 (0.072)	0.454 (0.088)	0.567 (0.109)
$J_{T,\tau}(0.5)$ (p-value)	1.249 (0.536)	0.253 (0.880)	0.445 (0.801)	0.434 (0.805)
$J_{T,\tau}(\hat{\alpha}_{T,\tau})$ (p-value)	0.056 (0.813)	0.022 (0.882)	0.171 (0.679)	0.056 (0.813)
[SPF, real-time]				
$\hat{\alpha}_{T,\tau}$ $(se(\hat{\alpha}_{T,\tau}))$	0.594 (0.059)	0.651 (0.070)	0.570 (0.067)	0.749 (0.065)
$J_{T,\tau}(0.5)$ (p-value)	2.873 (0.238)	5.200 (0.074)	1.279 (0.528)	14.009 (0.001)
$J_{T,\tau}(\hat{\alpha}_{T,\tau})$ (p-value)	0.320 (0.572)	0.568 (0.451)	0.190 (0.663)	0.483 (0.487)
[SPF, revised]				
$\hat{\alpha}_{T,\tau}$ $(se(\hat{\alpha}_{T,\tau}))$	0.484 (0.059)	0.568 (0.081)	0.499 (0.082)	0.699 (0.081)
$J_{T,\tau}(0.5)$ (p-value)	0.367 (0.833)	0.821 (0.663)	0.729 (0.729)	6.281 (0.043)
$J_{T,\tau}(\hat{\alpha}_{T,\tau})$ (p-value)	0.292 (0.589)	0.121 (0.728)	0.632 (0.427)	0.166 (0.684)

Notes: The results in this table are for the full sample period with $T = n$ and $\tau = 1$. The asymptotic standard error $se(\hat{\alpha}_{T,\tau})$ of $\hat{\alpha}_{T,\tau}$, and the p-values of the J-statistics, are reported in parentheses. $V_t = (1\ y_{t-1})'$ is used.

4. Empirical analysis

We apply the EKT method to both the full data period and rolling windows of size $T = 40$. Three statistics are computed: the asymmetric loss parameter estimate $\hat{\alpha}_{T,\tau}$ with standard errors, the p-value of the statistic $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ for forecast rationality under asymmetric loss, and the p-value of the joint test statistic $J_{T,\tau}(0.5)$ for forecast rationality and loss symmetry. The results for the full data period are reported in Table 1, and the results for the rolling windows are reported in Figs. 1–4.

Table 1 presents results for the full data period. The estimates $\hat{\alpha}_{T,\tau}$ are followed by their standard errors $se(\hat{\alpha}_{T,\tau})$ in brackets, and $J_{T,\tau}(0.5)$ and $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ are followed by p-values in brackets. For the forecast value $\hat{f}_{t+h}$, both the Greenbook and SPF forecasts (of real output growth and the inflation rate) are included. For the realized values of y_{t+h}, both real-time and revised data are used. Results for both 1-step-ahead (one-quarter-ahead) forecasts and 4-step-ahead (one-year-ahead) forecasts are presented. Most of the estimates of the asymmetric loss parameter are near 0.5. The p-values of $J_{T,\tau}(0.5)$, which tests rationality under symmetry, are all larger than 0.05, except for those for the $h = 4$ SPF inflation forecast. This indicates that in most cases for the full data period, the asymmetry in the forecast loss functions of Greenbook and SPF are not statistically significant.

Figs. 1–4 present the results for the rolling sub-periods. Figs. 1 and 2 present results for each of the rolling windows, indexed by τ for real output growth forecasts with $h = 1$ and $h = 4$, respectively, while Figs. 3 and 4 do the same for the inflation rate forecasts. Each figure has a 4×2 array of time series graphs over τ. Panels (a), (c), (e) and (g) report "Estimates" $\hat{\alpha}_{T,\tau}$ with 95% asymptotic confidence intervals at each τ. Panels (b), (d), (f) and (h) report the p-values of two "Tests", $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ and $J_{T,\tau}(0.5)$. The p-values of the statistic $J_{T,\tau}(\hat{\alpha}_{T,\tau})$ are in red and the p-values of the statistic $J_{T,\tau}(0.5)$ are in black. Each figure has four rows, with different pairs of forecast and realized values.[4] In the figures, the time stamp on the horizontal axis, denoted by τ, is the *beginning* of the 10-year rolling windows, as was discussed in the previous section.

