Econometrics and Statistics 000 (2016) 1–22

Contents lists available at ScienceDirect

Econometrics and Statistics

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

Simple robust tests for the specification of high-frequency predictors of a low-frequency series

J. Isaac Miller

Department of Economics, University of Missouri, 118 Professional Building, Columbia, MO 65211, United States

ARTICLE INFO

Article history:
Received 18 January 2016
Revised 10 September 2016
Accepted 10 September 2016
Available online xxx

JEL Classification:
C12
C22
Q41

Keywords:
Temporal aggregation
Mixed-frequency model
MIDAS
Variable addition test
Forecasting model comparison
Retail gasoline prices

ABSTRACT

Two variable addition test statistics aimed at the specification of a high-frequency predictor of a series observed at a lower frequency are proposed. Under the null, the high-frequency predictor is aggregated to the low frequency versus mixed-frequency alternatives. The first test statistic is similar to those in the extant literature, but its robustness to conditionally biased forecast error and cointegrated and deterministically trending covariates is shown. It is feasible and consistent even if estimation is not feasible under the alternative. However, its size is not robust to nuisance parameters when the high-frequency predictor is stochastically trending, and size distortion may be severe. The second test statistic is a simple modification of the first that sacrifices power in order to correct this distortion. An application to forecasting and nowcasting monthly state-level retail gasoline prices illustrates how the test statistics may be utilized when the presence of nuisance parameters and orders of integration are unknown.

1. Introduction

The possibility of deleterious effects from aggregation on the properties of time series and forecasting models have long been known (e.g., Lütkepohl, 1987; Working, 1960). In order to circumvent such problems and exploit both higher frequency and real-time data in forecasting, a large number of recent papers have used a mixed-frequency (MF) approach, exploiting high-frequency (HF) data – usually financial data – to predict low-frequency (LF) data – usually macroeconomic data. These papers include Tay (2007), Clements and Galvão (2008), Clements and Galvão (2009), Hogrefe (2008), Schumacher and Breitung (2008), Armesto et al. (2010), Andreou et al. (2011), Andreou et al. (2013), Marcellino and Schumacher (2010), Kuzin et al. (2011), Götz et al. (2014), Miller (2014) and Breitung and Roling (2015).

Many of these papers employ a parsimonious distributed lag (DL) structure under the MIDAS moniker, for "MIxed DAta Sampling" (Ghysels et al., 2004; 2006). At one extreme, the most parsimonious DL structure reduces to temporal aggregation or systematic sampling of the HF data. With only a single coefficient to estimate in that case, a researcher must necessarily decide how to choose the vector of aggregation weights. A specific aggregation weight vector is the "fixed" null of the tests proposed in this paper. At the other extreme, the least parsimonious, an unrestricted DL specification, has at least as many parameters to estimate as the ratio m of frequencies in the model. When this ratio is large – e.g., for daily predictors of an annual target – the unrestricted specification is infeasible.

E-mail address: millerjisaac@missouri.edu

http://dx.doi.org/10.1016/j.ecosta.2016.09.001

The problem of specifying a mixed-frequency DL model is not limited to forecasting. Any regression model with a LF regressand and HF regressors may also have data of different types (stocks vs. flows) or different types of aggregation. A HF data-generating process coupled with a particular aggregation scheme for the regressand may suggest a particular scheme for the regressors, in which case no test is needed. That is not generally the case in forecasting, where optimality is attained by minimizing a loss function, such as mean squared forecast error. The aggregation weight vector that does so may not have any structural interpretation at all, in which case a test is useful.

In forecasting models, out-of-sample testing of predictive accuracy along the lines of West (1996) and Clark and McCracken (2001) provides one approach to this problem. In contrast, Diebold (2015) emphasizes the advantage of using in-sample tests to compare forecasting models in order to exploit the full sample available to a researcher. The Diebold (2015) approach may be viewed as pretesting in the sense that the tests are conducted before making the forecasts, while the out-of-sample tests are conducted after making forecasts.

