K.7

Eliciting GDP Forecasts from the FOMC's Minutes Around the Financial Crisis

Neil R. Ericsson

Please cite paper as:
Ericsson, Neil R. (2015). Eliciting GDP Forecasts from the FOMC's Minutes Around the Financial Crisis International Finance Discussion Papers 1152.

http://dx.doi.org/10.17016/IFDP.2015.1152

International Finance Discussion Papers
Board of Governors of the Federal Reserve System

Number 1152
November 2015

Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 1152

November 2015

Eliciting GDP Forecasts from the FOMC's Minutes Around the Financial Crisis

Neil R. Ericsson

NOTE: International Finance Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References to International Finance Discussion Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors. Recent IFDPs are available on the Web at www.federalreserve.gov/pubs/ifdp/. This paper can be downloaded without charge from the Social Science Research Network electronic library at www.ssrn.com.

Eliciting GDP Forecasts from the FOMC's Minutes Around the Financial Crisis

Neil R. Ericsson*

Abstract: Stekler and Symington (2016) construct indexes that quantify the Federal Open Market Committee's views about the U.S. economy, as expressed in the minutes of the FOMC's meetings. These indexes provide insights on the FOMC's deliberations, especially at the onset of the Great Recession. The current paper complements Stekler and Symington's analysis by showing that their indexes reveal relatively minor bias in the FOMC's views when the indexes are reinterpreted as forecasts. Additionally, these indexes provide a proximate mechanism for inferring the Fed staff's Greenbook forecasts of the U.S. real GDP growth rate, years before the Greenbook's public release.

Keywords: Autometrics, bias, Fed, financial crisis, FOMC, forecasts, GDP, Great Recession, Greenbook, impulse indicator saturation, projections, Tealbook, United States.

JEL classifications: E58, C53.

---

*Forthcoming in the *International Journal of Forecasting*. The author is a staff economist in the Division of International Finance, Board of Governors of the Federal Reserve System, Washington, DC 20551 USA, and a Research Professor of Economics, Department of Economics, The George Washington University, Washington, DC 20052 USA. He may be reached at ericsson@frb.gov and ericsson@gwu.edu. The views in this paper are solely the responsibility of the author and should not be interpreted as necessarily representing or reflecting the views of the Federal Open Market Committee, its principals, the Board of Governors of the Federal Reserve System, or of any other person associated with the Federal Reserve System. This paper uses publicly available information, and only publicly available information. It does *not* use any internal or confidential Federal Reserve Board information, either directly or indirectly. The author is grateful to David Hendry, Aaron Markiewitz, Jaime Marquez, Ellen Meade, J Seymour, Tara Sinclair, Herman Stekler, and Joyce Zickler for helpful discussions and comments. All numerical results were obtained using PcGive Version 14.0B3, Autometrics Version 1.5e, and Ox Professional Version 7.00 in 64-bit OxMetrics Version 7.00: see Doornik and Hendry (2013) and Doornik (2009).

1 Introduction

Monetary policy decisions by the Fed's Federal Open Market Committee (FOMC) have attracted considerable attention in recent years, especially with quantitative easing through large-scale asset purchases and with the introduction of forward guidance; see Bernanke (2012) and Yellen (2012) *inter alia*. The FOMC's decisions are based in part on the Greenbook forecasts, which are economic forecasts produced by the Fed's staff. Romer and Romer (2008), Sinclair, Joutz, and Stekler (2010), Nunes (2013), and Ericsson, Hood, Joutz, Sinclair, and Stekler (2013) *inter alia* have extensively analyzed the Greenbook forecasts; and Banternghansa and McCracken (2009) and Sheng (2015) have examined the FOMC participants' own forecasts.

Stekler and Symington (2016) propose a creative and insightful innovation on such existing studies of U.S. monetary policy. Stekler and Symington employ a textual analysis of the minutes of the FOMC meetings for 2006–2010, developing quantitative indexes that reflect the extent of optimism or pessimism expressed in the FOMC minutes themselves on the current and future outlook for the U.S. economy. Through these indexes and accompanying research, Stekler and Symington provide key insights on the views of the FOMC, especially those held at the onset of the recent financial crisis and Great Recession. Specifically, Stekler and Symington find temporary inaccurate assessments by the FOMC of the economy at that time—in part due to inaccurate advance estimates of GDP growth.

The current paper shows that Stekler and Symington's indexes reveal even more. With few exceptions, Stekler and Symington's indexes imply relatively minor bias in the FOMC's views when the indexes are reinterpreted as forecasts. Furthermore, these indexes very closely track the Greenbook forecasts of the current-quarter and one-quarter-ahead U.S. real GDP growth rates. Stekler and Symington's indexes thus provide a proximate mechanism for inferring these Greenbook forecasts, years in advance of the public release of the Greenbook. The minutes of an FOMC meeting are published *three weeks* after the meeting itself, whereas the Greenbook is not released to the public until at least *five years* after it is presented to the FOMC.

This paper is organized as follows. Section 2 describes Stekler and Symington's indexes, the Greenbook forecasts, and the data being forecast. Section 3 discusses different approaches to testing for forecast bias, and it proposes impulse indicator saturation as a generic test of potentially time-varying forecast bias. Section 4 presents evidence on forecast bias; and it constructs "post-casts" of the 2010 Greenbook forecasts, conditional on Stekler and Symington's indexes. Section 5 concludes.

Before proceeding, it is important to note that the presence of forecast bias (or lack thereof) is consequential, both economically and statistically. That said, the particular sense in which forecast bias is consequential depends in part on whether the Greenbook forecast and Stekler and Symington's indexes are interpreted as "forecasts" or as "projections", where "projections" are in the sense of being policy sim-

ulations conditional upon a certain set of assumptions. If they are interpreted *qua* forecasts, then forecast bias implies potential room for improvement in terms of standard performance measures, although forecast improvement may be feasible for unbiased forecasts as well. If the forecasts are interpreted *qua* projections, then forecast bias implies limited usefulness of the forecasts as representing interesting hypothetical paths for economic policy. With that in mind, the Greenbook forecasts and Stekler and Symington's indexes are always referred to as "forecasts" below, while recognizing that sometimes they may be more usefully viewed as projections. This broader usage of the term "forecast" is also in line with Clements and Hendry (2002, p. 2): "A forecast is any statement about the future".

2 The FOMC Minutes Index, Forecasts, and Data

This section describes the indexes constructed by Stekler and Symington (2016) from the minutes of the FOMC, some related forecasts, and the data on the U.S. real GDP growth rates being forecast. The construction of these indexes may involve truncation and nonlinearities, so this section also discusses those issues.

FOMC Minutes Index. Stekler and Symington (2016) (hereafter, S&S) employ a focused textual analysis of the minutes for the 40 FOMC meetings during 2006–2010. From their textual analysis, S&S construct quantitative indexes that gauge the FOMC's views on the current and future strength of the U.S. economy, as expressed in the FOMC minutes themselves; see S&S (Sections 3.1.1, 3.1.2, and 4.1) in particular. S&S regard these indexes as "measuring the FOMC's views about the state of the economy" (Section 3.1.1); and the indexes are scaled such that they correspond to GDP growth rates. To design their indexes, S&S examine certain sections of the minutes that discuss:

(i) *the current economic outlook*, typically in a paragraph or paragraphs beginning "The information reviewed at the ... meeting suggested that ...”; and

(ii) *the future economic outlook*, typically in a paragraph or paragraphs beginning "In their discussion of the economic situation and outlook, meeting participants ...".

In these sets of paragraphs, Stekler and Symington search for select keywords that characterize views on the outlook. Keywords range from "strong", "robust", "considerable", "upbeat", "brisk", and "surge" for a very optimistic outlook to "recession", "contraction", and "sharp and widespread decline" for a very pessimistic one. S&S (Table 2) provide details. From the frequencies of occurrence of the keywords, Stekler and Symington create two indexes, one for the current outlook and one for the future outlook. These indexes are called FOMC Minutes Indexes (or FMIs) below. S&S then

Table 1: A Schematic for the Generation of the FOMC Minutes Index.

Information or Transformation	Description	Units, Range, or Source
g_t, x_{0t}	Greenbook forecast (g_t), and certain qualitative and quantitative information (x_{0t})	Numerical and textual
$\downarrow$ $x_{1t} = \phi_0(g_t, x_{0t}, e_{0t})$	(a) Writing of the FOMC minutes	FOMC
x_{1t}	Text of the FOMC minutes	Text
$\downarrow$ $x_{2t} = \phi_1(x_{1t}, e_{1t})$	(b) Quantification of the FOMC minutes	S&S (Table 3)
x_{2t}	Stekler and Symington's score index	$[-1, +1]$
$\downarrow$ $f_t = \phi_2(x_{2t}, e_{2t})$	(c) Calibration of the score to forecasts of the GDP growth rate	S&S (Table 4)
f_t	FOMC Minutes Index (FMI)	$[-1\%, +4\%]$

document properties of their indexes, including through comparisons with the Greenbook forecasts and with the forecasts from the Survey of Professional Forecasters (SPF). Section 4 below analyzes additional properties of S&S's indexes.

