Benchmarking judgmentally adjusted forecasts

Philip Hans Franses

Bert de Bruijn

Econometric Institute

Erasmus School of Economics

Econometric Institute Report 2015-36

Abstract

Many publicly available macroeconomic forecasts are judgmentally-adjusted model-based forecasts. In practice usually only a single final forecast is available, and not the underlying econometric model, nor are the size and reason for adjustment known. Hence, the relative weights given to the model forecasts and to the judgment are usually unknown to the analyst.

This paper proposes a methodology to evaluate the quality of such final forecasts, also to allow learning from past errors. To do so, the analyst needs benchmark forecasts. We propose two such benchmarks. The first is the simple no-change forecast, which is the bottom line forecast that an expert should be able to improve. The second benchmark is an estimated model based forecast, which is found as the best forecast given the realizations and the final forecasts. We illustrate this methodology for two sets of GDP growth forecasts, one for the US and for the Netherlands. These applications tell us that adjustment appears most effective in periods of first recovery from a recession.

JEL code: C20, C51

Key words: Forecast decomposition; Expert adjustment; Total Least Squares

This version: November 2015

Thanks are due to Richard Paap, Christiaan Heij and Tom Wansbeek for various helpful comments.

Introduction

Many publicly available macroeconomic forecasts are judgmentally-adjusted model-based forecasts. Econometric models can be multiple-equation systems with hundreds of variables or identities, or Bayesian vector autoregressions or even simple extrapolation tools. An illustration of the first is given in Franses, Kranendonk and Lanser (2011), where all the forecasts from the large macroeconomic model of the Netherlands Bureau for Economic Policy Analysis (CPB) are manually adjusted by experts with domain-specific knowledge.

In many situations it can be beneficial to adjust model-based forecasts. When experts foresee that a prediction error is to be made with the model, then adjustment can help to improve accuracy. For example, adjustment can be needed due to measurement issues in the explanatory variables at the forecast origin or due to anticipated changes, not included in the model at the forecast origin.

Despite the potential success of expert adjustment it is rarely documented what an expert does and why certain decisions have been made. This hampers a straightforward evaluation of forecast errors, as it is usually unknown which part of the error could be due to the econometric model and which part to the manual adjustment. In other words, the relative weights given to the econometric model forecasts and to the judgment are usually unknown to the analyst.

In this paper we propose a methodology that allows to study the merits of the relative contribution of an expert. In fact, our methodology allows to indicate when, that is, for which years or quarters, did the expert make the final forecast better than an underlying model forecast and when did the expert touch harm that forecast quality? For this methodology we need benchmark econometric model forecasts. Now, typically, one resorts to the simplest benchmark possible, and this is the no-change forecast, see Vuchelen and Gutierrez (2005) and also recently Franses and Maassen (2015). The idea is that an expert would not show much expertise if this trivial forecast cannot be beaten. In the present paper we additionally propose another benchmark forecast, and this associates with in some sense a “best model-based” forecast. We derive this best forecast from the final forecasts and the realizations, and use the technique called Total Least Squares (TLS), which here in our setting of forecasts and realizations boils down to the so-called Deming regression (Deming, 1943). We illustrate our methodology using two sets of forecasts for growth in Gross Domestic Product (GDP), one for the Netherlands and one for the USA. Zooming in on successful contributions of the experts we find that they have in common that they have been particularly successful in the first periods of recovery from a recession as then the experts' added valuable information to the model forecast.

The outline of our paper is as follows. Section 2 introduces the two benchmark model-based forecasts, where most attention will be given to the "best model-based" forecast. Section 3 presents a detailed illustration of our methodology, and Section 4 concludes.

Benchmark model-based forecasts

When an analyst wants to evaluate the quality of forecasts, say from the IMF, OECD, the World Bank, or, as in our illustration below, wants to analyse the qualities of the Econometric Institute Current Indicator of the Economy (EICIE), then a benchmark is needed. In some situations, typically in business forecasting, there is the availability of the actual model-based forecasts, see Franses (2014) for a review, but in many other situations, typically in macroeconomics, such model-based forecasts are not available.

