Received: 9 June 2020 | Revised: 4 November 2020 | Accepted: 4 January 2021

DOI: 10.1002/for.2765

RESEARCH ARTICLE

WILEY

Testing bias in professional forecasts

Philip Hans Franses

Econometric Institute, Erasmus School of Economics, Rotterdam, The Netherlands

Correspondence
Philip Hans Franses, Econometric Institute, Erasmus School of Economics, PO Box 1738, Rotterdam NL-3000 DR, The Netherlands.
Email: franses@ese.eur.nl

Abstract

Professional forecasters can rely on econometric models, on their personal expertise or on both. To accommodate for adjustments to model forecasts, this paper proposes to use two stage least squares (TSLS) (and not ordinary least squares [OLS]) for the familiar Mincer–Zarnowitz regression when examining bias in professional forecasts, where the instrumental variable is the consensus forecast. An illustration for 15 professional forecasters with their quotes for real gross domestic product (GDP) growth, inflation and unemployment for the United States documents the usefulness of this new estimation method. It also shows that TSLS suggests less bias than OLS does.

KEYWORDS

bias, forecast evaluation, Mincer–Zarnowitz regression, OLS, TSLS

1 | INTRODUCTION AND MOTIVATION

One way to examine bias in forecasts is to consider the so-called Mincer and Zarnowitz (1969) (MZ) regression. Given forecasts m_t from an econometric model for a variable y_t, this MZ regression reads as

$$y_t = \alpha + \beta m_t + \varepsilon_t,$$

and the focus is on the hypothesis that $\alpha = 0$ and $\beta = 1$, which entails unbiasedness. In this paper this regression is used to examine bias in forecasts made by professional forecasters.

Usually the parameters in the MZ regression are estimated using ordinary least squares (OLS), although alternative estimators are proposed in, for example, Lovell (1986) and Jeong and Maddala (1991). These alternative estimators rely on instrumental variables as it is hypothesized that a forecast can have measurement errors. Jeong and Maddala (1991) propose to use a second forecast for the same target variable as an instrument.

In a sense, this paper extends this notion of measurement errors by advancing the idea that professional forecasters do not solely rely on an econometric model but also add their own adjustment. The "measurement error" is then associated with the adjustment of an econometric model forecast, which makes the measurement error interpretable. As there are many forecasters around, it seems most sensible to choose a consensus forecast as the instrumental variable in all MZ regressions. This also avoids multiple testing problems.

In sum, in this paper the, OLS estimation method is challenged because it is uncertain if each professional

Thanks are due to an anonymous reviewer, Christiaan Heij and Michel van der Wel for their helpful suggestions. All computations are performed using EViews (version 11).

forecaster relies only on an econometric model, which in principle aims at unbiased forecasts. In fact, it is not unlikely that professional forecasters consult the outcome of an econometric model and modify this outcome using their experience or intuition. In this paper, it is argued that this modification of the model forecast can be interpreted as introducing a measurement error in the m_t variable in the MZ regression. This measurement error makes that the finally observed forecast is endogenous. It is therefore proposed that a proper estimation method for the MZ regression is two stage least squares (TSLS). The instrumental variable, next to an intercept, that is necessary for TSLS is the average forecast of all involved forecasters, a consensus forecast. This consensus forecast is correlated with each of the individual forecasts, and it shall be uncorrelated with the individual adjustment. Comparing the OLS and TSLS estimates also allows for evaluating how much adjustment is exercised by each of the professional forecasters and whether such adjustment is similar across variables.

The new estimation method is illustrated for the quotes of 15 professional forecasters for the USA economy, who give their quotes via Consensus Economics.[1] The forecasts concern real gross domestic product (GDP) growth, (consumer price index based) inflation and the unemployment rate for year T, where the forecasts are created in the 12 months in the preceding year $T-1$.

Section 2 provides more details about the estimation method. Section 3 deals with an empirical illustration and shows that the choice of estimation method matters for inference. Section 4 concludes.

2 | METHODOLOGY

To examine forecast bias, one can use the MZ regression. Consider a professional forecaster who uses an econometric model to create forecasts for y_t, and denote such a model forecast as m_t. If a professional forecaster has adjusted this model forecast for some specific reasons, then the evaluation does not involve m_t but an observed forecast f_t. This final forecast is

$$f_t = m_t + a_t, \tag{1}$$

where a_t is the adjustment.[2] The MZ regression then is not equal to

$$y_t = \alpha + \beta m_t + \varepsilon_t, \tag{2}$$

but it is

$$y_t = \alpha + \beta f_t + \varepsilon_t.$$

The test regression thus becomes

$$y_t = \alpha + \beta m_t + \varepsilon_t + \beta a_t.$$

As the covariance between m_t and $\varepsilon_t + \beta a_t$ equals $-\beta\sigma_a^2 \neq 0$, where σ_a^2 is the variance of the adjustment, the variable m_t is endogenous. This is the familiar errors-in-variables setting, for which it holds that the OLS based estimator $\hat{\beta}$ has the property:

$$\hat{\beta}_{\text{OLS}} \overset{p}{\Rightarrow} \frac{\beta}{1 + \sigma_a^2/\sigma_m^2}. \tag{3}$$

Hence, the OLS-based estimated β is smaller than the true β.