The empirical results in S&S and those below are subject to an important caveat. Specifically, the *context* of the paragraphs for (i) and (ii) in the FOMC minutes affects the interpretation of the FMIs. Starting with the January 2009 minutes, the paragraphs for (i) and (ii) respectively appear under separate sections titled:

"*Staff Review* of the Economic Situation", and

"*Participants' Views* on Current Conditions and the Economic Outlook",

or similar (italics added). Thus, the "current-outlook FMI" draws on text about the Fed staff's views, whereas the "future-outlook FMI" ostensibly reflects the views of the FOMC participants on both current conditions and future outlook. These nuances are germane to the interpretation of the FMIs when compared with (e.g.) the Fed staff's Greenbook forecasts, as in S&S (Section 4.1 and Figure 2). Section 4.3 discusses these subtleties further. See Danker and Luecke (2005) for a valuable perspective on the evolution of the FOMC's minutes.

Table 1 presents a schematic for the generation of the FMIs, formalizing the creation of the FOMC minutes and the derivation of the FMI therefrom, as follows.

(a) *Writing of the FOMC minutes.* The FOMC draws on the Greenbook forecast (g_t), certain qualitative and quantitative information (x_{0t}), and possibly additional information (e_{0t}) in writing the text of the FOMC minutes (x_{1t}). That transformation is denoted $\phi_0(g_t, x_{0t}, e_{0t})$. The subscript t is the time index that dates the FOMC meetings.

The variable x_{0t} includes information on the housing and financial markets; and the Greenbook forecast itself draws on that information as well. This presence of x_{0t} motivates in part S&S's interest in ascertaining x_{0t}'s roles in influencing the minutes x_{1t}. The text of the minutes also provides the basis for S&S's assessment of the FOMC's views on risk and uncertainty to the economy.

(b) *Quantification of the FOMC minutes as a score.* S&S (Table 2) then analyze the text in the minutes of the FOMC meetings to create a score index (x_{2t}), as given in S&S (Table 3). That quantification of the text may involve additional information (e_{1t}). The implied transformation is $\phi_1(x_{1t}, e_{1t})$.

(c) *Calibration of the score.* Finally, S&S (Table 4) calibrate that score index to forecasts of the GDP growth rate, again possibly involving additional information (e_{2t}). The implied transformation is $\phi_2(x_{2t}, e_{2t})$, and the resulting FMI is denoted f_t. The variable being forecast is the U.S. real GDP growth rate, denoted y_t.

The FMI thus is an implicit function of the Greenbook forecast, certain qualitative and quantitative information, and possible additional information involved in (a) the writing of the FOMC minutes, (b) the quantification of the FOMC minutes as a score, and (c) the calibration of that score to forecasts of the GDP growth rate. This description applies to the FMIs for both the current outlook and the future outlook, albeit with the time horizon (current or future) being implicit in the notation.

Equation (1) articulates the implicit function for the FMI in three ways.

$$\begin{aligned} f_t &= \phi_2(\phi_1(\phi_0(g_t, x_{0t}, e_{0t}), e_{1t}), e_{2t}) \\ &= \phi_m(g_t, x_{0t}, e_{0t}, e_{1t}, e_{2t}) \\ &\approx g_t \,. \end{aligned} \tag{1}$$

The first line in equation (1) expresses the FMI as an explicit repeated function of its information sources. The second line shows the implied implicit function $\phi_m(\cdot)$. The third line indicates that the FMI may be approximated to some degree by the Greenbook forecast alone.

Comparing the FMI directly with the Greenbook forecast is thus of interest; see Section 4.2. Equation (1) also clarifies that the FMI is a forecast in its own right. Hence, it can be evaluated as a forecast, such as with tests of unbiasedness and efficiency; see Section 3. Moreover, the third line approximates the nonlinear function $\phi_m(g_t, x_{0t}, e_{0t}, e_{1t}, e_{2t})$ by a simple linear one ($= 0 + 1g_t$), per White (1980). Impulse

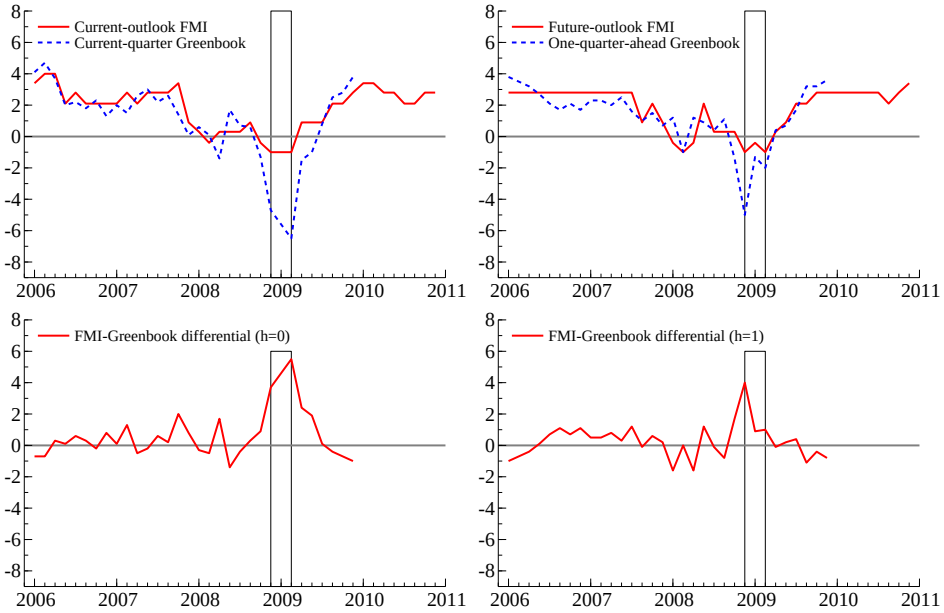

Figure 1: The current-outlook and future-outlook FMIs, the Greenbook forecasts of the current-quarter and one-quarter-ahead U.S. real GDP growth rates, and the differentials between the FMI and Greenbook forecasts.

indicator saturation can help determine the degree to which that approximation holds; see Section 3.

Greenbook forecasts. Figure 1 compares the FMI with the Greenbook forecasts in a 2×2 panel of graphs, with each graph labeled by a suffix a, b, c, or d, as follows: $\begin{bmatrix} \text{a} & \text{b} \\ \text{c} & \text{d} \end{bmatrix}$. Figure 1a plots two "nowcasts": the FMI for the current outlook, and the Greenbook forecast for the current quarter. Figure 1b plots two forecasts: the FMI for the future outlook, and the Greenbook forecast for one quarter ahead.[1] Figure 1c plots the difference between the FMI for the current outlook and the Greenbook forecast for the current quarter; and Figure 1d plots the difference between the FMI for the future outlook and the Greenbook forecast for one quarter ahead. In Figure 1 (and likewise in Figure 3 below), the time t is the date of the FOMC meeting to which a Greenbook is submitted and from which an FMI is constructed. The graphs

[1] S&S (footnote 12) compare the FMI for the future outlook with *two*-quarter-ahead Greenbook forecasts. One-quarter-ahead Greenbook forecasts are employed herein, noting the behavior of the root mean squared forecast errors for the Greenbook forecasts as a function of the forecast horizon; see Ericsson, Hood, Joutz, Sinclair, and Stekler (2013), Ericsson, Fiallos, and Seymour (2014), and Chang and Hanson (2015). That said, the forecast-encompassing tests described below permit sorting out which horizon or horizons are relevant. Additional analysis (not reported herein) suggests that combinations of Greenbook forecasts at different horizons may be particularly useful for explaining the FMI for both the current outlook and the future outlook.

Table 2: Dates of FOMC meetings for which the FMI may have truncated the FOMC's forecast, and (in percent per annum) the corresponding FMI (f_t^h), Greenbook forecast (g_t^h), and their differential ($f_t^h - g_t^h$) at horizons $h = 0$ and $h = 1$.

Date t	f_t^0	g_t^0	$f_t^0 - g_t^0$	f_t^1	g_t^1	$f_t^1 - g_t^1$
March 2006	+4.0	4.7	–0.7	—	—	—
May 2006	+4.0	3.7	+0.3	—	—	—
March 2008	—	—	—	–1.0	–1.0	+0.0
December 2008	–1.0	–4.7	+3.7	–1.0	–5.0	+4.0
January 2009	–1.0	–5.6	+4.6	—	—	—
March 2009	–1.0	–6.5	+5.5	–1.0	–2.0	+1.0

in Figure 1 show that the FMI and the Greenbook forecasts are generally very close numerically, whether for the current outlook or for the future outlook.

There are, however, large deviations between the FMI and the Greenbook forecast for the FOMC meetings in December 2008, January 2009, and March 2009: see the boxed-in areas in Figure 1. For these meetings, S&S's score is at its minimum ($= -1$), the most pessimistic outlook. However, the FMI corresponding to that score is only a 1.0% per annum rate of decline in GDP: see S&S (Table 4), which maps the score to the FMI. The Greenbook forecasts were typically much more negative for those meetings. This discrepancy (or "deviation") between the FMI and the Greenbook forecast arises because the score (and hence the FMI) truncates both tails of the distribution of the implicit forecast in the FOMC minutes.