The no-change forecast

A first and simple benchmark forecast is of course the no-change forecast. That is, if we consider a variable y_t that needs to be predicted, then the one-step-ahead no-change forecast is y_{t-1}.

Denoting the final expert-adjusted forecast as f_t, Vuchelen and Gutierrez (2005) advocate the use of this no-change forecast in their auxiliary regression

$$y_t = \mu + \gamma_1 y_{t-1} + \gamma_2 (f_t - y_{t-1}) + \varepsilon_t \tag{1}$$

where they advocate a Wald test for the composite null hypothesis that $\mu = 0, \gamma_1 = 1, \gamma_2 = 1$. Under this null hypothesis, the model-based forecast is unbiased and the expert-adjustment on top of that no-change forecast is then unbiased too. If the null hypothesis is rejected, one can have a closer look at the estimated parameter values of γ_1 and γ_2.

The best model-based forecast

To arrive at a method to retrieve an estimator of the "best model-based" forecast, we somehow need to make assumptions. A first assumption is that an observed expert-adjusted forecast f_t is a forecast of a variable y_t^*, which is the true variable of interest, but that this true variable is measured with error, hence y_t. Next, we assume that f_t amounts to a concerted outcome of an econometric model forecast f_t^M and an expert touch f_t^E, with

$$f_t = f_t^M + f_t^E. \tag{2}$$

The third assumption is that f_t^E and f_t^M are independent. This assumption corresponds with an optimal situation, as when it does not hold, the expert is adding something to the model forecast that is already in there, and this amounts to double counting.

Our simple method to estimate f_t^M and f_t^E from f_t and the realizations y_t relies on the familiar regression

$$y_t^* = \alpha + \beta f_t^M + \varepsilon_t \tag{3}$$

which is usually used to test if $\alpha = 0$ and $\beta = 1$, where these parameter values associate with unbiased forecasts. Our method is now based on the assumption that the two variables in (3) are measured with error. First, as mentioned, for y_t^* we assume that

$$y_t = y_t^* + w_t \tag{4}$$

where w_t has variance σ_w^2 and where w_t is independent from y_t^* and the ε_t in (3). For f_t^M we introduce a measurement error via (2), that is, $f_t = f_t^M + f_t^E$, which thus treats the expert touch as a measurement error. The f_t^E has variance σ_E^2, f_t^M has variance σ_M^2 and, as said, we further assume that f_t^E and f_t^M are independent, so the variance of f_t is $\sigma_F^2 = \sigma_M^2 + \sigma_E^2$.

For practical purposes it is interesting to estimate f_t^E and f_t^M, and in particular the variances σ_E^2 and σ_M^2. It is also important to study the model-based forecast errors $y_t - f_t^M$ versus $y_t - f_t$ to learn about the contribution of the expert. That is, does the expert touch lead to better forecasts?

In sum, the key unobserved variable to estimate is f_t^M using data on y_t and f_t. We now propose a methodology to do so. The key problem that we face is estimating f_t^M, given that the true regression model is $y_t^* = \alpha + \beta f_t^M + \varepsilon_t$ and that the data are assumed to follow from $y_t = y_t^* + w_t$ and $f_t = f_t^M + f_t^E$, which is the case of measurement errors in two variables, the dependent and the independent variables. There are many techniques available which usually focus on obtaining consistent estimators of α and β, see for example Koopmans (1937), Fuller (1987) and Wansbeek and Meijer (2000). One technique, which goes back to Frisch (1933), is particularly useful as it delivers a simple estimator to predict the values of f_t^M. This method is called Total Least Squares and it is also sometimes coined as the Deming regression (Deming, 1943).