The asymmetric loss parameter estimates $\hat{\alpha}_{T,\tau}$ are apparently time-varying in both degree and direction, as the first column of the four figures shows. Unlike the results for the full data period in Table 1, forecast rationality under symmetry is often rejected in rolling periods, since the 95% confidence intervals often do not include $\alpha = 0.5$ (dashed line) and the p-values of $J_{T,\tau}(0.5)$ in Column 2 are often smaller than 0.05 (dashed line), especially for inflation forecasts. While the values of $\hat{\alpha}_{T,\tau}$ are mostly below 0.5 for real output growth forecasts and mostly above 0.5 for inflation rate forecasts, occasional crossings are also observed.

The second column of all four figures shows that the p-values of $J_{T,\tau}\left(\hat{\alpha}_{T,\tau}\right)$ for forecast rationality under asymmetric loss rarely go below 0.05 and are much larger than their counterparts, $J_{T,\tau}(0.5)$, even though the joint hypothesis of symmetry and rationality for many rolling periods is rejected due to small p-values of $J_{T,\tau}(0.5)$. Rejections of forecast rationality occur only in a very few rolling windows. This implies that forecast rationality under asymmetric loss is not rejected for either the real output growth or inflation rate forecasts, for most rolling periods, at both the one quarter and one year horizons, and for both real-time data and revised data. This finding confirms that the rejection of the joint test of symmetry and rationality is probably due to the rejection of symmetry rather than the rejection of forecast rationality.

There is a substantial difference between the patterns of time-varying asymmetry of the real output growth and

[4] Row 1 uses the forecasts $\hat{f}_{t+h}$ from Greenbook and the realized data y_{t+h} from real-time vintages. Row 2 uses the forecasts $\hat{f}_{t+h}$ from Greenbook and the realized data y_{t+h} from the last revised data vintage. Row 3 uses the forecasts $\hat{f}_{t+h}$ from SPF and the realized data y_{t+h} from real-time vintages. Row 4 uses the forecasts $\hat{f}_{t+h}$ from SPF and the realized data y_{t+h} from revised data vintages.