Truncation. To examine this issue further, Table 2 lists the dates of *all* FOMC meetings for which the score may have truncated the forecast implicit in the minutes, along with the FMI f, the Greenbook forecast g, and the differential between them. The forecast horizon h appears as a superscript, where $h = 0$ and $h = 1$ respectively denote the current outlook and future outlook for the FMI (and below, for the score x_2), or the current quarter and one quarter ahead for the Greenbook forecast.

The deviations for the current outlook are large for all three meetings with very pessimistic assessments; and the FMI is much larger than the corresponding Greenbook forecast, by 3%–6% per annum. The deviation for the future outlook is large for December 2008, but not for January 2009 (when the future outlook was less pessimistic) or for March 2009 (when the discrepancy was only 1% per annum). For the current outlook, the score is at its most *optimistic* value for the FOMC meetings in March 2006 and May 2006, with the FMI being +4.0% per annum; but the corresponding deviations are smaller than 1% per annum in absolute value. For the future outlook, there are no meetings for which the score (and hence the FMI) is at its most

optimistic value.

S&S's score algorithm can thus truncate the forecast implicit in the FOMC minutes. That algorithm appears to do so for three FOMC meetings at the onset of the Great Recession. These large deviations also could have arisen if the forecast implicit in the FOMC minutes itself deviated markedly from the Greenbook forecast; but that explanation seems less plausible in light of the close match between the FMI and the Greenbook forecast for other dates.

Linearity in calibration. The function $\phi_2(x_{2t}, e_{2t})$ calibrates S&S's score to have units comparable to forecasts of GDP growth rates. That calibration function is very close to linear, as a simple regression of the current-outlook FMI (f_t^0) on the current-outlook score (x_{2t}^0) demonstrates.

$$f_t^0 \quad = \quad \underset{(0.005)}{1.515} \quad + \quad \underset{(0.009)}{2.508}\ x_{2t}^0 \tag{2}$$

$$T = 40\ [2006\text{F1–}2010\text{F8}] \qquad \mathrm{R}^2 = 0.9995 \qquad \hat{\sigma} = 0.0315\%$$

Here and below, estimated standard errors are in parentheses $(\cdot)$ under coefficient estimates; T is the total number of observations; the letter "F" indexes the FOMC meeting within the year (F1, ..., F8), paralleling the common notation of "Q" for "quarter" (Q1, ..., Q4); R^2 is the squared multiple correlation coefficient; and $\hat{\sigma}$ is the estimated residual standard error. Equation (2) is a regression, and so the left-hand side variable is implicitly the *fitted value* of f_t^0, not f_t^0 itself.

To put the regression (2) in context, none of its residuals are larger than 0.05% per annum, whereas the actual GDP growth rates in S&S (Table 5) and the Greenbook forecasts thereof are reported to a tenth of a percent per annum. Thus, to a very good degree of approximation, the FMI f_t is 2.5 times S&S's score x_{2t}, plus 1.5. That is, for the current outlook, the transformation $f_t = \phi_2(x_{2t}, e_{2t})$ is essentially linear in the score x_{2t}; and e_{2t} is approximately zero. Results are numerically very similar for the future outlook.

Actual GDP growth rates. Figure 2 plots the FMIs and actual GDP growth rates in a 2×2 panel similar to the one in Figure 1, albeit with time t in Figure 2 being the date of the actual GDP growth rate, and with each FMI forecast aligned correspondingly. Figure 2a and 2b graph FMIs for the current outlook and future outlook respectively, each along with the actual GDP growth rate as measured 90 days after quarter end and as measured historically. Figure 2c and 2d plot the corresponding discrepancies, which are interpretable as forecast errors in this context. As S&S note, and as is apparent in Figures 2a–b, the actual values were revised markedly downward for several quarters at the beginning of the financial crisis, including by 2.2% for 2007Q3, 3.7% for 2008Q1, and 2.0% for 2008Q4. Figures 2c–d show that the large and negative FMI "forecast errors" for late 2008 and early 2009 reflect these data revisions and the effects of truncation discussed above.

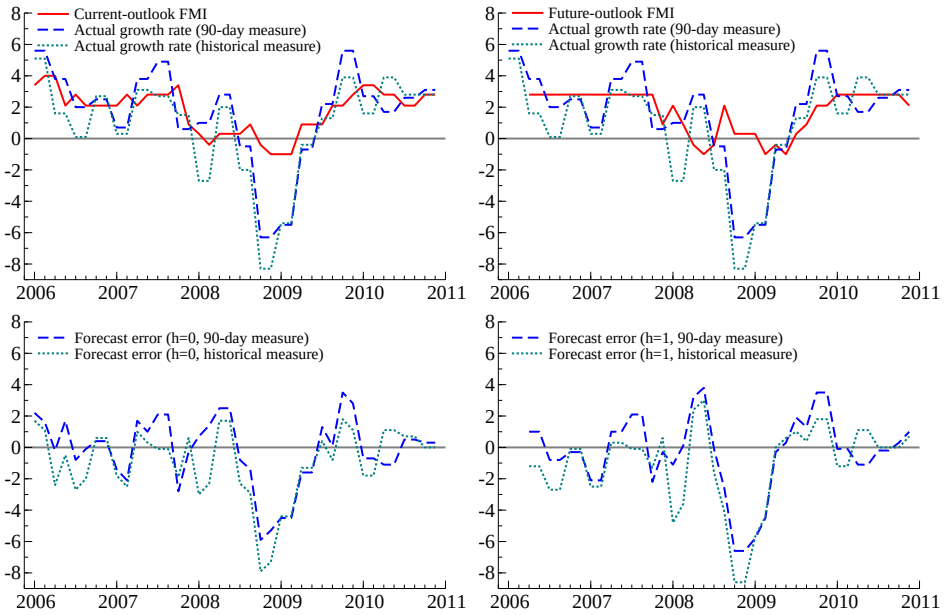

Figure 2: FMIs for the current outlook and future outlook, and the actual GDP growth rate as measured 90 days after quarter end and as measured historically; and the corresponding FMI forecast errors.

The actual GDP growth rates are quarterly, expressed at annual rates and spanning 2006–2010. For the sequential pair of FOMC meetings within a given quarter, the value of an actual GDP growth rate is repeated, thereby providing actual values for comparison with the FMI and Greenbook forecasts. See Ahumada (1992, especially Figure 3) for a similar interpolation of quarterly Argentine total final domestic expenditure to monthly values.

Data details. Sources and practical details for the FMIs, Greenbook forecasts, and actual GDP growth rates follow. The units for these series are quarterly growth rates, measured in percent per annum.

S&S construct their scores and FMIs for the 40 FOMC meetings during 2006–2010, a period that spans the financial crisis and leads into the Great Recession. The minutes of the FOMC meetings themselves are available online at:

www.federalreserve.gov/monetarypolicy/fomc_historical.htm (for 2006–2009) and
www.federalreserve.gov/monetarypolicy/fomccalendars.htm (for 2010).

In S&S, Table 3 lists the scores, Figures 1A and 1B plot them, Table 4 provides the algorithm for converting the scores to the FMIs, and Figure 2 plots the current-outlook score in its original units and in the units for the FMI.

The Fed staff's Greenbooks are available on the first website listed above. Currently, the Greenbooks are available through only December 2009, given the approx-

imate 5-year delay in their release to the public. There are thus 32 observations of Greenbook forecasts for a given horizon, with those forecasts obtained from Greenbooks spanning 2006–2009. S&S (Figure 2) plot the Greenbook forecasts of the current quarter through 2009, with comparable SPF forecasts for 2010 appended at the end. (The SPF forecasts are *not* considered in the analysis below.)

S&S (Table 5) list the "90-day estimate" and "historical" values (S&S's terminology) for the quarterly GDP growth rates over 2006–2010. These two measurements correspond to the U.S. Bureau of Economic Analysis's "third estimate" release, which is typically available about 90 days after the end of the relevant quarter; and to the complete set of values for 2006–2010 as reported by the Bureau of Economic Analysis when S&S was initially written; see Stekler and Symington (2014).

Remarks. Before proceeding to a detailed assessment of the FMIs, two observations about Stekler and Symington (2016) are germane. First, Stekler and Symington follow a common approach to textual analysis. For instance, Boukus and Rosenberg (2006) use latent semantic analysis to assess the roles of different themes in the FOMC's minutes; and Meade, Burk, and Josselyn (2015) calculate the changing frequencies of different quantitative words in the FOMC's minutes to ascertain the diversity of views among the FOMC members and participants. Quantitative textual analysis also has numerous precedents outside of economics, as in Cox and Brandwood (1959).

Second, Stekler and Symington's analysis is novel by quantifying qualitative text from the minutes on the *outlook* and comparing that quantification with the Greenbook forecasts. By contrast, Meade, Burk, and Josselyn focus on the disparity of views in the minutes, rather than some central tendency of views. Banternghansa and McCracken (2009, 2014) likewise focus on the disparity of views, albeit as measured by individual participants' economic forecasts. Yet other researchers have compared the Greenbook forecasts and FOMC participants' forecasts with each other and with other forecasts.