An alternative least squares estimator for β is the Total Least Squares (TLS) estimator, which seeks to minimize the squares of the *orthogonal* distances to the regression line. It is thus assumed that part of the error in the regression model corresponds with a measurement error in the dependent variable. Define

$$\delta = \frac{\sigma_\varepsilon^2 + \sigma_w^2}{\sigma_E^2} \tag{5}$$

see Carroll and Ruppert (1996), and define $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$ and $\bar{f} = \frac{1}{T}\sum_{t=1}^{T} f_t$, where T is the number of one-step-ahead forecasts. The TLS estimators for β and α now converge to

$$\hat{\beta}_{TLS} \rightarrow \frac{\sigma_y^2 - \delta\sigma_F^2 + \sqrt{(\sigma_y^2 - \delta\sigma_F^2)^2 + 4\delta\sigma_{Fy}^2}}{2\sigma_{Fy}} \tag{6}$$

$$\hat{\alpha}_{TLS} \rightarrow \bar{y} - \hat{\beta}_{TLS}\bar{f} \tag{7}$$

where we denote as σ_{Fy} the covariance between the observed series and its forecasts, see Deming (1943, page 184). In practice, these TLS estimators are of course based on the sample equivalents of the variances and covariance. The key feature of this method, which is relevant for our purposes, is that an interesting by-product of TLS is an estimator for the measurement-error-free explanatory variable, that is,

$$\hat{f}_t^M = f_t + \frac{\hat{\beta}_{TLS}}{\hat{\beta}_{TLS}^2 + \delta}(y_t - \hat{\alpha}_{TLS} - \hat{\beta}_{TLS} f_t) \tag{8}$$

see Linnet (1990). Our key assumption now is that we will coin this $\hat{f}_t^M$ as the “best model-based” forecast in our illustrations below.

The key parameter that one should set from the outset is δ in (5). Given our particular case of realizations and forecasts it may not be unreasonable to assume that $\sigma_w^2 = \sigma_E^2$. Then

$$\delta = \frac{\sigma_\varepsilon^2}{\sigma_E^2} + 1$$

Simulation results in Table 1 show that, in case the value of δ is known, the correlation between simulated f_t^M and estimated $\hat{f}_t^M$ ranges from around 0.8 to close to 1. The size of the unexplained part depends on the variances, and can range from 5% to close to 60%. The sample size does not seem to matter. Tables 2a and 2b present the results for the cases where the true value of δ is deliberately underestimated by a fraction $\frac{1}{2}$ and deliberately overestimated by a fraction 2, respectively. In general the correlations do not differ much from those in Table 1. For the explained part we see that overestimation leads to a slightly larger fraction of the unexplained part.

In the next section we apply our methodology to two cases, one concerning annually observed IMF forecasts for USA real GDP growth and one concerning quarterly forecasts for GDP growth in the Netherlands.

Illustrations

We first present the various relevant parameter estimates, and then turn to an evaluation of the forecast performance.

Benchmarks

Columns 2 and 3 of Table 3 presents the data on the Econometric Institute Current Index of the Economy (EICIE) (available from the website of the Erasmus School of Economics) and the second release data from Statistics Netherlands concerning year-to-year GDP growth observed per quarter. The available data range from 2004Q4 to 2015Q2. The second release data appear 90 days after the relevant quarter. The EICIE is published during the relevant quarter, and hence in fact amounts to a nowcast. Figure 1 gives a graphical impression of the data. An application of Ordinary Least Squares (OLS) to the regression model as in (3) for the observable data, that is,

$$y_t = \alpha + \beta f_t + \varepsilon_t \qquad (6)$$

gives $\hat{\alpha}$ = -0.582 with standard error 0.229, and $\hat{\beta}$ = 1.248 with standard error 0.124. The Wald test value for the joint hypothesis that $\alpha = 0, \beta = 1$ is 6.811, with a p value of 0.033. This suggests that the EICIE delivers biased forecasts.

This bias is reinforced by looking at the estimation results for the regression (1), see the second column of Table 5. The estimated γ_1 is quite close to 1, but the estimated γ_2 is not. The model fit is substantial (0.770), but the Wald test on $\mu = 0, \gamma_1 = 1, \gamma_2 = 1$ results in a p value of 0.001. Hence, on average, the added contribution of the expert, on top of a no-change model forecast, apparently does not improve the final forecast.

The estimated TLS parameters for the regression (3) appear in the left-hand side panel of Table 6. The variance σ_y^2 is estimated as 4.865, the variance σ_F^2 is 2.227, and the covariance between the CBS data (Statistics Netherlands) and the EICIE forecasts is estimated as 2.778. The average observed growth rate is 0.791 and the average nowcast is 1.1. Table 6 reports on the TLS estimates for various values of δ, ranging from 0.7 to 1.3. Clearly, the estimated parameter values do not change much across this range of δ.