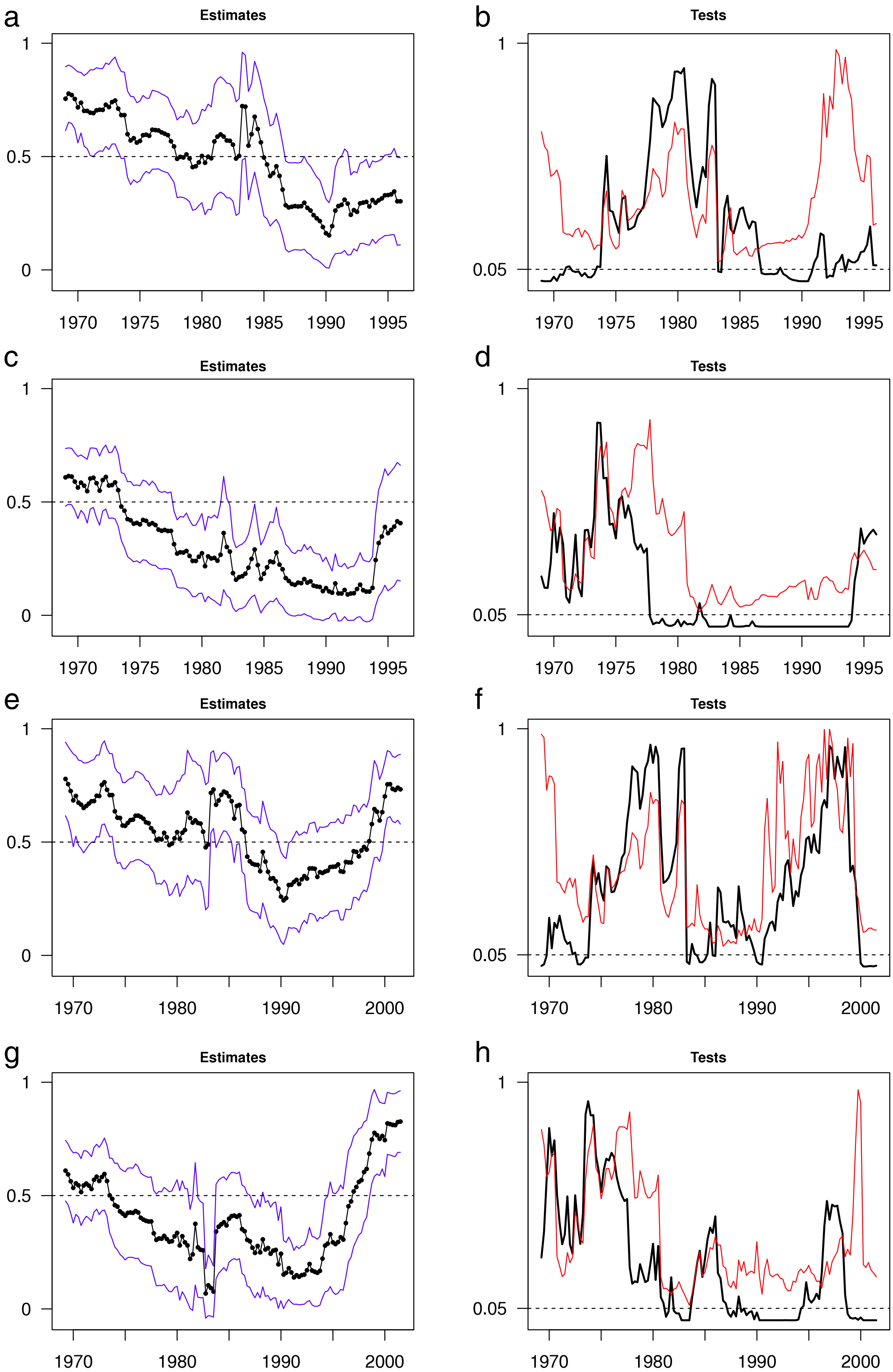

Fig. 1. **Output growth forecasts** ($h = 1$). Estimates and tests with 1-step-ahead real output growth forecasts: (a) and (b) with Greenbook and real-time data; (c) and (d) with Greenbook and revised data; (e) and (f) with SPF and real-time data; (g) and (h) with SPF and revised data. $V_t = (1\ y_{t-1})'$ is used.

inflation rate forecasts. For real output growth in Figs. 1 and 2, the level of loss asymmetry is generally mild, except in those periods around $\tau = 1990$Q1, when the estimates of $\hat{\alpha}_{T,\tau}$ are significantly below 0.5 in Column 1. This suggests that an asymmetric loss function which punished over-prediction more than under-prediction was used in the 1990s. For the inflation rate, there is a much greater degree of asymmetry in terms of both magnitude and frequency.