Stekler and Symington (2016) thus apply an innovative and creative quantitative textual analysis to elicit a measure of the nowcasts and forecasts implicit in the minutes of the FOMC meetings. That measure is the FMI and, because it is a forecast *elicited* from the FOMC minutes, it is denoted an "el-cast". Even though the text that Stekler and Symington examine includes little or no quantitative information, their FMIs reveal much about the thinking of the FOMC participants and (as Section 4 below shows) about the Fed staff's input to the FOMC meetings.

3 Assessing and Comparing Forecasts

This section considers different approaches for assessing and comparing forecasts, starting with the standard test of forecast bias by Mincer and Zarnowitz (1969). The

Mincer–Zarnowitz framework is then extended to allow for (a) the suspected truncations noted in Section 2, (b) general time variation in forecast bias, and (c) comparison of multiple forecasts, rather than of a single forecast against the outcome. These extensions include many standard procedures for forecast evaluation as special cases.

Mincer–Zarnowitz. Mincer and Zarnowitz (1969, pp. 8–11) suggest testing for forecast bias by regressing the forecast error on an intercept and testing whether the intercept is statistically significant. That is, for a variable y_t at time t and its forecast $\hat{y}_t$, estimate the equation:

$$(y_t - \hat{y}_t) \;=\; a \;+\; u_t\,, \qquad t = 1, \ldots, T, \tag{3}$$

where a is the intercept, and u_t is the error term at time t. In the present context, y_t is the growth rate of real GDP (either the 90-day estimate or the historical value), and $\hat{y}_t$ is either the FMI or the Greenbook forecast.

A test of $a = 0$ is interpretable as a test that the forecast $\hat{y}_t$ is unbiased for the variable y_t. For one-step-ahead forecasts, the error u_t may be serially uncorrelated, in which case a standard t-statistic or F-statistic may be appropriate. For multi-step-ahead forecasts, u_t generally will be serially correlated; hence inference about the intercept a may require some accounting for that autocorrelation.

Suspected systematic effects. The potential truncations documented in Table 2 may distort the FMI as a measure of the forecast implicit in the FOMC minutes. If a measured forecast is distorted in certain time periods, it may be advantageous to account for those distortions in the Mincer–Zarnowitz regression (3), both to correct for any biases that those distortions induce in estimating a, and to improve the efficiency of the estimation of a. Direct adjustment of regression (3) is feasible for known dates of potential distortions:

$$(y_t - \hat{y}_t) \;=\; \textstyle\sum_{j \in \mathcal{J}} d_j J_{jt} \;+\; a \;+\; u_t\,, \qquad t = 1, \ldots, T, \tag{4}$$

where J_{jt} is a dummy variable that is unity for $t = j$ and zero otherwise, d_j is the corresponding coefficient for J_{jt}, and $\mathcal{J}$ is the set of dates for which distortions are suspected. Equation (4) thus provides the basis for calculating the Mincer–Zarnowitz test while correcting for suspected distortions in the measurement of the forecast.

Arbitrarily time-varying forecast bias. Equation (4) is also interpretable as allowing for a specific form of time dependence in the bias of the *measured* forecast, where that bias depends on the degree of truncation. That time dependence could be completely general, while also explicitly allowing for the date-specific potential distortions:

$$\begin{aligned}(y_t - \hat{y}_t) \;&=\; \textstyle\sum_{j \in \mathcal{J}} d_j J_{jt} \;+\; a_t \;+\; u_t \\ &=\; \textstyle\sum_{j \in \mathcal{J}} d_j J_{jt} \;+\; \sum_{i=1}^{T} c_i I_{it} \;+\; u_t\,, \qquad t = 1, \ldots, T.\end{aligned} \tag{5}$$

The intercept (now denoted a_t) is explicitly time-dependent; the impulse indicator I_{it} is a dummy variable that is unity for $t = i$ and zero otherwise; and c_i is the corresponding coefficient for I_{it}. Because the $\{c_i\}$ may have any values whatsoever, the intercept a_t in (5) may vary arbitrarily over time. In this context, a test that all coefficients c_i are equal to zero is a generic test of forecast unbiasedness, corrected for potential truncation distortions. In determining the form of a_t, equation (5) thus retains some variables (the $\{J_{jt}\}$) without selection, paralleling the model-selection approach in Hendry and Johansen (2015). Because equation (5) includes at least T coefficients, equation (5) cannot be estimated unrestrictedly. However, the question being asked *can* be answered with impulse indicator saturation.

Impulse indicator saturation. Impulse indicator saturation (IIS) uses the zero-one dummies $\{I_{it}\}$ to analyze properties of a model. Unrestricted inclusion of all T dummies in the model (thereby "saturating" the sample) is infeasible. However, blocks of dummies *can* be included, and statistically significant dummies can be retained from those blocks. That insight provides the basis for IIS. See Ericsson and Reisman (2012) for an intuitive non-technical exposition of IIS, and Hendry and Doornik (2014) for extensive analysis in the context of automatic model selection.

As background, IIS provides a general procedure for robust estimation and for model evaluation—in particular, for testing parameter constancy. IIS is a generic test for an unknown number of structural breaks, occurring at unknown times, with unknown duration and magnitude, anywhere in the sample. IIS is a powerful empirical tool for both evaluating and improving existing empirical models. Hendry (1999) proposes IIS as a procedure for testing parameter constancy. Further discussion, recent developments, and applications appear in Hendry, Johansen, and Santos (2008), Doornik (2009), Johansen and Nielsen (2009, 2013, 2015), Hendry and Santos (2010), Ericsson (2011a, 2011b, 2012), Ericsson and Reisman (2012), Bergamelli and Urga (2013), Hendry and Pretis (2013), Hendry and Doornik (2014), Pretis, Mann, and Kaufmann (2015), and Castle, Doornik, Hendry, and Pretis (2015). Ericsson (2015) proposes a new application for IIS—as a generic test for time-varying forecast bias. Section 4 applies IIS to test for potential bias in the FMI as a forecast; and IIS addresses measurement errors in the FMI due to suspected truncation biases.

Direct comparison of two forecasts. Equation (5) also provides the basis for direct comparison of forecasts. Specifically, consider an alternative forecast of y_t, denoted $\tilde{y}_t$. The corresponding forecast error $(y_t - \tilde{y}_t)$ is equivalent to $\sum_{i=1}^{T}(y_t - \tilde{y}_t) \cdot I_{it}$, which is a weighted sum of the impulse indicators. Equation (5) may thus be rewritten as:

$$(y_t - \hat{y}_t) \;=\; \textstyle\sum_{j\in\mathcal{J}} d_j^* J_{jt} \;+\; (y_t - \tilde{y}_t) \;+\; \sum_{i=1}^{T} c_i^* I_{it} \;+\; u_t\,, \qquad t = 1, \ldots, T. \tag{6}$$

Subtracting $(y_t - \tilde{y}_t)$ from both sides of equation (6) obtains:

$$\begin{aligned}(\tilde{y}_t - \hat{y}_t) \;&=\; \textstyle\sum_{j\in\mathcal{J}} d_j^* J_{jt} \;+\; \sum_{i=1}^{T} c_i^* I_{it} \;+\; u_t \\ &=\; \textstyle\sum_{j\in\mathcal{J}} d_j^* J_{jt} \;+\; a_t^* \;+\; u_t\,, \qquad\qquad t = 1, \ldots, T.\end{aligned} \tag{7}$$

The coefficients a_t^*, c_i^*, and d_j^* in equations (6)–(7) need not equal the coefficients a_t, c_i, and d_j in equation (5). Also, tests on $\{c_i^*, d_j^*\}$ (or, equivalently, on $\{a_t^*, d_j^*\}$) in equation (7) allow direct comparison of the two forecasts $\hat{y}_t$ and $\tilde{y}_t$.

Remarks. Equations (3)–(7) are directly related to several common forecast tests. First, Mincer and Zarnowitz (1969, p. 11) propose a variation on equation (3) in which the coefficient on $\hat{y}_t$ is estimated rather than imposed. See Holden and Peel (1990), Stekler (2002), and Sinclair, Stekler, and Carnow (2015) *inter alia*; and note that this modified formulation generates a test of forecast efficiency as well as a test of forecast unbiasedness. Second, if $\tilde{y}_t$ $(= \sum_{i=1}^{T} \tilde{y}_t \cdot I_{it})$ rather than $(y_t - \tilde{y}_t)$ is included on the right-hand side of equation (6), Chong and Hendry's (1986) forecast-encompassing test emerges. Third, as Ericsson (1992b) discusses, a necessary condition for forecast encompassing is having the smallest root mean squared forecast error (RMSFE); and Granger (1989) and Diebold and Mariano (1995) propose tests of whether one model's RMSFE is less than another model's RMSFE. Fourth, the regression for the forecast-encompassing test can be extended to compare several forecasts at once, in which case the right-hand side of the regression includes the differential of each alternative forecast relative to the "primary" forecast; see Ericsson and Marquez (1993). Fifth, the regression for the forecast-encompassing test may include forecast revisions across multiple horizons, paralleling the direct extension of forecast encompassing to include a comparison with multiple forecasts. Nordhaus (1987) proposes this revisions-based test as a test of efficiency; see also Dovern, Fritsche, Loungani, and Tamirisa (2015) and Coibion (2015). Tversky and Kahneman (1974) earlier described "anchoring" as a potential phenomenon exhibited by forecast revisions; see Campbell and Sharpe (2009) for empirical evidence on anchoring. In a related vein, equation (5) extends the framework for regression-based tests of predictive accuracy in West and McCracken (1998).