Table 4 presents the IMF forecasts for US real GDP growth for the years 1991 to and including 2013, the columns 2 and 3. Figure 2 gives a graphical impression of the data. The right-hand column of Table 5 shows that final expert forecasts do add something relevant to the no-change forecasts, as the p value of the Wald test on $\mu = 0, \gamma_1 = 1, \gamma_2 = 1$ is 0.695. Moreover, the estimated value of γ_2 is 0.9543, which is quite close to 1. So, the contribution of the IMF experts is unbiased and relevant. The variance σ_y^2 is estimated as 3.024, the variance σ_F^2 is 0.610, and the covariance between the actual data and the IMF forecasts is estimated as 0.781. The average observed growth rate is 2.483 and the average forecast is 2.421. Table 6 reports on the TLS estimates for various values of δ, ranging from 0.7 to 1.3, and again the estimated parameter values do not change much across this range of δ.

Forecast performance

For further analysis, we now set $\delta = 1$. Table 7 presents the fraction of times that forecasts have the lowest absolute forecast error across three forecasts, that is, the final judgmentally adjusted forecast, the no-change forecast and the best-model forecast. As could be expected, and by creation, the best-model forecast is best in about half the cases across these sets of forecasts. The no-change forecast seems on average about equally good as the final expert forecast. But still, in 1 of 4 quarters or years, the expert touch does seem to improve on both benchmark forecasts.

Table 8 zooms in on the quarters and years where the expert forecasts were more accurate than the benchmarks. Clearly, the quarters 2009Q3, 2009Q4, 2012Q2 and 2012Q3 as well as the years 2002-2005 and 2013 are recovery quarters and years. So, it seems that the expert adjustment was most useful in these recovery periods. Apparently, econometric models can need the help of experts, particularly in these business cycle episodes.

Conclusion

We have proposed a simple methodology to benchmark final macroeconomic forecasts. This is necessary as those final forecasts are typically the combination of an econometric model-based forecast and a manual modification by an expert. The analyst usually does not know the specific weights in the combination. Illustrations to two sets of GDP growth forecasts showed the merits of the methodology.

Table 1: Average correlation between the predicted measurement-error-free explanatory variable and its true observations, and the percentage unexplained of the true observations. The setting is

$$y_t = y_t^* + w_t, \text{ with } \sigma_w^2$$
$$x_t = x_t^* + v_t, \text{ with } \sigma_v^2$$
$$\text{DGP: } y_t^* = -1 + 2x_t^* + \varepsilon_t, \text{ with } \sigma_\varepsilon^2$$

where $y_t^*, w_t, x_t^*, v_t, \varepsilon_t$ are draws from a N(0,1) distribution. Simulations are for samples T = 100 and 500, and the number of replications is 10000. It is assumed that δ is known.

	T	Correlation	Percentage unexplained
σ_w^2 = 1, σ_v^2 = 1, σ_ε^2 = 0, $\sigma_{x^*}^2$ = 1	100	0.912	22.4
	500	0.913	20.5
σ_w^2 = 1, σ_v^2 = 1, $\boldsymbol{\sigma_\varepsilon^2}$ **= 1** $\sigma_{x^*}^2$ = 1	100	0.864	36.2
	500	0.866	33.8
σ_w^2 = 1, σ_v^2 = 1**,** $\boldsymbol{\sigma_\varepsilon^2}$ **= 2**, $\sigma_{x^*}^2$ = 1	100	0.798	58.8
	500	0.801	56.1
$\boldsymbol{\sigma_w^2}$ **= 2**, σ_v^2 = 1, σ_ε^2 = 0, $\sigma_{x^*}^2$ = 1	100	0.813	52.9
	500	0.816	50.6
σ_w^2 = 1, $\boldsymbol{\sigma_v^2}$ **= 2,** σ_ε^2 = 0, $\sigma_{x^*}^2$ = 1	100	0.898	22.9
	500	0.899	25.6
σ_w^2 = 1, σ_v^2 = 1, σ_ε^2 = 0, $\boldsymbol{\sigma_{x^*}^2}$ **= 2**	100	0.976	5.5
	500	0.976	5.1

