



Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: [www.elsevier.com/locate/jeconom](http://www.elsevier.com/locate/jeconom)

# Rolling window selection for out-of-sample forecasting with time-varying parameters



Atsushi Inoue<sup>a,\*</sup>, Lu Jin<sup>b</sup>, Barbara Rossi<sup>c,d</sup>

<sup>a</sup> Department of Economics, Vanderbilt University, Nashville, TN 37235-1819, United States

<sup>b</sup> StataCorp LP, 4905 Lakeway Drive, College Station, TX 77845-4512, United States

<sup>c</sup> ICREA-Universitat Pompeu Fabra, Barcelona GSE and CREI, Carrer Ramon Trias Fargas 25–27, 08005 Barcelona, Spain

<sup>d</sup> CEPR, United Kingdom

## ARTICLE INFO

### Article history:

Received 14 June 2014

Received in revised form

1 January 2016

Accepted 24 March 2016

Available online 7 October 2016

### JEL classification:

C14

C22

C53

### Keywords:

Macroeconomic forecasting

Parameter instability

Nonparametric estimation

Bandwidth selection

## ABSTRACT

There is strong evidence of structural changes in macroeconomic time series, and the forecasting performance is often sensitive to the choice of estimation window size. This paper develops a method for selecting the window size for forecasting. Our proposed method is to choose the optimal size that minimizes the forecaster's quadratic loss function, and we prove the asymptotic validity of our approach. Our Monte Carlo experiments show that our method performs well under various types of structural changes. When applied to forecasting US real output growth and inflation, the proposed method tends to improve upon conventional methods, especially for output growth.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Parameter instability is widely recognized as a crucial issue in forecasting (Stock and Watson, 1996; Rossi, 2013; Giacomini and Rossi, 2009; Paye and Timmermann, 2006; Koop and Potter, 2004; Goyal and Welch, 2003; Clements and Hendry, 1998). The empirical evidence of parameter instability is widespread in financial forecasting (Goyal and Welch, 2003), exchange rate prediction (Schinasi and Swamy, 1989; Wolff, 1987), and macroeconomic forecasting (Stock and Watson, 1996, 2003, 2007), to name a few. To handle such instability, instead of using all available observations, it is quite common to use only the most recent observations estimate the parameters (the so-called “rolling estimation” method). Examples of rolling estimation include: in finance, Goyal and Welch (2003) to evaluate the power of dividend ratios in predicting stock market returns and the equity premium; in macroeconomics, Swanson (1998) to investigate the extent

to which fluctuations in the money stock predict fluctuations in real income; in exchange rate forecasting, Molodtsova and Papell (2009) to investigate the predictability of models that incorporate Taylor rule fundamentals for exchange rate.

In rolling out-of-sample forecasting, one produces a sequence of pseudo out-of-sample forecasts using a fixed number of the most recent data at each point of time. One practical issue with rolling out-of-sample forecasting is how many recent observations should be used in the estimation. The number of the recent observations used in estimation is referred to as the window size. Conventionally, the window size is either arbitrarily determined by forecasters or based on past experience. For instance, Molodtsova and Papell (2009) use a 10-year window of monthly data to predict exchange rates; and Stock and Watson (2007) forecast inflation with a 10-year window of quarterly observations. However, we often find that the forecasting performance of the rolling scheme is sensitive to the choice of the window size (see Inoue and Rossi, 2012).

While the problem of selecting the estimation window size is similar to the problem of bandwidth selection in nonparametric estimation, methods to select the window size in rolling out-of-sample forecasting have received little attention. Among recent

\* Corresponding author. Fax: +1 615 343 8495.

E-mail addresses: [atsushi.inoue@vanderbilt.edu](mailto:atsushi.inoue@vanderbilt.edu) (A. Inoue), [jl817cn@gmail.com](mailto:jl817cn@gmail.com) (L. Jin), [barbara.rossi@upf.es](mailto:barbara.rossi@upf.es) (B. Rossi).

<http://dx.doi.org/10.1016/j.jeconom.2016.03.006>

0304-4076/© 2016 Elsevier B.V. All rights reserved.

papers focusing on how to determine the optimal window size: Pesaran and Timmermann (2007) propose five methods to select the window size when the forecasting model is subject to one or multiple discrete breaks; Pesaran et al. (2013) derive optimal weights under continuous and discrete breaks; and Giraitis et al. (2013) develop a cross-validation-based method to select a tuning parameter to downweight older data in the presence of structural change.

In this paper, we develop a new approach for selecting the size of the rolling estimation window for forecasting in models with potential breaks. More specifically, parameters are specified as smooth functions of time and the functional forms are unknown. This setting, in which structural changes may occur in every point in time and are small, is consistent with empirical findings of small instability in some forecasting areas, such as forecasting inflation (Stock and Watson, 1999). This setup is also adopted in the nonparametric literature, for example, Robinson (1989), Cai (2007) and Chen and Hong (2012).