Many common procedures for evaluating forecasts are thus derivable from the generic framework in equation (5), which characterizes general and potentially arbitrarily time-varying forecast bias. Equally, these common procedures are interpretable as specific forms of potentially time-varying forecast bias. As the simplest case, the original Mincer–Zarnowitz test from (3) is a test of constant (i.e., *time-invariant*) forecast bias. This observation highlights the strength of the Mincer–Zarnowitz test (that it focuses on detecting a constant nonzero forecast bias) and also its weakness (that it assumes that the forecast bias *is* constant over time). These characteristics of the Mincer–Zarnowitz test bear directly on the empirical results in Section 4.

4 The FMI and the Greenbook Forecasts

This section examines the FMI as an elicited forecast (el-cast) of the U.S. real GDP growth rate over 2006–2010, and it considers the FMI's relationship to the Greenbook

forecast. Section 4.1 shows that standard (Mincer–Zarnowitz) tests of forecast bias detect economically small and statistically insignificant biases for the FMI compared against the *90-day estimate* of the growth rate; and IIS-type tests find similar results. However, standard tests find economically large and statistically significant biases of the FMI relative to the *historical* growth rates, whereas IIS-type tests find much smaller and less significant bias. Those differences arise because the standard tests ignore truncation, whereas the IIS-type tests adjust for it. Section 4.2 shows how the FMI can post-cast as-yet-unreleased Greenbook forecasts. Section 4.3 considers some implications.

4.1 Forecast Biases for the FMI

The FMI is a relatively unbiased forecast of the U.S. real GDP growth rate, once accounting for truncation. Tables 3 and 4 respectively report analyses of the FMI relative to the 90-day estimate of the growth rate, and of the FMI relative to the historical growth rate. Each table lists regression coefficients, estimated standard errors, and various summary statistics for regressions in which the FMI forecast error $(y-f)$ is the dependent variable. Results are reported for the current-outlook ($h=0$) and future-outlook ($h=1$) FMIs, and for regressions with and without truncation adjustment ("TA"). The sample periods are 2006F1–2009F8 ($T=32$) for $h=0$ and 2006F3–2009F8 ($T=30$) for $h=1$, noting that each quarter has two FOMC meetings. The sample periods used in these regressions end in 2009F8 in order to maintain comparability with publicly available Greenbook forecasts. (Section 4.2 considers the FMI in 2010.) Dates on impulse indicator dummies correspond to the dates of actual GDP growth rates.

In Table 3, the first column of numbers lists results for the standard Mincer–Zarnowitz regression (3) for the current-outlook FMI, with no truncation adjustment. The estimated bias is -0.17% and is statistically insignificant with an F-statistic of 0.15 (p-value of 0.704); $\hat{\sigma}$ is 2.445%; and (for comparison) the RMSFE of the FMI is 2.412%.

The first column also lists diagnostic statistics for testing against various alternative hypotheses, thereby helping reveal the equation's strengths and weaknesses. These statistics test for autocorrelation in the residuals (AR), residual autoregressive conditional heteroscedasticity ($ARCH$), and residual nonnormality as characterized by skewness and excess kurtosis (*Normality*); see Godfrey (1978), Engle (1982), and Jarque and Bera (1980) and Doornik and Hansen (2008) respectively. For a given diagnostic statistic, the entries within a block of numbers are the value of the statistic for testing the null hypothesis against the designated maintained hypothesis, its tail probability (in square brackets) calculated for the approximate distribution under the null hypothesis, and that distribution (either $\chi^2(\cdot)$ or $F(\cdot,\cdot)$, with degrees of freedom in parentheses). Superscript asterisks * and ** denote rejections of the null hypothesis

Table 3: Coefficients, estimated standard errors, and summary statistics for regressions with FMI forecast errors $(y - f)$ for the U.S. real GDP growth rate, with the growth rate as measured 90 days after quarter end.

Regressor or statistic	$h = 0$, no TA	$h = 0$, with TA	$h = 1$, no TA	$h = 1$, with TA
Intercept	-0.17	0.31	-0.36	0.23
	(0.43)	(0.38)	(0.51)	(0.44)
J_{2008F8}	–	-5.61	–	-6.83
		(2.06)		(2.30)
J_{2009F1}	–	-4.81	–	-6.03
		(2.06)		(2.30)
J_{2009F2}	–	-4.81	–	-4.73
		(2.06)		(2.30)
$\hat{\sigma}$	2.445%	2.025%	2.803%	2.260%
RMSFE	2.412%	2.014%	2.779%	2.230%
Mincer–Zarnowitz	0.15	0.68	0.49	0.28
(F-statistic)	[0.704]	[0.416]	[0.491]	[0.602]
	$F(1, 31)$	$F(1, 28)$	$F(1, 29)$	$F(1, 26)$
F-statistic for	–	5.73**	–	6.20**
truncation adjustment		[0.003]		[0.003]
		$F(3, 28)$		$F(3, 26)$
AR(3) LM statistic	7.18**	2.14	14.7**	6.77**
	[0.001]	[0.120]	[0.000]	[0.002]
	$F(3, 28)$	$F(3, 25)$	$F(3, 26)$	$F(3, 23)$
ARCH(3) LM statistic	4.00*	0.27	9.31**	0.33
	[0.018]	[0.845]	[0.000]	[0.804]
	$F(3, 26)$	$F(3, 26)$	$F(3, 24)$	$F(3, 24)$
Normality statistic	5.75	6.04*	3.66	7.18*
	[0.056]	[0.049]	[0.160]	[0.028]
	$\chi^2(2)$	$\chi^2(2)$	$\chi^2(2)$	$\chi^2(2)$
Dates of additional	$2008F7$,	$2008F7$	$2008F7$,	$2008F7$
impulses selected by	$2008F8$,		$2008F8$,	
IIS (1% target size)	$2009F1$,		$2009F1$	
	$2009F2$			
F-statistic for the	9.96**	14.4**	9.56**	14.4**
additional impulses	[0.000]	[0.001]	[0.000]	[0.001]
	$F(4, 27)$	$F(1, 27)$	$F(3, 26)$	$F(1, 25)$

Notes. Regressions are for current-outlook ($h = 0$) and future-outlook ($h = 1$) FMIs, without and with truncation adjustment ("TA").

Table 4: Coefficients, estimated standard errors, and summary statistics for regressions with FMI forecast errors $(y - f)$ for the U.S. real GDP growth rate, with the growth rate as measured historically.

Regressor or statistic	$h = 0$, no TA	$h = 0$, with TA	$h = 1$, no TA	$h = 1$, with TA
Intercept	-1.23 (0.43)	-0.80 (0.39)	-1.46 (0.53)	-0.93 (0.48)
J_{2008F8}	–	-6.50 (2.13)	–	-7.67 (2.54)
J_{2009F1}	–	-3.60 (2.13)	–	-4.77 (2.54)
J_{2009F2}	–	-3.60 (2.13)	–	-3.47 (2.54)
$\hat{\sigma}$	2.445%	2.095%	2.915%	2.490%
RMSFE	2.701%	2.209%	3.215%	2.613%
Mincer–Zarnowitz (F-statistic)	8.08^{**} [0.008] $F(1, 31)$	4.23^{*} [0.049] $F(1, 28)$	7.49^{*} [0.010] $F(1, 29)$	3.73 [0.064] $F(1, 26)$
F-statistic for truncation adjustment	–	4.73^{**} [0.009] $F(3, 28)$	–	4.59^{*} [0.010] $F(3, 26)$
AR(3) LM statistic	4.06^{*} [0.016] $F(3, 28)$	1.42 [0.262] $F(3, 25)$	8.55^{**} [0.000] $F(3, 26)$	4.12^{*} [0.018] $F(3, 23)$
ARCH(3) LM statistic	4.33^{*} [0.013] $F(3, 26)$	0.04 [0.990] $F(3, 26)$	9.00^{**} [0.000] $F(3, 24)$	0.32 [0.812] $F(3, 24)$
Normality statistic	6.20^{*} [0.045] $\chi^2(2)$	9.36^{**} [0.009] $\chi^2(2)$	5.29 [0.071] $\chi^2(2)$	7.09^{*} [0.029] $\chi^2(2)$
Dates of additional impulses selected by IIS (1% target size)	$2008F7$, $2008F8$	$2008F7$	$2008F7$, $2008F8$, $2009F1$	$2008F7$
F-statistic for the additional impulses	12.8^{**} [0.000] $F(2, 29)$	19.9^{**} [0.000] $F(1, 27)$	10.1^{**} [0.000] $F(3, 26)$	15.3^{**} [0.001] $F(1, 25)$

Notes. Regressions are for current-outlook ($h = 0$) and future-outlook ($h = 1$) FMIs, without and with truncation adjustment ("TA").

at the 5% and 1% levels respectively.