Table 2a: Average correlation between the predicted measurement-error-free explanatory variable and its true observations, and the percentage unexplained of the true observations. The setting is

$$y_t = y_t^* + w_t, \text{ with } \sigma_w^2$$
$$x_t = x_t^* + v_t, \text{ with } \sigma_v^2$$
$$\text{DGP: } y_t^* = -1 + 2x_t^* + \varepsilon_t, \text{ with } \sigma_\varepsilon^2$$

where $y_t^*, w_t, x_t^*, v_t, \varepsilon_t$ are draws from a N(0,1) distribution. Simulations are for samples T = 100 and 500, and the number of replications is 10000. **It is assumed that $\boldsymbol{\delta}$ is incorrectly specified as $\frac{1}{2}\boldsymbol{\delta}$**

	T	Correlation	Percentage unexplained
σ_w^2 = 1, σ_v^2 = 1, σ_ε^2 = 0, $\sigma_{x^*}^2$ = 1	100	0.907	20.6
	500	0.908	18.8
σ_w^2 = 1, σ_v^2 = 1, $\boldsymbol{\sigma_\varepsilon^2}$ **= 1** $\sigma_{x^*}^2$ = 1	100	0.852	31.6
	500	0.853	29.8
σ_w^2 = 1, σ_v^2 = 1**,** $\boldsymbol{\sigma_\varepsilon^2}$ **= 2**, $\sigma_{x^*}^2$ = 1	100	0.765	48.9
	500	0.767	47.0
$\boldsymbol{\sigma_w^2}$ **= 2**, σ_v^2 = 1, σ_ε^2 = 0, $\sigma_{x^*}^2$ = 1	100	0.786	44.7
	500	0.788	42.8
σ_w^2 = 1, $\boldsymbol{\sigma_v^2}$ **= 2,** σ_ε^2 = 0, $\sigma_{x^*}^2$ = 1	100	0.897	29.4
	500	0.898	22.1
σ_w^2 = 1, σ_v^2 = 1, σ_ε^2 = 0, $\boldsymbol{\sigma_{x^*}^2}$ **= 2**	100	0.974	5.6
	500	0.975	5.2

Table 2b: Average correlation between the predicted measurement-error-free explanatory variable and its true observations, and the percentage unexplained of the true observations. The setting is

$$y_t = y_t^* + w_t, \text{ with } \sigma_w^2$$
$$x_t = x_t^* + v_t, \text{ with } \sigma_v^2$$
$$\text{DGP: } y_t^* = -1 + 2x_t^* + \varepsilon_t, \text{ with } \sigma_\varepsilon^2$$

where $y_t^*, w_t, x_t^*, v_t, \varepsilon_t$ are draws from a N(0,1) distribution. Simulations are for samples T = 100 and 500, and the number of replications is 10000. **It is assumed that $\boldsymbol{\delta}$ is incorrectly specified as $\boldsymbol{2\delta}$**

	T	Correlation	Percentage unexplained
σ_w^2 = 1, σ_v^2 = 1, σ_ε^2 = 0, $\sigma_{x^*}^2$ = 1	100	0.898	33.2
	500	0.899	30.9
σ_w^2 = 1, σ_v^2 = 1, $\boldsymbol{\sigma_\varepsilon^2}$ **= 1** $\sigma_{x^*}^2$ = 1	100	0.840	54.1
	500	0.842	52.0
σ_w^2 = 1, σ_v^2 = 1**,** $\boldsymbol{\sigma_\varepsilon^2}$ **= 2**, $\sigma_{x^*}^2$ = 1	100	0.772	78.8
	500	0.774	76.7
$\boldsymbol{\sigma_w^2}$ **= 2**, σ_v^2 = 1, σ_ε^2 = 0, $\sigma_{x^*}^2$ = 1	100	0.787	73.8
	500	0.788	71.6
σ_w^2 = 1, $\boldsymbol{\sigma_v^2}$ **= 2,** σ_ε^2 = 0, $\sigma_{x^*}^2$ = 1	100	0.888	58.5
	500	0.890	47.2
σ_w^2 = 1, σ_v^2 = 1, σ_ε^2 = 0, $\boldsymbol{\sigma_{x^*}^2}$ **= 2**	100	0.973	6.6
	500	0.973	6.1