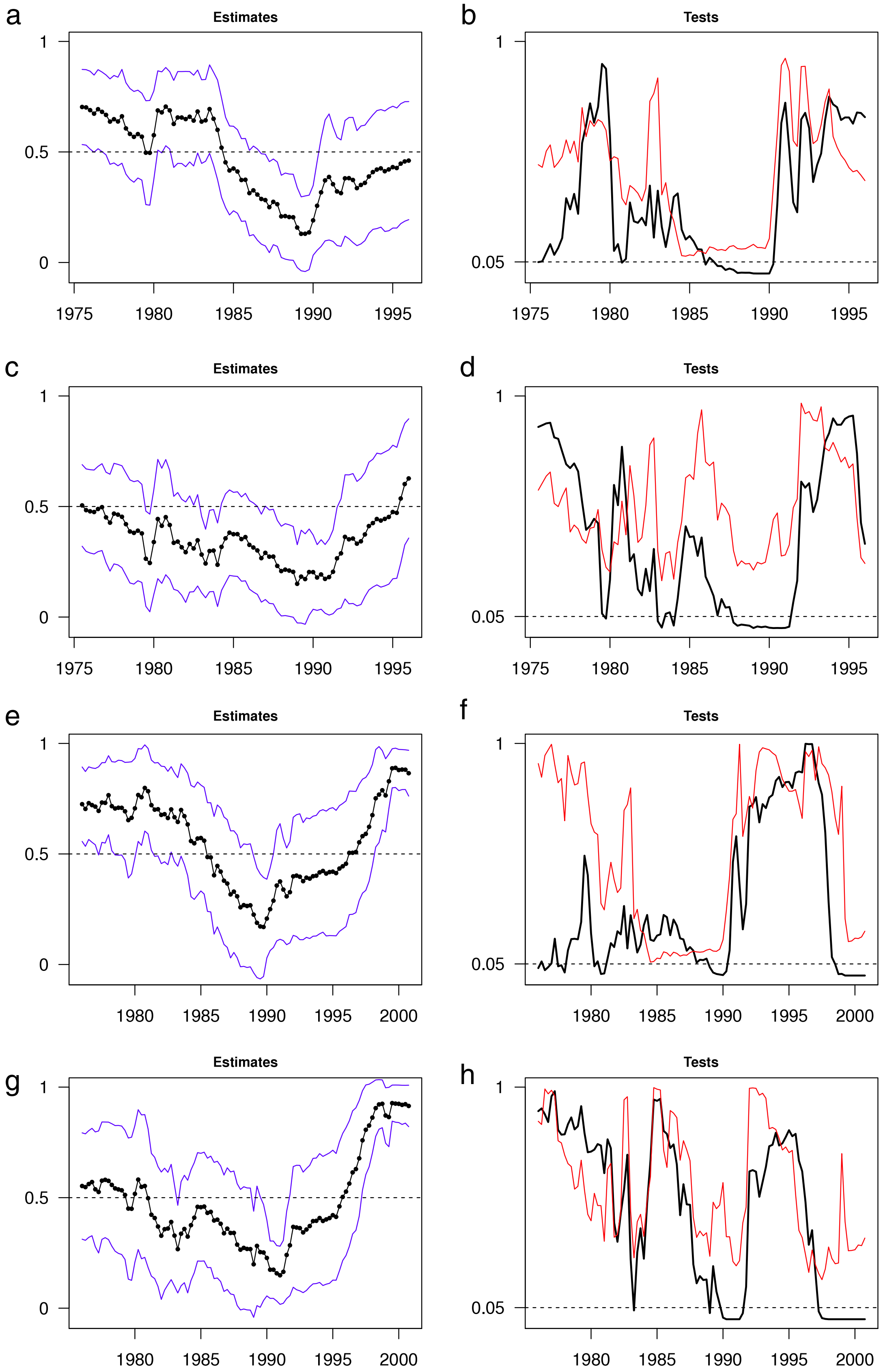

Fig. 2. Output growth forecasts ($h = 4$). Notes. Estimates and tests with 4-step-ahead real output growth forecasts: (a) and (b) with Greenbook and real-time data; (c) and (d) with Greenbook and revised data; (e) and (f) with SPF and real-time data; (g) and (h) with SPF and revised data. $V_t = (1\, y_{t-1})'$ is used.