Finally, the column lists dates of impulse dummies selected by IIS at a 1% target size, and the F-statistic for those retained impulses. Notably, IIS selects impulse dummies for dates for which truncation is suspected (2008F8, 2009F1, and 2009F2), and for the date immediately preceding (2008F7). Even without focusing the regression on dates with possible truncation, IIS detects those dates.

The second numerical column in Table 3 reports the same regression, but where impulse dummies J_{jt} are included for the three suspected truncation dates. Those dummies are highly significant (F-statistic of 5.73, with a p-value of 0.003); and the only additional impulse dummy selected by IIS is for 2008F7. The estimated bias is +0.31% and is statistically insignificant. The third and fourth numerical columns report results for the future outlook, both without and with truncation adjustment; and those results are similar to those for the current outlook.

Table 4 repeats the analysis in Table 3 but, in calculating the forecast errors, uses the historical values of the GDP growth rate rather than the 90-day estimates. As S&S note, the historical values and the 90-day estimates differ markedly for several periods. Those differences are reflected in Table 4's regressions. For the current outlook and future outlook, the "pure" Mincer–Zarnowitz regression (3) detects large and highly significant negative biases: of −1.23% and −1.46% respectively. However, with adjustment for possible truncation, estimated biases are only about two thirds of those values (−0.80% and −0.93%); and they are only marginally significant or insignificant at the 5% level. Standard tests of time-invariant forecast bias can and do mislead. Also, IIS again detects 2008F7 as an anomaly.

4.2 Post-casting the Greenbook Forecasts

The FMI very closely approximates the Greenbook forecast, once accounting for the FMI's truncation in three observations. Table 5 reports regressions of the form in equation (7), with the dependent variable being the differential between the FMI and the Greenbook forecast ($f - g$). The sample period is 2006F1–2009F8 ($T = 32$) for both $h = 0$ and $h = 1$. Each date in that sample period corresponds to the date of an FOMC meeting, to which a Greenbook is submitted and from which an FMI is constructed.

The first numerical column in Table 5 indicates a statistically significant differential of 0.66% between the FMI and the Greenbook forecast for the current quarter ($h = 0$). However, once adjusted for possible truncation effects (second numerical column), that estimated bias is only 0.26% and is statistically insignificant. For $h = 1$, no statistically significant or numerically large differential is detectable, whether or not the regression is adjusted for possible truncation effects.

For both $h = 0$ and $h = 1$, the (truncation-adjusted) FMI and the Greenbook forecast are very close numerically. Deviations between them are small, typically less

Table 5: Coefficients, estimated standard errors, and summary statistics for regressions of $(f - g)$, the deviation between the FMI and the Greenbook forecast.

Regressor or statistic	$h = 0$, no TA	$h = 0$, with TA	$h = 1$, no TA	$h = 1$, with TA
Intercept	0.66 (0.28)	0.26 (0.17)	0.27 (0.19)	0.09 (0.16)
J_{2008F8}	–	3.44 (0.95)	–	3.91 (0.86)
J_{2009F1}	–	4.34 (0.95)	–	0.81 (0.86)
J_{2009F2}	–	5.24 (0.95)	–	0.91 (0.86)
$\hat{\sigma}$	1.582%	0.938%	1.078%	0.850%
RMSE	1.692%	0.956%	1.094%	0.840%
Mincer–Zarnowitz (F-statistic)	5.61* [0.024] $F(1, 31)$	2.15 [0.154] $F(1, 28)$	1.94 [0.173] $F(1, 31)$	0.32 [0.574] $F(1, 28)$
F-statistic for truncation adjustment	–	20.1** [0.000] $F(3, 28)$	–	7.28** [0.001] $F(3, 28)$
AR(3) LM statistic	7.97** [0.000] $F(3, 28)$	0.48 [0.699] $F(3, 25)$	0.91 [0.446] $F(3, 28)$	1.13 [0.356] $F(3, 25)$
ARCH(3) LM statistic	5.37** [0.005] $F(3, 26)$	1.31 [0.292] $F(3, 26)$	0.12 [0.948] $F(3, 26)$	0.29 [0.830] $F(3, 26)$
Normality statistic	18.3** [0.000] $\chi^2(2)$	2.65 [0.265] $\chi^2(2)$	10.6** [0.005] $\chi^2(2)$	0.60 [0.741] $\chi^2(2)$
Dates of additional impulses selected by IIS (1% target size)	$2008F8$, $2009F1$, $2009F2$	None	$2008F8$	None
F-statistic for the additional impulses	20.1** [0.000] $F(3, 28)$	–	20.0** [0.000] $F(1, 30)$	–

Notes. The dependent variable is either the current-outlook FMI minus the Greenbook forecast of the current-quarter U.S. real GDP growth rate ($h = 0$), or the future-outlook FMI minus the Greenbook forecast of the one-quarter-ahead U.S. real GDP growth rate ($h = 1$). Regressions are without and with truncation adjustment ("TA").

than 1% per annum and often less than 0.5% per annum. The root mean squared errors (RMSEs) and residual standard errors are correspondingly small, under 1% per annum. From IIS, the adjustment for possible truncation in the FMI appears sufficient to correct for that truncation; and only that truncation appears to need correction. Notably, the one-quarter-ahead RMSEs are *smaller* than the current-quarter RMSEs. That is, S&S's indexes are more accurate at inferring the one-quarter-ahead Greenbook forecast than the current-quarter Greenbook forecast.

This empirical analysis establishes a close relationship between the FMI and the Greenbook forecast. This relationship presents a special opportunity for "post-casting" Greenbook forecasts, conditional on the FMI, noting that the minutes of an FOMC meeting are available three weeks after the meeting, whereas the Greenbook forecasts are not released to the public until approximately five years after the Greenbook itself is presented to the FOMC. Specifically, the regressions in Table 5 suggest a potentially viable mechanism for post-casting the Greenbook forecasts in 2010, a year for which the Greenbook forecasts have not yet been released publicly. Noting that the regressions' dependent variable is the differential $(f-g)$, the Greenbook forecast g can be post-cast, conditional on the calculated value of the FMI f. While inversion of a statistical relationship—as between f and g—generally is inadvisable, inversion of a deterministic or near-deterministic relationship may be useful; see Ericsson (1992a, Section 2C) and Ericsson, Hendry, and Hood (2015, Section 4.6).

Figures 3a and 3b plot the FMI-based post-casts of the 2010 Greenbook forecasts, with ± 1 standard error bands. These post-casts range between 2.1% and 3.4% for both $h = 0$ and $h = 1$, with somewhat different dynamics for each forecast horizon h. The forecast standard error is under 1%. Figures 3c and 3d plot the observed in-sample discrepancies between the FMI and the Greenbook forecasts for 2006–2009 with truncation adjustment.

4.3 Remarks and Implications

Several implications follow directly from the results above. First, and foremost, Stekler and Symington's indexes provide a proximate mechanism for inferring Greenbook forecasts of the current-quarter and one-quarter-ahead U.S. real GDP growth rates, years in advance of the public release of the Greenbook.

Second, the current-outlook FMI draws on text about the Fed *staff*'s views, whereas the future-outlook FMI ostensibly reflects the views of the FOMC participants on *both* current conditions and future outlook. In principle, these nuances are germane to the interpretation of the FMIs when compared with the Fed staff's Greenbook forecasts. However, Section 4.2 implies that these distinctions are relatively unimportant over the sample considered, Nunes's (2013) analysis notwithstanding.