Table 3: The EICIE forecasts and realizations

Quarter	Actuals	EICIE	No change forecast	Best model forecast
2004Q4	1.6	1.1	NA	1.465293
2005Q1	-0.5	1.0	1.6	0.488282
2005Q2	1.3	-1.5	-0.5	0.587519
2005Q3	1.3	1.6	1.3	1.472545
2005Q4	1.6	1.8	1.3	1.665138
2006Q1	2.9	2.3	1.6	2.395027
2006Q2	2.8	2.5	2.9	2.406960
2006Q3	2.7	1.9	2.8	2.190500
2006Q4	2.7	2.3	2.7	2.304697
2007Q1	2.5	2.5	2.7	2.271466
2007Q2	2.6	3.1	2.5	2.487926
2007Q3	4.2	2.8	2.6	3.124916
2007Q4	4.5	3.5	4.2	3.460255
2008Q1	3.3	3.5	4.5	2.918277
2008Q2	3.0	2.7	3.3	2.554389
2008Q3	1.8	1.8	3.0	1.755467
2008Q4	-0.6	1.4	1.8	0.557314
2009Q1	-4.5	-1.3	-0.6	-1.974944
2009Q2	-5.4	-2.0	-4.5	-2.581272
2009Q3	-3.7	-2.6	-5.4	-1.984764
2009Q4	-2.2	-1.3	-3.7	-0.936152
2010Q1	0.6	2.0	-2.2	1.270588
2010Q2	2.2	0.6	0.6	1.593536
2010Q3	1.9	1.6	2.2	1.743534
2010Q4	2.5	2.0	1.9	2.128720
2011Q1	2.8	2.0	2.5	2.264214
2011Q2	1.6	3.1	2.8	2.036277
2011Q3	0.9	2.2	1.6	1.463181
2011Q4	-0.6	0.9	0.9	0.414568
2012Q1	-1.1	-0.5	-0.6	-0.210945
2012Q2	-0.4	0.0	-1.1	0.247955
2012Q3	-1.5	-1.0	-0.4	-0.534350
2012Q4	-1.7	-0.7	-1.5	-0.539032
2013Q1	-1.8	-0.5	-1.7	-0.527099
2013Q2	-1.7	0.4	-1.8	-0.224991
2013Q3	-0.4	1.1	-1.7	0.561996
2013Q4	0.8	1.1	-0.4	1.103975
2014Q1	0.0	1.0	0.8	0.714107
2014Q2	1.1	0.6	0.0	1.096723
2014Q3	1.2	1.8	1.1	1.484478
2014Q4	1.4	0.1	1.2	1.089472
2015Q1	2.5	1.2	1.4	1.900326
2015Q2	1.8	1.2	2.5	1.584172

Table 4: The IMF forecasts and realizations

Year	Actuals	IMF	No change forecast	Best model forecast
1991	-0.1	1.678600	NA	1.660111
1992	3.6	3.001200	-0.1	2.770674
1993	2.7	3.141500	3.6	2.537561
1994	4.0	2.555300	2.7	2.843543
1995	2.7	2.459800	4.0	2.482932
1996	3.8	2.022900	2.7	2.746576
1997	4.5	2.349000	3.8	2.962763
1998	4.4	2.581600	4.5	2.954253
1999	4.7	2.030900	4.4	2.991573
2000	4.1	2.598900	4.7	2.874187
2001	1.0	3.155700	4.1	2.077139
2002	1.8	2.189200	1.0	2.216891
2003	2.8	2.559400	1.8	2.518064
2004	3.8	3.914200	2.8	2.898140
2005	3.3	3.540100	3.8	2.732408
2006	2.7	3.268400	3.3	2.547731
2007	1.8	2.922400	2.7	2.275648
2008	-0.3	1.939000	1.8	1.626678
2009	-2.8	0.054956	-0.3	0.796930
2010	2.5	1.518100	-2.8	2.353165
2011	1.6	2.312600	2.5	2.172479
2012	2.3	1.782300	1.6	2.320036
2013	2.2	2.116400	2.3	2.319659

