ARTICLE IN PRESS

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

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Interpretation and use of sensitivity in econometrics, illustrated with forecast combinations

Jan R. Magnus[a], Andrey L. Vasnev[b,*]

[a] Department of Econometrics & Operations Research, VU University Amsterdam, The Netherlands

[b] University of Sydney, New South Wales, Australia

ARTICLE INFO

Keywords:
Sensitivity analysis
Reconciling forecasts
Yield curve prediction

ABSTRACT

Sensitivity analysis is important both for its own sake and in combination with diagnostic testing. We consider the question of how to use sensitivity statistics in practice, and in particular, how to judge whether the sensitivity is large or small. For this purpose, we distinguish between absolute and relative sensitivity, and highlight the context-dependent nature of sensitivity analysis. The relative sensitivity is then applied to forecast combinations, and sensitivity-based weights are introduced. All of the concepts are illustrated using the European yield curve. In this context, it is natural to consider the sensitivity to autocorrelation and normality assumptions. Different forecasting models are combined using equal, fit-based and sensitivity-based weights, and compared with the multivariate and random walk benchmarks. We show that the fit-based and sensitivity-based weights are complementary, but that the sensitivity-based weights perform better than other weights for long-term maturities.

1. Introduction

The majority of papers in applied econometrics concentrate on the fit of the models and the statistical significance of the coefficients, while sensitivity analysis is often not reported at all, or is reported only tangentially. This is unfortunate, because sensitivity analysis is at least as important as diagnostic testing. While diagnostic testing attempts to answer the question: is it true? (for example, that a coefficient is zero), sensitivity analysis addresses the question: does it matter? (that we set the coefficient to zero). At first glance, the two questions seem to be closely related, but Magnus and Vasnev (2007) showed that such is not the case. In fact, the two concepts are essentially orthogonal.

Fig. 1 shows the potential danger of ignoring sensitivity. The sample is given by three points, (x_1, y_1), (x_2, y_2), and (x_3, y_3), and two models are fitted. The horizontal line, given by the average value of the dependent variable $\bar{y} = (y_1 + y_2 + y_3)/3$, provides a minimal fit, but it is not sensitive to autocorrelation, non-normality, or other model assumptions. The other model provides a perfect fit, but it can only be used in a very small neighborhood of the sample points. It is unstable outside the data range $[x_1, x_3]$, and even within this range it produces unjustified values that are bigger than the maximum in the observed data. In this situation, the simple non-sensitive model is more reliable.

There are also situations in which one might be interested in a model with a high rather than a low sensitivity. For example, if we are interested in detecting a crisis or abnormalities in the market, then we prefer a model which is maximally sensitive, even to small indications of a crisis.

Magnus and Vasnev (2007) provide an overview of the sensitivity literature, and prove the asymptotic independence of the commonly-used diagnostic tests and the sensitivity statistic formally. Diagnostic tests and sensitivity statistics are therefore complementary, and

* Corresponding author.
E-mail address: andrey.vasnev@sydney.edu.au (A.L. Vasnev).

0169-2070/$ – see front matter Crown Copyright © 2013 Published by Elsevier B.V. on behalf of International Institute of Forecasters. All rights reserved.
http://dx.doi.org/10.1016/j.ijforecast.2013.08.001

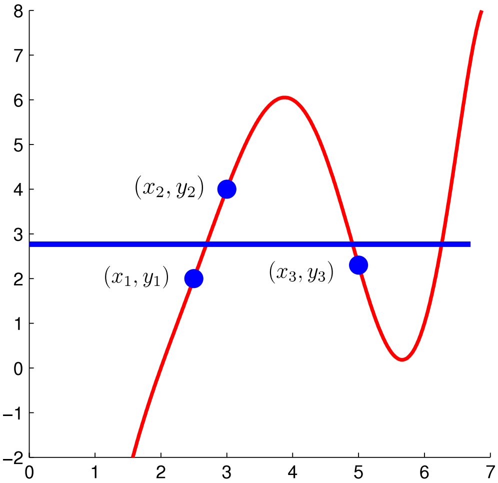

Fig. 1. The sample is given by three points. The horizontal line provides a minimal fit, but is not sensitive to model assumptions. The curve gives a perfect fit, but is very unstable.

both require our attention when analyzing a model. It is possible to derive sensitivity statistics, and several papers have suggested local and global sensitivity measures. However, it is more difficult to answer the question of when a sensitivity statistic is large or small, a question which is addressed in the current paper. The paper gives practical recommendations with regard to the way in which sensitivity statistics can be used. We shall see that the use of sensitivity is context-dependent, as is also emphasized by Severini (1996), so that we need to consider sensitivity in relation to the problem under consideration.