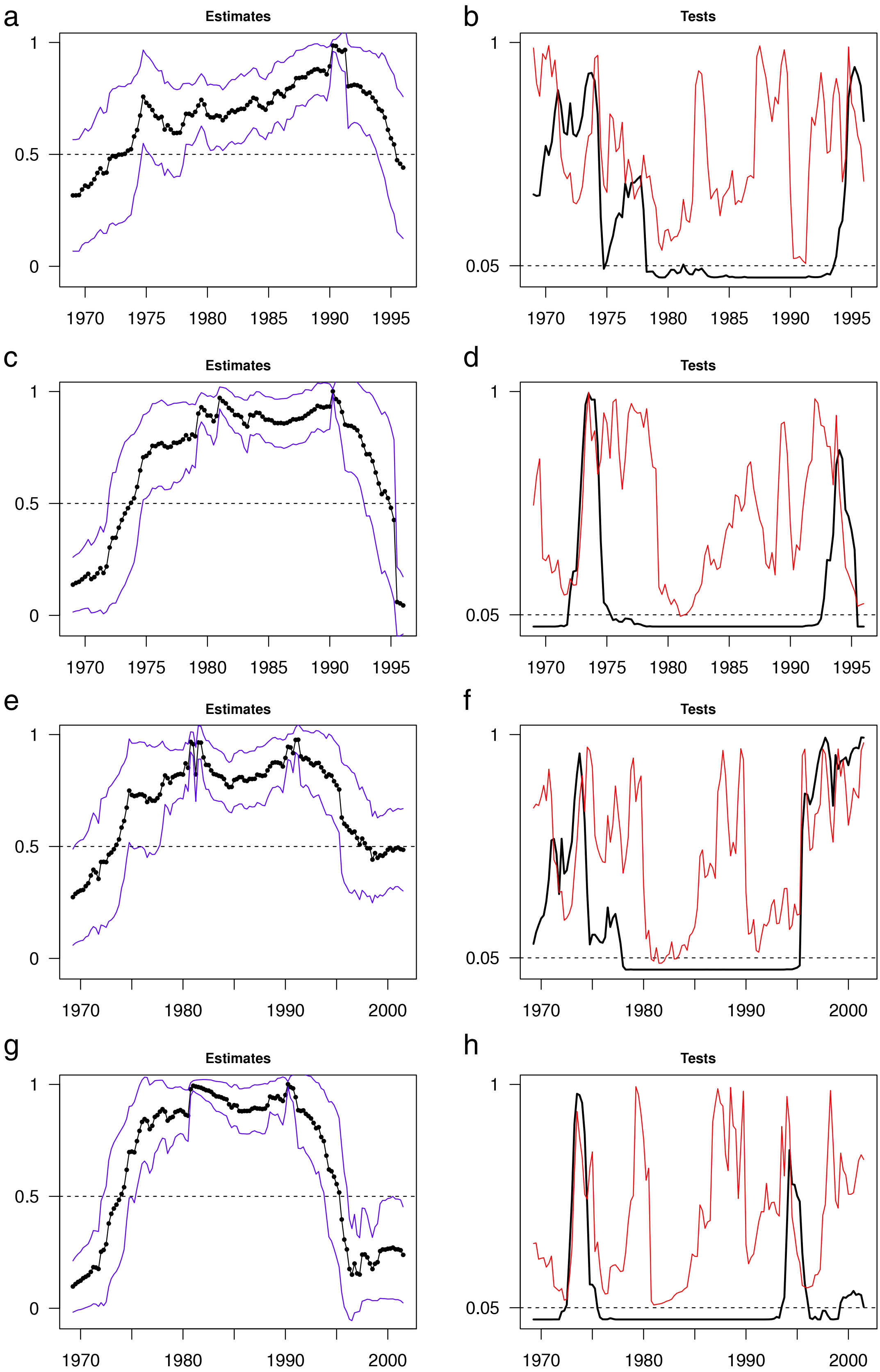

Fig. 3. Inflation forecasts ($h = 1$). Notes. Estimates and tests with 1-step-ahead inflation rate forecasts: (a) and (b) with Greenbook and real-time data; (c) and (d) with Greenbook and revised data; (e) and (f) with SPF and real-time data; (g) and (h) with SPF and revised data. $V_t = (1\, y_{t-1})'$ is used.

In Figs. 3 and 4, forecast rationality under symmetric loss is consistently rejected over a wide range of periods, including the 1980s and 1990s. The estimates of $\hat{\alpha}_{T,\tau}$ in Column 1 are significantly above 0.5. This indicates that