Third, it may seem surprising that FOMC participants' views for the future outlook—as measured by the future-outlook FMI—are well-captured by the *one-*

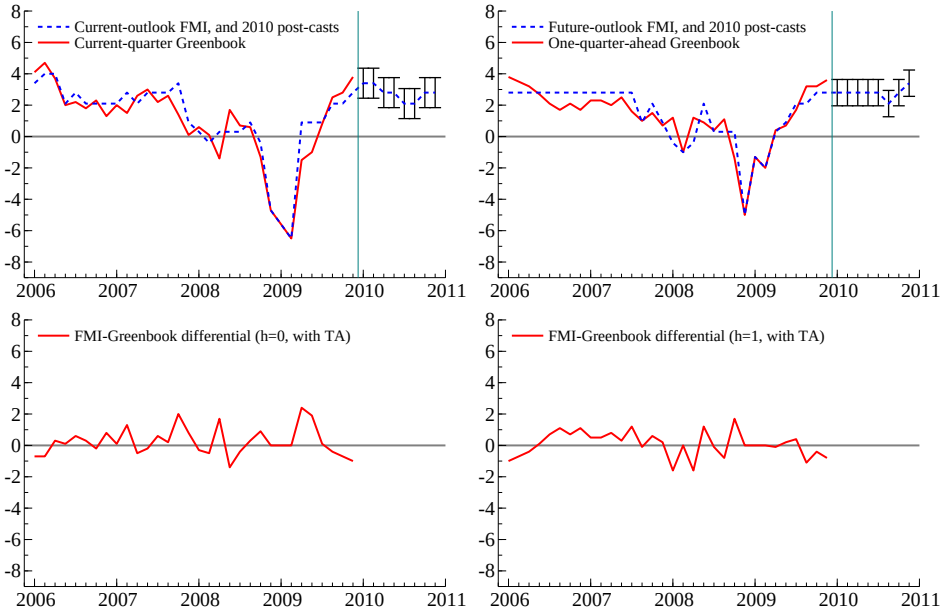

Figure 3: The post-casts of the 2010 Greenbook forecasts for current-quarter and one-quarter-ahead U.S. real GDP growth rates, with ± 1 standard error bands; and the observed in-sample discrepancies between the FMI and Greenbook forecasts for 2006–2009 with truncation adjustment.

quarter-ahead Greenbook forecast, noting that the policy-relevant horizon may be somewhat longer than one quarter ahead. The future-outlook FMI may thus be an even better proxy for Greenbook forecasts at longer horizons. Or, participants' views may down-weight Greenbook forecasts at longer horizons, noting that Greenbook forecasts at two or more quarters ahead appear relatively uninformative when compared with naive forecasts such as a random walk or an unconditional mean growth rate. See Ericsson, Fiallos, and Seymour (2014) and Chang and Hanson (2015), noting that the former includes evaluation of U.S. as well as foreign growth. Also, because the future-outlook FMI quantifies the participants' views for the future outlook in the minutes, comparison of the future-outlook FMI with the FOMC's Summary of Economic Projections may be of interest.

Fourth, the FMI itself could be improved, extended, and expanded. The measurement errors in the FMI due to truncation could be reduced or eliminated by expanding the range of Stekler and Symington's score. The FMI could be constructed on a finer grid, noting that it currently has intervals of 0.6% or 0.7% per annum between adjacent values; see S&S (Table 4). And, the FMI could be extended through 2015, from which post-casts of the Greenbook for 2011 through 2015 could be constructed. The FMI also could be extended to meetings prior to 2006. Indexes over those periods

might be more or less accurate than those over 2006–2009; and the Greenbooks from 2010 onward have yet to be released publicly.[2] FMIs also could be constructed for other economic variables, such as inflation and unemployment.

Fifth, IIS serves as a diagnostic and exploratory tool for detecting potential missing information in a regression. IIS also serves to correct or adjust for such anomalies: here, specifically, by robustifying the results to the truncation effects. Focusing on specific known or suspected dates can enhance the analysis, albeit with supplemental IIS offering some protection from unsuspected events. IIS thus helps detect when the FMI deviated substantially from outcomes, and from the Greenbook forecasts. Clearly, "rejection of the null doesn't imply the alternative". However, for time series data, the date-specific nature of IIS-type procedures can aid in identifying important sources of forecast error. Use of these tests in forecast development is consistent with a progressive modeling approach; see White (1990).

Sixth, the analysis above implicitly involves factorizing the joint distribution of the FMI, the Greenbook forecast, and the actual GDP growth rate into conditional and marginal distributions. As Engle, Hendry, and Richard (1983) and Ericsson (1992a) discuss, some factorizations may prove more fruitful than others; and this aspect merits further investigation.

Seventh, textual analysis of Fed documents such as the Greenbook may be fruitful. For instance, while the Greenbook does include numerical forecasts, it also includes extensive discussion of the factors underpinning those forecasts. The Greenbook constitutes what Hendry (2001, p. 18) calls "fore-diction", a combination of numerical forecasts with a narrative.

Eighth, the largest inaccuracies in the FMI occur at the onset of the financial crisis and Great Recession—in effect, at a turning point in the business cycle. As Fildes and Stekler (2002) and others document, turning points have been difficult to forecast; and the FMI's inaccuracies may reflect that challenge. From an institutional perspective, it may be useful to isolate the causes of the forecast errors according to the various assumptions made. Such an analysis could lead to improved forecasts, or at least provide a deeper understanding of the sources of the inaccuracies. Equally, robust forecasting procedures could reduce or eliminate systematic forecast errors; see Hendry (2006), Castle, Fawcett, and Hendry (2010), and Castle, Clements, and Hendry (2015) *inter alia*.

5 Conclusions

Stekler and Symington (2016) propose and construct ingenious quantitative indexes that measure the extent of optimism or pessimism expressed in the FOMC's minutes

[2] Also, starting in 2010, the Greenbook forecasts appear in a Fed document called the Tealbook, which combines the previous Fed documents called the Greenbook and the Bluebook.

on the current and future outlook for the U.S. economy. These indexes help provide insights on the views of the FOMC, especially those held at the onset of the recent financial crisis and Great Recession. The current paper complements Stekler and Symington's analysis by showing that, aside from a few specific instances, their indexes reveal relatively minor bias in the FOMC's views when the indexes are reinterpreted as forecasts. These indexes also provide a proximate mechanism for inferring Greenbook forecasts of the U.S. real GDP growth rate, years before the public release of the Greenbook. Impulse indicator saturation proves central to obtaining these results.

References

Ahumada, H. (1992) "A Dynamic Model of the Demand for Currency: Argentina 1977–1988", *Journal of Policy Modeling*, 14, 3, 335–361.

Banternghansa, C., and M. W. McCracken (2009) "Forecast Disagreement Among FOMC Members", Federal Reserve Bank of St. Louis Working Paper No. 2009–059A, Research Division, Federal Reserve Bank of St. Louis, St. Louis, Missouri, December.

Banternghansa, C., and M. W. McCracken (2014) "The Effect of FOMC Forecast Disagreement on U.S. Treasuries", Presentation, International Symposium on Forecasting, Rotterdam, The Netherlands, July 1.

Bergamelli, M., and G. Urga (2013) "Detecting Multiple Structural Breaks: A Monte Carlo Study and an Application to the Fisher Equation for the US", draft, Cass Business School, London, March.

Bernanke, B. S. (2012) "U.S. Monetary Policy and International Implications", remarks at the seminar "Challenges of the Global Financial System: Risks and Governance under Evolving Globalization", Bank of Japan, Tokyo, Japan, October 14.

Boukus, E., and J. V. Rosenberg (2006) "The Information Content of FOMC Minutes", mimeo, Federal Reserve Bank of New York, New York, July.

Campbell, S. D., and S. A. Sharpe (2009) "Anchoring Bias in Consensus Forecasts and Its Effect on Market Prices", *Journal of Financial and Quantitative Analysis*, 44, 2, 369–390.

Castle, J. L., M. P. Clements, and D. F. Hendry (2015) "Robust Approaches to Forecasting", *International Journal of Forecasting*, 31, 1, 99–112.

Castle, J. L., J. A. Doornik, D. F. Hendry, and F. Pretis (2015) "Detecting Location Shifts During Model Selection by Step-indicator Saturation", *Econometrics*, 3, 2, 240–264.

Castle, J. L., N. W. P. Fawcett, and D. F. Hendry (2010) "Forecasting with Equilibrium-correction Models During Structural Breaks", *Journal of Econometrics*, 158, 1, 25–36.

Chang, A. C., and T. J. Hanson (2015) "The Accuracy of Forecasts Prepared for the Federal Open Market Committee", Finance and Economics Discussion Series Paper No. 2015–062, Board of Governors of the Federal Reserve System, Washington, D.C., July.

Chong, Y. Y., and D. F. Hendry (1986) "Econometric Evaluation of Linear Macroeconomic Models", *Review of Economic Studies*, 53, 4, 671–690.

Clements, M. P., and D. F. Hendry (2002) "An Overview of Economic Forecasting", Chapter 1 in M. P. Clements and D. F. Hendry (eds.) *A Companion to Economic Forecasting*, Blackwell Publishers, Oxford, 1–18.

Coibion, O. (2015) "Comments on Dovern, Fritsche, Loungani and Tamirisa", *International Journal of Forecasting*, 31, 1, 155–156.

Cox, D. R., and L. Brandwood (1959) "On a Discriminatory Problem Connected with the Works of Plato", *Journal of the Royal Statistical Society, Series B*, 21, 1, 195–200.

Danker, D. J., and M. M. Luecke (2005) "Background on FOMC Meeting Minutes", *Federal Reserve Bulletin*, 2005, Spring, 175–179.

Diebold, F. X., and R. S. Mariano (1995) "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, 13, 3, 253–263.

Doornik, J. A. (2009) "Autometrics", Chapter 4 in J. L. Castle and N. Shephard (eds.) *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford, 88–121.

Doornik, J. A., and H. Hansen (2008) "An Omnibus Test for Univariate and Multivariate Normality", *Oxford Bulletin of Economics and Statistics*, 70, supplement, 927–939.

Doornik, J. A., and D. F. Hendry (2013) *PcGive 14*, Timberlake Consultants Press, London (3 volumes).

Dovern, J., U. Fritsche, P. Loungani, and N. Tamirisa (2015) "Information Rigidities: Comparing Average and Individual Forecasts for a Large International Panel", *International Journal of Forecasting*, 31, 1, 144–154.