Our approach has three advantages over existing methods. First, the error term and the regressors can be weakly dependent, and the regressors can include both exogenous and lagged dependent variables, while existing methods rely on more stringent assumptions. The five window selection methods developed in Pesaran and Timmermann (2007) require serially uncorrelated errors and strictly exogenous regressors. Pesaran et al.'s (2013) approach needs independent errors and exogenous regressors. Giraitis et al. (2013) focus on models without regressors. Thus, our approach can be used for a wider range of forecasting models than existing methods. Second, our approach allows multiple-step-ahead forecasting, while existing methods only consider one-step-ahead. Third, we propose a feasible solution to approximate forecasters' quadratic loss function, and we also prove the asymptotic validity of this feasible approximation.

Our new approach proposes to choose the optimal window size that minimizes the conditional mean square forecast error (MSFE), which is commonly used as the forecasters' loss function. Since the conditional MSFE is infeasible, we construct an approximate conditional MSFE by replacing the unknown parameters in the conditional MSFE with estimates from local linear regressions, and then choose the window size that minimizes this approximate conditional MSFE. We show that choosing the optimal window size based on our approximate criterion is asymptotically equivalent to choosing the window size based on the infeasible one. Choosing the window size for the conditional MSFE as opposed to the integrated MSFE and establishing its asymptotic justification under the aforementioned general framework are our new contributions to the literature. Our Monte Carlo simulations suggest that using the window size selected by our procedure can improve upon the forecasting performance vis-à-vis an ad-hoc choice of the window size.

Moreover, we empirically assess the practical value of our procedure in forecasting real output growth and inflation. As shown in Stock and Watson (2003, 2007), the predictive ability of standard forecasting models suffers from instability; that is, finding a predictor useful at one point in time does not guarantee that the same predictor will forecast well in later periods. In our empirical analysis, we examine whether we can improve forecasts by using our proposed window selection procedure. Our results suggest that asset prices, unemployment and monetary measures have useful predictive content for forecasting output growth at short horizons. When forecasting inflation, measures of unemployment are useful at long horizons, confirming the usefulness of the unemployment-based Phillips curve for inflation forecasting in the presence of parameter instability. In general, the forecast improvements generated by the optimal window size are more substantial when forecasting output growth than inflation,

since, as we show, parameters are more likely to vary in the former than in the latter.

When the optimal window sizes are used, the number of building permits has useful predictive content for long-term output growth forecasts, and measures of unemployment are useful for inflation forecasts. One possible economic interpretation is that building constructions typically take a long time to complete, so investment in the construction sector has a long-term effect on output growth. The unemployment-based Phillips curve is useful in predicting inflation, possibly because the non-accelerating inflation rate of unemployment (NAIRU) is unstable and the optimal window size captures time variation.

The rest of the paper is organized as follows. Section 2 presents a model, motivates our problem and describes our proposed window selection procedure. Section 3 provides theoretical justifications for our window selection procedure. Section 4 reports Monte Carlo simulation results. Section 5 applies our procedure to forecasting output growth and inflation in the United States, and Section 6 concludes. The appendix provides lemmas and proofs of the theorems in Section 3. The working paper version of this paper, Inoue et al. (2015, hereafter IJR), includes extended Monte Carlo and empirical results as well as the proof of the lemmas.

## 2. Motivation and setup

Assume the data generating process (DGP) is:

$$y_{t+h} = \beta'_{h,t} x_t + u_{t+h}, \quad t = 1, 2, \dots, T, \quad (1)$$

where  $x_t$  is a  $p \times 1$  vector of stochastic regressors,  $\beta_{h,t}$  is a  $p \times 1$  vector of time-varying parameters;  $u_{t+h}$  is an unobservable disturbance;  $h$  denotes the forecast horizon; and  $T$  denotes the full sample size. The regressor vector  $x_t$  may include exogenous explanatory variables and lagged values of the dependent variable. Our interest is to predict  $y_{T+h}$  using information available at time  $T$ .

As in Robinson (1989) and Cai (2007), the time variation in the parameters is represented by a smooth function of the current period  $t$ . Thus Eq. (1) can be rewritten as:

$$y_{t+h} = \beta_h(t/T)' x_t + u_{t+h}, \quad (2)$$

where  $\beta_h(t/T)$  is a  $p \times 1$  vector of unknown functions of  $t$  that is defined on an equally spaced grid over  $(0, 1]$ , which becomes finer as  $T \rightarrow \infty$ . According to Robinson (1989), this requirement is important for deriving consistent nonparametric estimates, since the amount of local information on which an estimator depends increases suitably as the sample size  $T$  increases.

The rolling OLS estimator is commonly used in forecasting because parameters are often found to be time-varying. While the rolling OLS estimator may look like a parametric estimator, it is a local constant estimator and thus it is a nonparametric estimator of  $\beta_h(\cdot)$  in Eq. (2), where the estimation window size plays the role of the bandwidth.<sup>1</sup>

We focus on how to determine the size of the estimation window for forecasting in the framework described above. Our new approach chooses the optimal window size that minimizes the conditional MSFE. The conditional MSFE is a commonly used

<sup>1</sup> The rolling window estimator is a local constant estimator with the truncated kernel that assigns 0–1 to the observations. While such weights may not be optimal, we focus on the rolling window estimator because it is widely used in practice. We refer to Pesaran et al. (2013) for the analysis of optimal weights.

Download English Version:

<https://daneshyari.com/en/article/5095654>

Download Persian Version:

<https://daneshyari.com/article/5095654>

[Daneshyari.com](https://daneshyari.com)