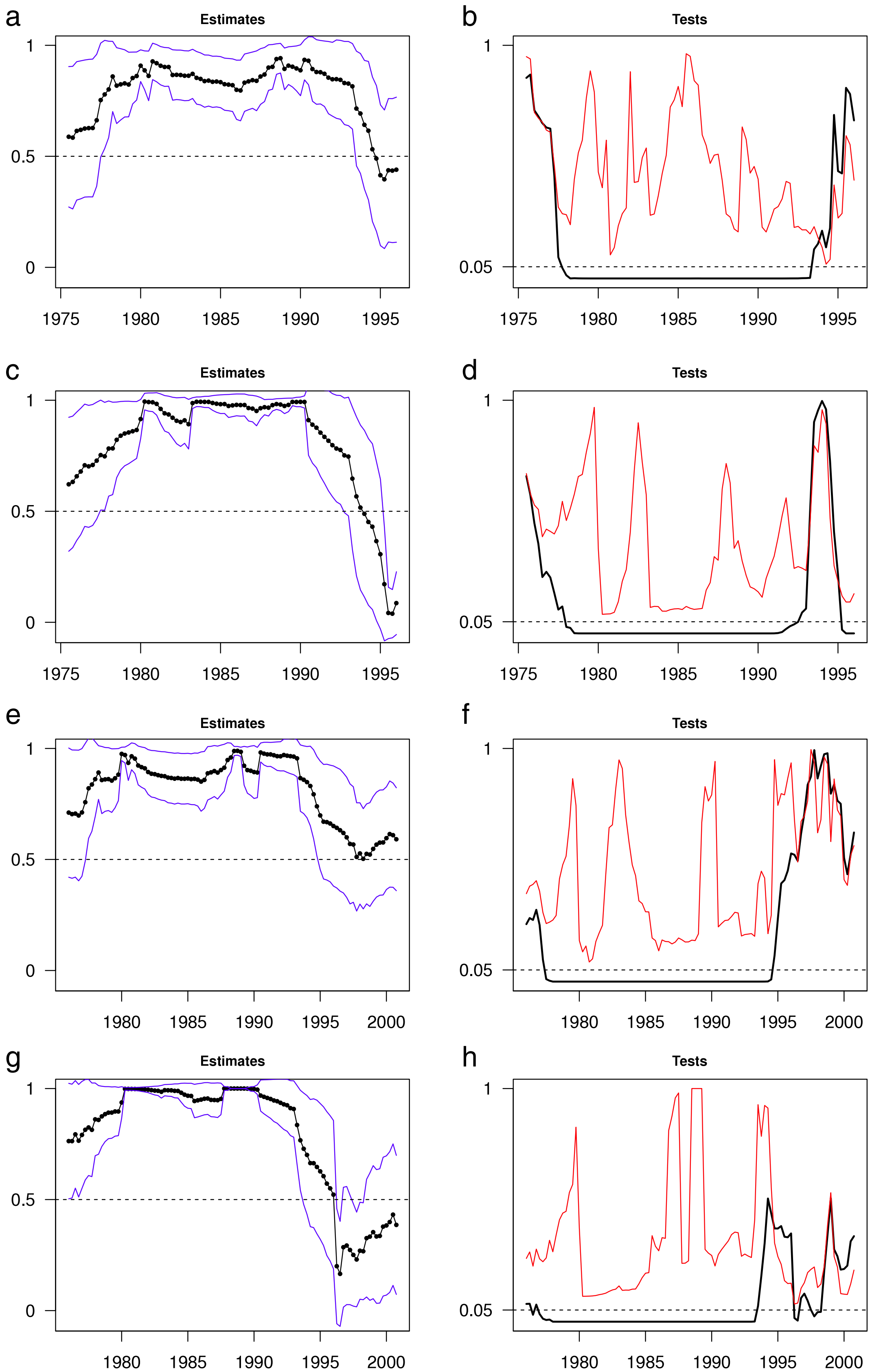

Fig. 4. Inflation forecasts ($h = 4$). Notes. Estimates and tests with 4-step-ahead inflation rate forecasts: (a) and (b) with Greenbook and real-time data; (c) and (d) with Greenbook and revised data; (e) and (f) with SPF and real-time data; (g) and (h) with SPF and revised data. $V_t = (1\ y_{t-1})'$ is used.

the forecaster consistently punishes under-prediction of the inflation rate much more than over-prediction during the 1980s and 1990s. Notably, in inflation forecasting, long-horizon inflation forecasts exhibit a greater degree of loss asymmetry than short-horizon inflation forecasts. We have computed the same statistics (not reported here, to save space) for Greenbook and SPF inflation nowcasts ($h = 0$), and find that the level of loss asymmetry is much milder in inflation nowcasts than in longer-horizon inflation forecasts.

The graphs show a surprising degree of similarity between the asymmetric loss preferences of the Greenbook and SPF forecasts in most rolling periods. For all pairs of Greenbook and SPF forecasts, the estimates $\hat{\alpha}_{T,\tau}$ in Column 1 generally share the same pattern. For example, compare Panels (a) and (e) of Fig. 4. Both of the estimates rise from below 0.5 to above 0.5, become stable at above 0.5 at nearly the same time τ, and drop below 0.5 at nearly the same time τ.

We summarize our results as follows. (1) Over rolling periods, the degree and direction of the asymmetry in forecast loss functions are time-varying, and forecast rationality under symmetry is often rejected. (2) Nearly all rolling windows fail to reject forecast rationality under asymmetric loss for both real output growth and the inflation rate. (3) Real output growth is consistently under-predicted in the 1990s, while the inflation rate is consistently over-predicted in the 1980s and 1990s. (4) Inflation forecasts, especially at longer horizons, exhibit a greater degree of loss asymmetry in both magnitude and frequency than output growth forecasts. (5) The above results are similar for both Greenbook and SPF forecasts.

These results are obtained from using the other sets of instrumental variables similarly, as was discussed in the previous section. Although we do not report them here for space reasons (though they are available in the supplementary appendix), we make some observations on the results from using the different instruments. In the sets of instruments we considered, the standard errors of $\hat{\alpha}_{T,\tau}$ are typically small, but they sometimes become smaller when using $V_t = \left(1\, \hat{f}_{t+h}\right)'$ or $V_t = \left(1\, \hat{\varepsilon}_{t-1}\right)'$. The power of the forecast rationality test also becomes slightly higher than when using $V_t = (1\, y_{t-1})'$. Meanwhile, the estimates $\hat{\alpha}_{T,\tau}$ make more apparent abrupt changes in adjacent rolling periods when the instrument $V_t = \left(1\, \hat{f}_{t+h}\right)'$ is used, especially in inflation forecasts. These abrupt changes in $\hat{\alpha}_{T,\tau}$ may reflect the genuine time-varying nature of the asymmetry, or may be due to the quality of the instruments, and/or the small rolling window size $T = 40$. Instruments with three elements also give similar results in the estimates of $\hat{\alpha}_{T,\tau}$ and in the test p-values. For example, the results from using $V_t = \left(1\, y_{t-1}\, \hat{f}_{t+h}\right)'$ are similar to those from using $V_t = \left(1\, \hat{f}_{t+h}\right)'$ or $V_t = (1\, y_{t-1})'$. Adding y_{t-1}^2, we consider $V_t = \left(1\, y_{t-1}\, y_{t-1}^2\right)'$, which leads to a slight increase in the power of the forecast rationality test, and smaller standard errors.