Table 5: Regression of actuals on past actuals and differences between judgmental forecasts and past actuals, that is, $y_t = \mu + \gamma_1 y_{t-1} + \gamma_2 (f_t - y_{t-1}) + \varepsilon_t$ and the Wald test on $\mu = 0, \gamma_1 = 1, \gamma_2 = 1$.

	EICIE	IMF
Parameters		
μ	-0.329 (0.230)	-0.236 (1.040)
γ_1	1.135 (0.120)	1.152 (0.404)
γ_2	0.691 (0.215)	0.943 (0.469)
R^2	0.770	0.338
P value of Wald test	0.001	0.695

Table 6: TLS parameter estimates for various values of δ

	EICIE		IMF	
δ	$\hat{\alpha}_{TLS}$	$\hat{\beta}_{TLS}$	$\hat{\alpha}_{TLS}$	$\hat{\beta}_{TLS}$
0.7	-0.993	1.622	-6.053	3.526
0.8	-0.977	1.608	-5.940	3.479
0.9	-0.963	1.594	-5.828	3.433
1.0	-0.949	1.582	-5.719	3.388
1.1	-0.936	1.570	-5.610	3.343
1.2	-0.924	1.559	-5.506	3.300
1.3	-0.912	1.549	-5.405	3.258

Table 7: Forecast performance. Fraction that forecasts have the lowest absolute forecast error across three forecasts, that is, the judgmentally adjusted forecast, the no change forecast and the best model forecast

	Judgment	No change	Best model
EICIE	21%	35%	44%
IMF	27%	23%	50%

Table 8: Quarters and years in which the final forecasts improve on both the no-change forecast and the best model forecast

EICIE	Quarters:	2007Q1, 2008Q1, 2008Q2, 2008Q3, 2009Q3, 2009Q4 2012Q2, 2012Q3, 2013Q4
IMF	Years:	1992, 2002, 2003, 2004, 2005, 2013

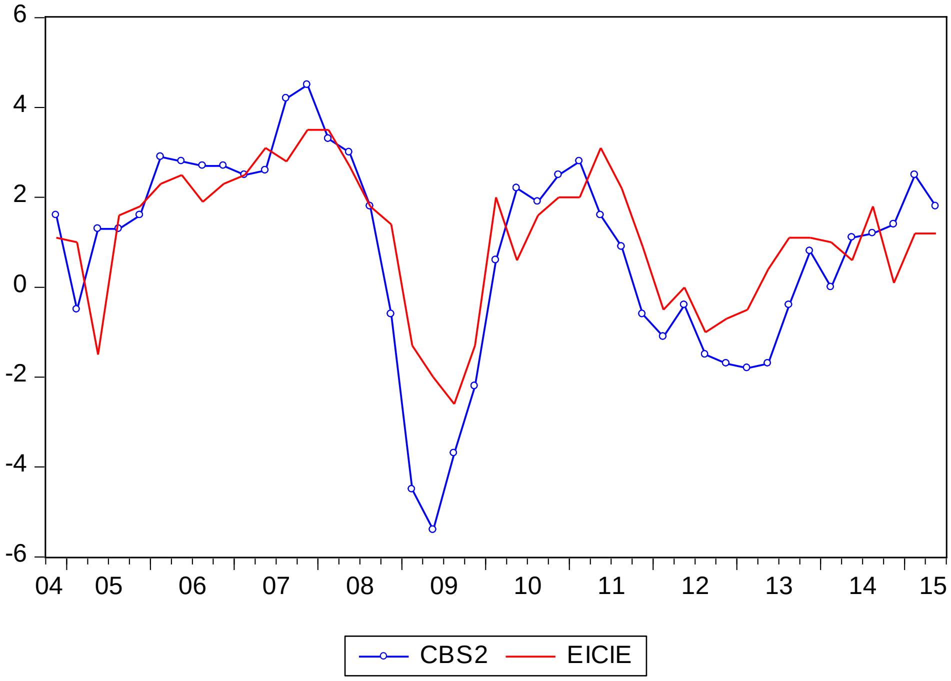

Figure 1: EICIE forecasts and actual quarterly GDP growth in the Netherlands (CBS2 concerns the second release data from Statistics Netherlands)