In some situations, the value of the sensitivity statistic is important, requiring a threshold in order to decide whether the model is sensitive or not; we call this case 'absolute sensitivity'. In other situations, only the relative magnitude is important, and we call this case 'relative sensitivity'. In both cases, it is essential to realize that sensitivity (unlike a diagnostic test) is context-dependent, and will be closely related to the estimator we are analyzing or the dependent variable we are modeling. To bring out this dependence, we illustrate all of the concepts introduced in this paper in a specific application, namely the forecasting of the Euro yield curve.

We show that, when several forecasts are available, the weights based on relative sensitivity perform well, and are complementary to the fit-based weights. The main purpose of combining forecasts is to improve the forecast accuracy (Bates & Granger, 1969), but the choice of weights is still an open question. Timmermann (2006) provides a thorough overview of the sizable forecast combination literature, but in practice the optimal weights have to be estimated, and this affects their actual performance. The adaptive weights seem to work well in many situations, but sometimes a simple alternative with equal weights gives better results, as was shown by Stock and Watson (2004). This fact is explained by Winkler and Clemen (1992) as being due to the instability of the estimated weights used in generating the combined forecast.

The paper is organized as follows. Section 2 introduces the practical aspects of sensitivity analysis and provides a brief overview of the sensitivity literature. It highlights the context-dependent nature of sensitivity analysis (Section 2.1), and distinguishes between absolute (Section 2.2) and relative (Section 2.3) sensitivity. Section 3 applies the concept of relative sensitivity to forecast combinations, and introduces sensitivity-based weights. The empirical Euro yield curve illustration is given in Section 4, and a detailed description of the data is given in the Data Appendix. Section 5 concludes.

2. Sensitivity analysis in practice

The concept of sensitivity is closely related to the concept of robustness, with which readers may be more familiar. Robustness was introduced by Hampel, Ronchetti, Rousseeuw, and Stahel (1986), and has since been studied extensively in the literature; see Kitamura, Otsu, and Evdokimov (2013) for a recent contribution. Robustness deals primarily with the effects of slight perturbations in the observed data, and often uses the influence function as a tool. Sensitivity analysis deals not only with data perturbations, but also with model perturbations. Sensitivity to the model specification studies the changes in the model output (often an estimator, test or predicted value) when one or more of the assumptions underlying the model are perturbed. The simplest example is given by Banerjee and Magnus (1999), who studied the sensitivity of the ordinary least squares estimator to autocorrelation in the regression errors. In this paper, we concentrate on model sensitivity, and refer to it simply as sensitivity.

Magnus and Vasnev (2007) introduced local sensitivity through a Taylor expansion. If the variable (or parameter) of interest, say y, depends on a nuisance parameter, say θ, then $\hat{y}(\theta)$ denotes the estimator of y for each given value of θ. Special cases are the 'restricted' estimator $\hat{y}(0)$, obtained by setting $\theta = 0$, and the 'unrestricted' estimator $\hat{y}(\hat{\theta})$, obtained by setting θ equal to its estimated value $\hat{\theta}$. The function $\hat{y}(\theta)$ provides not only these two special cases, but the whole *sensitivity curve*, given by the estimates of y for each given value of θ.

The first-order Taylor expansion of the sensitivity curve at the restricted point is given by

$$\hat{y}(\theta) = \hat{y}(0) + S\,\theta + O(\theta^2), \tag{1}$$

where

$$S = \left.\frac{\partial \hat{y}(\theta)}{\partial \theta}\right|_{\theta=0} \tag{2}$$

is the first derivative at the restricted point $\theta = 0$, and is called the *local sensitivity statistic*, or simply the sensitivity.

Sensitivity is computed for maximum likelihood estimators by Magnus and Vasnev (2007), and, in general, it can be expressed in terms of the Hessian. In the cases of mean, variance, and distribution misspecification, the sensitivity statistics allow tractable representations. This is particularly the case for the B_s and D_s statistics of Banerjee and Magnus (1999), and the sensitivity of GLS estimators in panel data derived by Vasnev (2010).

One might think that the sensitivity statistic and the corresponding diagnostics would be highly correlated.