In addition, the robustness of our results is checked in several dimensions. Our results are robust to the choice of p: the results with $p = 1$ are similar to those with $p = 2$, but with a milder degree of loss asymmetry. Our results are also robust to the choice of T, the size of the rolling windows. A rolling window of 40 quarters may be a value which is not so long as to erase the differences between different time periods, yet not so short as to undermine the power of the tests and the asymptotic results. With a short rolling window like $T = 20$, the power of the rationality test might be severely undermined, while a long rolling window like $T = 80$ can smooth out the potential time variation in $\hat{\alpha}_{T,\tau}$. The results in the figures presented are for $T = 40$ (quarters). The results for $T = 20, 80$ are similar to those presented, and so are not reported, to save space.

Various different vintages are used for computing the realized values of real output growth and the inflation rate. While our main results are qualitatively robust to the choice of data vintages, a major difference is observed for the realized values of different vintages, especially for the real output growth forecasts. In both Table 1 and Figs. 1–4, the estimates of the loss function parameter for real output growth forecasts are generally smaller when using revised realized values for the forecast target variable than when using real-time realized values. This difference can be explained by the upward trend of data revisions after the second revision: the average real output growth from 1968Q4 to 2005Q4 in revised data exceeds that in real-time data by 0.52% in absolute value, while the difference is only 0.16% for inflation forecasts. The upward trend of data revisions in real output growth is large enough to have a significant effect on the tests. Because the realized values (y) in revised data are generally larger than those in real-time data, the forecast error ($\varepsilon = y - f$) tends to be larger with revised data, so there are more under-predictions and fewer over-predictions with revised data, which generally amounts to smaller estimated values of α in rolling periods. It may be more appropriate to use the real-time data in actual forecasting practice, as "the data (in latest vintage) they are using have been revised over time and differ significantly from the data used by forecasters in real time" (Croushore & Stark, 2001); therefore, using the revised data for realized real output growth would give the estimates a stronger level of asymmetry in the forecast producer's loss function than is actually appropriate. Specifically, the difference is most pronounced in Fig. 1, the one-step-ahead real output growth forecast of both Greenbook and SPF, where we can see big differences depending on which version of the data is used.

5. Conclusions

This paper examines the asymmetry in forecast loss functions of the Greenbook and SPF forecasts over rolling periods. We find both the degree and the direction of the asymmetry in the forecast loss functions of Greenbook and SPF to be time-varying over rolling samples. This implies that using the full sample (as in Table 1) or assuming a constant level of asymmetry may be misleading. Capistran (2008) found a huge difference in the loss function parameter of the Greenbook inflation forecasts between two sub-periods. This paper examines the time-variation of the loss function parameter of forecast producers using rolling windows, which offers a better picture of the asymmetry in direction, magnitude and frequency over time. This paper also confirms that Greenbook and SPF forecasts of real output growth and the inflation rate can be rationalized if asymmetry is permitted in the loss functions of these forecast agents. This finding is in line with the finding of Elliott et al. (2008) that asymmetry in loss functions is necessary in order to obtain forecast rationality for SPF forecasts of nominal and real output growth. These time-varying asymmetry results are similar to different sets of instrumental variables for the estimation of the asymmetric loss and for the test of forecast rationality.

One interesting result is the asymmetry of under-prediction in real output growth during the 1990s and of over-prediction in the inflation rate during the 1980s and 1990s. For Greenbook, this long period of significant over-prediction of the inflation rate coincides with a strict monetary policy during the 1980s and 1990s, and with consequent low inflation. Romer and Romer (2004) argue that "the well-tempered monetary policies of the 1950s and of the 1980s and 1990s stemmed from a conviction that inflation has high costs and few benefits, together with realistic views about the sustainable level of unemployment and the determinants of inflation". According to this, Greenbook's over-prediction of inflation, reflected in its asymmetric loss, is preemptive, leading to a policy that is intended to lower inflation.