In the recent MF literature, Andreou et al. (2010) propose two statistics for the null of aggregation with equal weights against a MIDAS alternative, while Kvedaras and Zemlys (2012) and Miller (2014) propose statistics for a MIDAS null against the unrestricted DL alternative. One of the statistics proposed by Andreou et al. (2010) and that proposed by Miller (2014) are variable addition test (VAT) statistics. The use of VAT statistics for specification testing is often attributed to Wu (1973) and is of course utilized in a wide variety of econometric settings, most commonly for testing instrumental variable and panel data specifications (Hausman, 1978), but also for testing cointegration (Park, 1990), for example.

A strong advantage of the VAT framework is its simplicity in implementation, requiring only two steps involving least squares. Another particularly useful aspect of the VAT approach in the context of MF series is that the test is implemented without having to estimate the model under the alternative – an impossible task against an unrestricted alternative when m is large relative to the LF sample size T.

In this paper, I introduce two simple VAT statistics for the null of a fixed aggregation scheme against a MF scheme, such as MIDAS. The most common aggregation schemes are flat sampling (equal weights to all HF observations) and skip sampling (end-of-period or beginning-of-period only), and tests with these nulls are useful to rule them out in favor of a model that better employs HF data. I show that when the HF predictor is I(0), the first VAT statistic, which is similar to that of Andreou et al. (2010) and Miller (2014), is robust to conditionally biased forecast error, infeasible alternatives, and deterministic and stochastic trends in the model covariates as long as the model specification is such that the forecast error remains I(0).

The possibility that the HF predictor is I(1) – and that the limiting distribution of the VAT statistic contains nuisance parameters as a result – motivates a second, modified VAT statistic. The modified VAT statistic augments the regressand by a term that increases mildly with the sample size. This strategy corrects size distortion from nuisance parameters in the limiting distribution, but at the expense of power. Although power generally decreases with the modification, the test remains consistent against fixed and some local alternatives.

Using the second test, the order of integration and number and/or structure of cointegrating relationships do *not* need to be known. Pretesting for unit roots and cointegration is not required, although a number of recent papers have made advances in testing for these in MF and LF environments. For example, Chambers (2015) analyzes tests for unit roots using LF data, and Ghysels and Miller (2015) and Miller and Wang (2016) analyze tests for cointegration using LF and MF data.

When the predictor is known to be I(0), the modified VAT statistic does not suffer from any size distortion. Small-sample results suggest that the power loss resulting from the modification is too large to justify the modification. Hence, the unmodified VAT statistic should be used. In contrast, small-sample results reveal size distortions from rejection rates up to 100%(!) using the unmodified VAT statistic test with an I(1) predictor and conditionally biased forecast error, but they show little (if any) power loss from the modification. If we can safely rule out unit roots or conditionally biased forecast error, the unmodified VAT statistic is quite appropriate. Otherwise, the modified VAT statistic exhibits much better size control and the expense of power loss may be acceptable even against local alternatives – especially when m is relatively small.

As a demonstration of the proposed tests, I consider an application to forecasting and nowcasting monthly available gasoline prices for the state of Missouri using weekly Midwest retail prices and daily Gulf Coast spot prices. Such an exercise would be valuable to policy makers considering a state-wide excise tax. At the time of this writing, May 2016, the Missouri Senate has in fact voted to hold a referendum on increasing the current tax of $0.17/gallon, one of the lowest in the US, by $0.059/gallon. Not surprisingly, I find that the most recent weekly or daily prices are the best predictors of future monthly Missouri prices. In contrast, the specification chosen by tests of a nowcasting model is more complicated and may reflect the construction of the monthly series.

The rest of the paper is organized as follows. In Section 2, I detail a general MF forecasting model, in which (at least) one of the predictors is available at a higher frequency than the target, and I motivate and present the two test statistics. Analysis of the large-sample properties of the test statistics – a limiting χ-squared distribution under the null and consistency under the alternative – follows in Section 3. Section 4 explores the small-sample size and power of the statistics under a range of assumptions, while Section 5 introduces a sequential testing strategy and applies that strategy to Missouri retail gasoline prices. Section 6 concludes. A technical appendix contains proofs of useful lemmas and the main theoretical results.