ARTICLE IN PRESS

International Journal of Forecasting ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Interpreting estimates of forecast bias

Neil R. Ericsson *

Division of International Finance, Board of Governors of the Federal Reserve System, Washington, DC 20551, USA
Department of Economics, The George Washington University, Washington, DC 20052, USA

ARTICLE INFO

Keywords:
Bias
Debt
Federal government
Forecasts
Impulse indicator saturation
Heteroscedasticity
United States

ABSTRACT

This paper resolves differences in results and interpretation between Ericsson's (2017) and Gamber and Liebner's (2017) assessments of forecasts of U.S. gross federal debt. As Gamber and Liebner (2017) discuss, heteroscedasticity could explain the empirical results in Ericsson (2017). However, the combined evidence in Ericsson (2017) and Gamber and Liebner (2017) supports the interpretation that these forecasts have significant time-varying biases. Both Ericsson (2017) and Gamber and Liebner (2017) advocate using impulse indicator saturation in empirical modeling.

Published by Elsevier B.V. on behalf of International Institute of Forecasters.

1. Introduction

Using impulse indicator saturation (IIS), Ericsson (2017) tests for and detects economically large and statistically highly significant time-varying biases in forecasts of U.S. gross federal debt over 1984–2012, particularly at turning points in the business cycle. Gamber and Liebner (2017) discuss Ericsson (2017), obtaining different empirical results and offering a different interpretation. The current paper resolves those differences through a re-examination of IIS.

Gamber and Liebner (2017) examine Ericsson's (2017) choice of IIS's significance level and interpretation of the estimated bias, concluding that the empirical basis for *time-varying* bias *per se* is weaker than claimed, and that the outliers detected by IIS could easily arise from heteroscedasticity rather than from time-varying bias. Because IIS does have power to detect heteroscedasticity, heteroscedasticity could explain the IIS results in Ericsson (2017). However, as Sections 2 and 3 below show, time-varying bias is more consistent with the combined evidence in Ericsson (2017) and Gamber and Liebner (2017). Section 4 comments further on modeling with IIS.

2. Analysis of alternative model specifications

Ericsson (2017) and Gamber and Liebner (2017) assess forecasts of U.S. federal debt, focusing on the economic and statistical bases for the selected impulse indicators from IIS. Although Ericsson (2017) and Gamber and Liebner (2017) evaluate the same set of forecasts, they obtain different empirical results and offer different interpretations of those results. Section 3 below resolves the differences in interpretation through a re-examination of IIS. The current section resolves the differences in the empirical results themselves—both qualitatively and quantitatively—through an encompassing approach by examining alternative model specifications.

In particular, encompassing analysis of an analytical example demonstrates how certain model specifications reduce the power of tests to detect impulse indicators, where that power depends directly on t-ratios for the indicators. The encompassing analysis implies that some relevant indicators may nonetheless appear unimportant in certain models, simply because those models omit relevant

* Correspondence to: Division of International Finance, Board of Governors of the Federal Reserve System, Washington, DC 20551, USA.
E-mail addresses: ericsson@frb.gov, ericsson@gwu.edu.

http://dx.doi.org/10.1016/j.ijforecast.2017.01.001
0169-2070/Published by Elsevier B.V. on behalf of International Institute of Forecasters.

Please cite this article in press as: Ericsson, N. R., Interpreting estimates of forecast bias. International Journal of Forecasting (2017), http://dx.doi.org/10.1016/j.ijforecast.2017.01.001

ARTICLE IN PRESS

variables, thereby increasing the residual standard error and hence reducing the t-ratios. The current section first presents the analytical example and then applies it to the disparate empirical results with IIS.

This type of assessment is sometimes called "mis-specification analysis" because some models analyzed omit certain relevant variables and hence are mis-specified, relative to the data generation process; see Sargan (1988, Chapter 8). Mizon and Richard (1986) propose a constructive utilization of mis-specification analysis—known as the encompassing approach—in which a given model (Model M0, below) is shown to explain or "encompass" properties of the other models (Models M1 and M2, below). In the current section, model properties include t-ratios, residual variances, and the selection of impulse dummies. See Bontemps and Mizon (2008), Davidson, Hendry, Srba, and Yeo (1978), and Mizon and Richard (1986) for further discussion.

Analytical example. To put the encompassing analysis in context, suppose that both blocks of observations for bare-bones IIS include impulse dummies that have nonzero coefficients in the data generation process (DGP). In bare-bones IIS, estimation of coefficients for dummies that saturate a given block then implies omission of the other block's relevant dummies in the corresponding model. These omitted dummies typically result in reduced power to detect the significance of included dummies. An analytical example illustrates.[1]

In a notation similar to that in Ericsson (2017, Example 2), let the DGP for the variable w_t be as follows.

$$\text{DGP}: \; w_t = \delta_0 + \delta_1 I_{1t} + \delta_2 I_{2t} + \varepsilon_t,$$
$$\varepsilon_t \sim NID(0, \sigma^2), \; t = 1, \ldots, T. \tag{1}$$

That is, w_t is normally and independently distributed with a constant mean δ_0 and constant variance σ^2 over T observations, except that w_t's mean is $\delta_0 + \delta_1$ in period $t = t_1$ (when the impulse indicator I_{1t} is nonzero) and $\delta_0 + \delta_2$ in period $t = t_2$ (when $I_{2t} \neq 0$). For expository purposes, assume that δ_1 and δ_2 are both strictly positive, and that t_1 and t_2 are in the first and second blocks of observations respectively.