Another interesting result is that, when different vintages are used for the realized value of real output growth, the level of loss asymmetry is more pronounced for real output growth forecasts when the last vintage (revised) data are used than when the real-time vintage data are used. If the real-time data are not used in evaluating forecasts of real output growth, the loss asymmetry can be exaggerated significantly.

Finally, it is useful to note the similarity in loss preferences between Greenbook and SPF, which may be a consequence of SPF's keeping up with Greenbook in terms of loss preferences, hoping to benefit the private sectors (producers of SPF forecasts) with the future monetary policy of FRB (the producer of the Greenbook). This similarity means that the information contained in SPF may be valuable, since the Greenbook forecasts are published with a five-year delay. It remains to be seen whether the recent directional change in SPF loss preferences can foretell the changes in the loss preferences of Greenbook.

Acknowledgments

The authors would like to thank the editor, Graham Elliott, and two anonymous referees for many helpful comments and suggestions.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.ijforecast.2013.07.017.

References

Capistran, C. (2008). Bias in Federal Reserve inflation forecasts: is the Federal Reserve irrational or just cautious? *Journal of Monetary Economics*, *55*(8), 1415–1427.

Capistran, C., & Timmermann, A. (2009). Disagreement and biases in inflation expectations. *Journal of Money, Credit and Banking*, *41*(2–3), 365–396.

Croushore, D. (2010). An evaluation of inflation forecasts from surveys using real-time data. *B.E. Journal of Macroeconomics*, *10*, Article 10.

Croushore, D., & Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, *105*, 111–130.

Diebold, F. X., & Lopez, J. (1996). Forecast evaluation and combination. In G. S. Maddala & C. R. Rao (Eds.), *Handbook of statistics* (pp. 241–268). Amsterdam: North-Holland.

Elliott, G., Komunjer, I., & Timmermann, A. (2005). Estimation and testing of forecast rationality under flexible loss. *Review of Economic Studies*, *72*(4), 1107–1125.

Elliott, G., Komunjer, I., & Timmermann, A. (2008). Biases in macroeconomic forecasts: irrationality or asymmetric loss. *Journal of European Economic Association*, *6*(1), 122–157.

Granger, C. W. J., & Newbold, P. (1986). *Forecasting economic time series* (2nd ed.). New York: Academic Press.

Komunjer, I., & Owyang, M. T. (2012). Multivariate forecast evaluation and rationality testing. *The Review of Economics and Statistics*, *94*(4), 1066–1080.

Mincer, J., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In J. Mincer (Ed.), *Economic forecasts and expectations* (pp. 81–111). New York: National Bureau of Economic Research.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*, 703–708.

Patton, A. J., & Timmermann, A. (2007a). Properties of optimal forecasts under asymmetric loss and nonlinearity. *Journal of Econometrics*, *140*(2), 884–918.

Patton, A. J., & Timmermann, A. (2007b). Testing forecast optimality under unknown loss. *Journal of the American Statistical Association*, *102*(480), 1172–1184.

Patton, A. J., & Timmermann, A. (2010). Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics*, *57*(7), 803–820.

Patton, A. J., & Timmermann, A. (2012). Forecast rationality tests based on multi-horizon bounds. *Journal of Business and Economic Statistics*, *30*(1), 1–17.

Romer, C. D., & Romer, D. H. (2000). Federal Reserve information and the behavior of interest rates. *American Economic Review*, *90*, 429–457.

Romer, C. D., & Romer, D. H. (2004). Choosing the Federal Reserve chair: lessons from history. *The Journal of Economic Perspectives*, *18*(1), 129–162.

Rossi, B. (2012). Comment on: Forecast rationality tests based on multi-horizon bounds. *Journal of Business and Economic Statistics*, *30*(1), 25–29.

Sims, C. A. (2002). The role of models and probabilities in the monetary policy process. *Brookings Papers on Economic Activity*, *2*, 1–62.

Varian, H. R. (1975). A Bayesian approach to real estate assessment. In S. E. Fienberg & A. Zellner (Eds.), *Studies in Bayesian econometrics and statistics in honor of Leonard J. Savage* (pp. 195–208). Amsterdam: North-Holland.

Yiyao Wang is a Ph.D. student at Booth School of Business at the University of Chicago. Wang earned his B.A. in finance from Fudan University in China (2013).

Tae-Hwy Lee is a professor of economics at the University of California, Riverside. Lee earned a B.A. in economics from Seoul National University (1985) and his Ph.D. in economics from the University of California, San Diego (1990).