Engle, R. F. (1982) "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50, 4, 987–1007.

Engle, R. F., D. F. Hendry, and J.-F. Richard (1983) "Exogeneity", *Econometrica*, 51, 2, 277–304.

Ericsson, N. R. (1992a) "Cointegration, Exogeneity, and Policy Analysis: An Overview", *Journal of Policy Modeling*, 14, 3, 251–280.

Ericsson, N. R. (1992b) "Parameter Constancy, Mean Square Forecast Errors, and Measuring Forecast Performance: An Exposition, Extensions, and Illustration", *Journal of Policy Modeling*, 14, 4, 465–495.

Ericsson, N. R. (2011a) "Improving Global Vector Autoregressions", draft, Board of Governors of the Federal Reserve System, Washington, D.C., June.

Ericsson, N. R. (2011b) "Justifying Empirical Macro-econometric Evidence in Practice", invited presentation, online conference *Communications with Economists: Current and Future Trends* commemorating the 25th anniversary of the *Journal of Economic Surveys*, November.

Ericsson, N. R. (2012) "Detecting Crises, Jumps, and Changes in Regime", draft, Board of Governors of the Federal Reserve System, Washington, D.C., November.

Ericsson, N. R. (2015) "How Biased Are U.S. Government Forecasts of the Federal Debt?", *International Journal of Forecasting*, forthcoming.

Ericsson, N. R., E. J. Fiallos, and J. E. Seymour (2014) "Assessing Greenbook Forecasts of Foreign GDP Growth", draft, Board of Governors of the Federal Reserve System, Washington, D.C., August.

Ericsson, N. R., D. F. Hendry, and S. B. Hood (2015) "Milton Friedman as an Empirical Modeler", Chapter 6 in R. Cord and J. D. Hammond (eds.) *Milton Friedman: Contributions to Economics and Public Policy*, Oxford University Press, Oxford, forthcoming.

Ericsson, N. R., S. B. Hood, F. Joutz, T. M. Sinclair, and H. O. Stekler (2013) "Greenbook Forecasts and the Business Cycle", draft, Board of Governors of the Federal Reserve System, Washington, D.C., December.

Ericsson, N. R., and J. Marquez (1993) "Encompassing the Forecasts of U.S. Trade Balance Models", *Review of Economics and Statistics*, 75, 1, 19–31.

Ericsson, N. R., and E. L. Reisman (2012) "Evaluating a Global Vector Autoregression for Forecasting", *International Advances in Economic Research*, 18, 3, 247–258.

Fildes, R., and H. O. Stekler (2002) "The State of Macroeconomic Forecasting", *Journal of Macroeconomics*, 24, 4, 435–468.

Godfrey, L. G. (1978) "Testing for Higher Order Serial Correlation in Regression Equations When the Regressors Include Lagged Dependent Variables", *Econometrica*, 46, 6, 1303–1310.

Granger, C. W. J. (1989) *Forecasting in Business and Economics*, Academic Press, Boston, Massachusetts, Second Edition.

Hendry, D. F. (1999) "An Econometric Analysis of US Food Expenditure, 1931–1989", Chapter 17 in J. R. Magnus and M. S. Morgan (eds.) *Methodology and Tacit Knowledge: Two Experiments in Econometrics*, John Wiley and Sons, Chichester, 341–361.

Hendry, D. F. (2001) "How Economists Forecast", Chapter 2 in D. F. Hendry and N. R. Ericsson (eds.) *Understanding Economic Forecasts*, MIT Press, Cambridge, 15–41.

Hendry, D. F. (2006) "Robustifying Forecasts from Equilibrium-correction Systems", *Journal of Econometrics*, 135, 1–2, 399–426.

Hendry, D. F., and J. A. Doornik (2014) *Empirical Model Discovery and Theory Evaluation: Automatic Selection Methods in Econometrics*, MIT Press, Cambridge, Massachusetts.

Hendry, D. F., and S. Johansen (2015) "Model Discovery and Trygve Haavelmo's Legacy", *Econometric Theory*, 31, 1, 93–114.

Hendry, D. F., S. Johansen, and C. Santos (2008) "Automatic Selection of Indicators in a Fully Saturated Regression", *Computational Statistics*, 23, 2, 317–335, 337–339.

Hendry, D. F., and F. Pretis (2013) "Anthropogenic Influences on Atmospheric CO_2", Chapter 12 in R. Fouquet (ed.) *Handbook on Energy and Climate Change*, Edward Elgar, Cheltenham, 287–326.

Hendry, D. F., and C. Santos (2010) "An Automatic Test of Super Exogeneity", Chapter 12 in T. Bollerslev, J. R. Russell, and M. W. Watson (eds.) *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, Oxford University Press, Oxford, 164–193.

Holden, K., and D. A. Peel (1990) "On Testing for Unbiasedness and Efficiency of Forecasts", *The Manchester School*, 58, 2, 120–127.

Jarque, C. M., and A. K. Bera (1980) "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals", *Economics Letters*, 6, 3, 255–259.

Johansen, S., and B. Nielsen (2009) "An Analysis of the Indicator Saturation Estimator as a Robust Regression Estimator", Chapter 1 in J. L. Castle and N. Shephard (eds.) *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford, 1–36.

Johansen, S., and B. Nielsen (2013) "Outlier Detection in Regression Using an Iterated One-step Approximation to the Huber-skip Estimator", *Econometrics*, 1, 1, 53–70.

Johansen, S., and B. Nielsen (2015) "Asymptotic Theory of Outlier Detection Algorithms for Linear Time Series Regression Models", *Scandinavian Journal of Statistics*, in press.

Meade, E. E., N. A. Burk, and M. Josselyn (2015) "The FOMC Meeting Minutes: An Assessment of Counting Words and the Diversity of Views", FEDS Note, Board of Governors of the Federal Reserve System, Washington, D.C., May 26.

Mincer, J., and V. Zarnowitz (1969) "The Evaluation of Economic Forecasts", Chapter 1 in J. Mincer (ed.) *Economic Forecasts and Expectations: Analyses of Forecasting Behavior and Performance*, National Bureau of Economic Research, New York, 3–46.

Nordhaus, W. D. (1987) "Forecasting Efficiency: Concepts and Applications", *Review of Economics and Statistics*, 69, 4, 667–674.

Nunes, R. (2013) "Do Central Banks' Forecasts Take Into Account Public Opinion and Views?", International Finance Discussion Paper No. 1080, Board of Governors of the Federal Reserve System, Washington, D.C., May.

Pretis, F., M. L. Mann, and R. K. Kaufmann (2015) "Testing Competing Models of the Temperature Hiatus: Assessing the Effects of Conditioning Variables and Temporal Uncertainties Through Sample-wide Break Detection", *Climatic Change*, 131, 4, 705–718.

Romer, C. D., and D. H. Romer (2008) "The FOMC versus the Staff: Where Can Monetary Policymakers Add Value?", *American Economic Review*, 98, 2, 230–235.

Sheng, X. (2015) "Evaluating the Economic Forecasts of FOMC Members", *International Journal of Forecasting*, 31, 1, 165–175.

Sinclair, T. M., F. Joutz, and H. O. Stekler (2010) "Can the Fed Predict the State of the Economy?", *Economics Letters*, 108, 1, 28–32.

Sinclair, T. M., H. O. Stekler, and W. Carnow (2015) "Evaluating a Vector of the Fed's Forecasts", *International Journal of Forecasting*, 31, 1, 157–164.

Stekler, H. O. (2002) "The Rationality and Efficiency of Individuals' Forecasts", Chapter 10 in M. P. Clements and D. F. Hendry (eds.) *A Companion to Economic Forecasting*, Blackwell Publishers, Oxford, 222–240.

Stekler, H. O., and H. Symington (2014) "How did the FOMC View the Great Recession As It Was Happening?: Evaluating the Minutes from FOMC Meetings, 2006–2010", RPF Working Paper No. 2014–005, Research Program on Forecasting, Center of Economic Research, Department of Economics, The George Washington University, Washington, D.C., September.

Stekler, H. O., and H. Symington (2016) "Evaluating Qualitative Forecasts: The FOMC Minutes, 2006–2010", *International Journal of Forecasting*, in press.

Tversky, A., and D. Kahneman (1974) "Judgment under Uncertainty: Heuristics and Biases", *Science*, 185, 4157, 1124–1131.

West, K. D., and M. W. McCracken (1998) "Regression-based Tests of Predictive Ability", *International Economic Review*, 39, 4, 817–840.

White, H. (1980) "Using Least Squares to Approximate Unknown Regression Functions", *International Economic Review*, 21, 1, 149–170.

White, H. (1990) "A Consistent Model Selection Procedure Based on m-testing", Chapter 16 in C. W. J. Granger (ed.) *Modelling Economic Series: Readings in Econometric Methodology*, Oxford University Press, Oxford, 369–383.

Yellen, J. L. (2012) "Perspectives on Monetary Policy", remarks at the Boston Economic Club Dinner, Federal Reserve Bank of Boston, Boston, Massachusetts, June 6.