Consider three models, denoted M0, M1, and M2. Model M0 is specified as the DGP (1) itself.

$$\text{Model M0}: \; w_t = \delta_0 + \delta_1 I_{1t} + \delta_2 I_{2t} + \varepsilon_t. \tag{2}$$

Models M1 and M2 entail omitted variables. Model M1 includes I_{1t} but omits I_{2t}.

$$\text{Model M1}: \; w_t = \delta_0 + \delta_1 I_{1t} + v_{1t}. \tag{3}$$

Model M2 includes I_{2t} but omits I_{1t}.

$$\text{Model M2}: \; w_t = \delta_0 + \delta_2 I_{2t} + v_{2t}. \tag{4}$$

For Model M1, the error v_{1t} is $(\delta_2 I_{2t} + \varepsilon_t)$, so Model M1's mean squared error σ_1^2 is:

$$\sigma_1^2 = (\sigma^2 + \delta_2^2/T), \tag{5}$$

which is larger than σ^2, the error variance for Model M0. Likewise, for Model M2, the error v_{2t} is $(\delta_1 I_{1t} + \varepsilon_t)$, and the mean squared error σ_2^2 is:

$$\sigma_2^2 = (\sigma^2 + \delta_1^2/T), \tag{6}$$

which also is larger than σ^2.

One possible consequence of model specifications such as M1 and M2 is to shrink t-ratios on included variables. As Eqs. (5) and (6) imply, the estimated residual variance in a model with an omitted relevant variable is typically larger than the estimated residual variance in the DGP. Hence, the estimated standard error on the coefficient of a variable included in that model is larger than the corresponding coefficient's estimated standard error in the DGP. That shrinks the coefficient's t-ratio in the model with the omitted variable.

For example, the t-ratio for I_{1t} in Model M1 uses $\hat{\sigma}_1$ in the coefficient's estimated standard error, rather than $\hat{\sigma}$, which would be used for its t-ratio in Model M0. Thus, I_{1t} might be significant in Model M0 but appear insignificant in Model M1, simply because Model M1 excludes I_{2t} and so $\hat{\sigma}_1 > \hat{\sigma}$. Likewise, the t-ratio for I_{2t} in Model M2 uses $\hat{\sigma}_2$ in the coefficient's estimated standard error, rather than $\hat{\sigma}$. Hence, I_{2t} might be significant in Model M0 but appear insignificant in Model M2 because Model M2 excludes I_{1t} and so $\hat{\sigma}_2 > \hat{\sigma}$. As Hendry and Doornik (2014, p. 243) summarize, "[w]hen there is more than a single break, a failure to detect one [break] increases the residual variance and so lowers the probability of detecting any others."

Empirical application. Gamber and Liebner (2017) discuss t-ratios, significance levels, and empirical power for IIS, illustrating with the CBO forecasts. To interpret these empirical results in an encompassing framework, consider a baseline specification that includes all seven impulse indicators selected in Ericsson (2017). The observed t-ratios on retained impulses in Gamber and Liebner's models are closely matched by t-ratios as numerically solved from an encompassing analysis that starts with that baseline seven-indicator model. This comparison appears in Table 1. Moreover, the retention (or not) of individual impulse indicators in Gamber and Liebner (2017) is consistent with the losses in power implied by the encompassing analysis.

Key empirical results can be summarized, as follows. Using the "bare-bones" implementation of IIS, Gamber and Liebner (2017, Section 3) detect the following impulse indicators in the second subsample (1998–2012):

(a) I_{2001}, I_{2008}, and I_{2009} (at a 1% significance level);
(b) I_{2008} only (at a 1% significance level, but re-selected from (a)); and
(c) I_{2001}, I_{2002}, I_{2003}, I_{2008}, I_{2009}, and I_{2010} (at a 5% significance level).

For the first subsample (1984–1997), Gamber and Liebner find that:

(d) I_{1990} is not significant, nor is any other impulse indicator.

Columns ##1–4 in Table 1 report the t-ratios from (a)–(d). Using IIS in Autometrics, Ericsson (2017, Table 3) detects seven impulse indicators:

[1] This analysis and its empirical application below ignore changes in the estimated coefficients that arise from the omitted impulse indicators. However, because impulse indicators are orthogonal, those changes should not be an important consideration here.

Please cite this article in press as: Ericsson, N. R., Interpreting estimates of forecast bias. International Journal of Forecasting (2017), http://dx.doi.org/10.1016/j.ijforecast.2017.01.001

Download English Version:

https://daneshyari.com/en/article/5106339

Download Persian Version:

https://daneshyari.com/article/5106339

Daneshyari.com