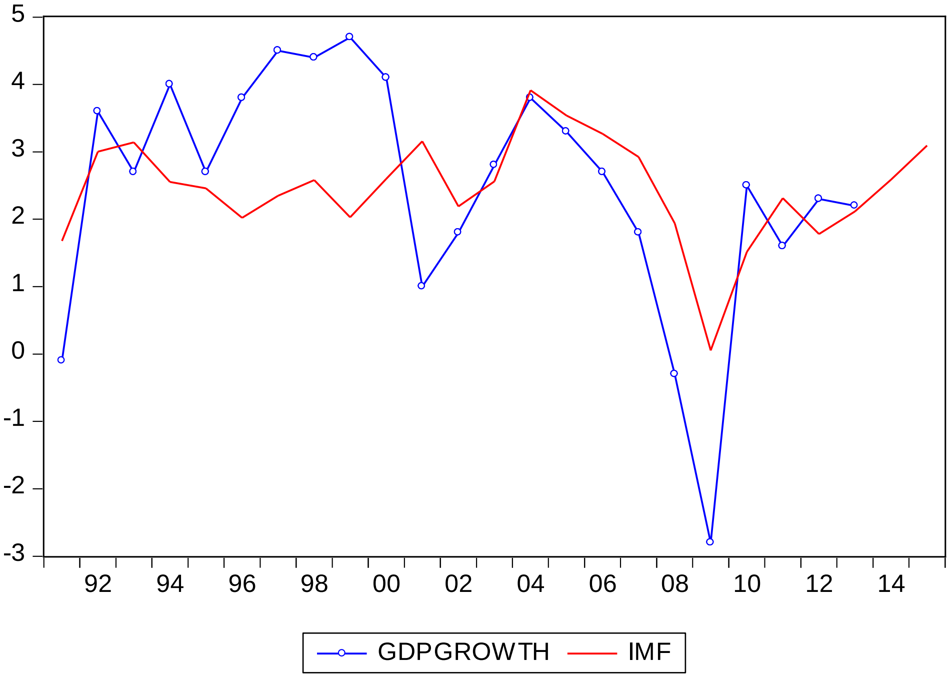

Figure 2: IMF forecasts and actual annual GDP growth rates (in the USA). Source is www.imf.org

References

Carroll, R.J. and D. Ruppert (1996), The use and misuse of orthogonal regression in linear errors-in-variables models, *The American Statistician*, 50, 1-6.

Deming, W.E. (1943), *Statistical Adjustment of Data*, London: Wiley.

Franses, P.H. (2014), *Expert Adjustments of Model Forecasts; Theory, Practice and Strategies for Improvement*, Cambridge: Cambridge University Press.

Franses, P.H., H. Kranendonk and D. Lanser (2011), One model and various experts: Evaluating Dutch macroeconomic forecasts, *International Journal of Forecasting*, 27, 482-495.

Franses, P.H. and N. Maassen (2015), Consensus forecasters: How good are they individually and why?, Econometric Institute Report 2015-21, Erasmus University Rotterdam.

Fuller, W.A. (1987), *Measurement Error Models*, New York: Wiley.

Koopmans, T. (1937), *Linear Regression Analysis of Economic Time Series*, Haarlem the Netherlands: De Erven F. Bohn NV.

Linnet, K. (1990), Estimation of the linear relationship between the measurements of two methods with proportional errors, *Statistics in Medicine*, 9, 1463-1473.

Vuchelen, J. and M.-I. Gutierrez (2005), A direct test of the information content of the OECD growth forecasts, *International Journal of Forecasting*, 21, 103-117.

Wansbeek, T. and E. Meijer (2000), *Measurement Error and Latent Variables in Econometrics*, Amsterdam: North-Holland.