A solution is now to use TSLS with an instrumental variable, additional to the intercept. This instrument should be correlated with each m_t and not with a_t. As we do not know the individual model-based forecasts m_t, a suitable choice in this setting of the analysis of individual forecasters is to take a consensus forecast as the instrument. This consensus forecast can be an unweighted average of all available forecasts. This variable will be correlated with the forecasts of each individual forecaster, as they are included in the consensus forecasts, but in theory, it will not be correlated with the adjustments made by each individual forecaster.

Next, a test for exogeneity can be carried out to examine if TSLS is indeed a more appropriate estimation method. When the Durbin Wu Hausman test indicates rejection of the null hypothesis of exogeneity, the difference between the two estimates is informative about the size of adjustment. When the TSLS estimator $\hat{\beta}_{\text{TSLS}}$ is interpreted as the "true" β and when the OLS estimator is denoted by $\hat{\beta}_{\text{OLS}}$, one can use 3 to infer the ratio

$$\frac{\sigma_a^2}{\sigma_m^2} = \frac{\hat{\beta}_{\text{TSLS}} - \hat{\beta}_{\text{OLS}}}{\hat{\beta}_{\text{OLS}}}. \tag{4}$$

One may expect that the larger is the variance in f_t, the larger is this ratio. This is because econometric model-based forecasts by their very nature do not fluctuate much. Another interesting feature to examine may be that in times of more uncertainty, for example in times of an economic crisis, forecasters may rely more on their judgment, as econometric models are not very good in predicting new regimes. On the other hand, it may also be that precisely in times of more uncertainty, one may want to stick closer to an econometric model.

3 | ILLUSTRATION

Table 1 lists the names and numbers of forecasters, all included in Consensus Economics, where the data cover 2000M01 to and including 2013M12, where the annual realizations from 2001 to 2014 are considered. When observations are missing, because forecasters did not deliver their quotes in certain months, these observations are not included in the regressions; that is, no interpolation is carried out. The focus will be on forecasts for real growth of GDP, CPI-based inflation, and the unemployment rate, all for the United States.

The MZ regression

$$y_t = \alpha + \beta f_t + \varepsilon_t$$

can involve serially correlated errors. In fact, inspection of the residuals indicates that the first order autocorrelation is close to 1. Hence, next to the MZ regression in levels, also the MZ regression in first differences is considered, that is,

$$y_t - y_{t-1} = \alpha + \beta(f_t - f_{t-1}) + \varepsilon_t.$$

Table 2a presents the results for the full sample for real GDP growth in the United States for the MZ regression in levels. The last column reports on a test for exogeneity of the regressor f_t. In only two cases is the null hypothesis of exogeneity not rejected, where a 5% cut-off point is adopted. In all cases where exogeneity is rejected, the $\hat{\beta}_{\text{TSLS}} > \hat{\beta}_{\text{OLS}}$, which confirms the notion of measurement errors. In words, if the professional forecasters each use an econometric model, they almost all adjust their model forecasts.

The median value across the 15 cases of

$$\frac{\hat{\beta}_{\text{TSLS}} - \hat{\beta}_{\text{OLS}}}{\hat{\beta}_{\text{OLS}}}$$

is 0.289, which means that the variance of the adjustments is 28.9% of the variance of the model forecasts.

Looking again at Table 2a, when using OLS, the number of times $\alpha = 0$ is in the 95% confidence interval is seven, and the number of times $\beta = 1$ is in the 95% confidence interval is 11. When using TSLS, the number of times $\alpha = 0$ is in the 95% confidence interval is two, and the number of times $\beta = 1$ is in the 95% confidence interval is 12. For GDP growth it can thus be seen that β is close to 1, but that the forecasts usually are too optimistic.

Table 2b presents the results for the full sample for real GDP growth in the United States for the MZ regression in differences. Exogeneity is rejected in 10 of the 15 cases. When using OLS, the number of times $\alpha = 0$ is in the 95% confidence interval is 15, and the number of times $\beta = 1$ is in the 95% confidence interval is 0. When using TSLS, the number of times $\alpha = 0$ is in the 95% confidence interval is again 15, and the number of times $\beta = 1$ is in the 95% confidence interval is 4. Hence, we can see that using TSLS results in less evidence of forecast bias. For the differences, it is obtained that the

TABLE 1 List of forecasters and sample size (not all forecasters give quotes in all months in the period 2000M01–2013M12)

	GDP growth	Inflation	Unemployment rate
DuPont	168	168	168
JP Morgan	162	162	162
Eaton Corporation	157	156	157
National Association of Home Builders	153	153	153
The Conference Board	153	153	153
Fannie Mae	151	151	151
General Motors	151	151	151
Wells Capital Management	149	149	149
Goldman Sachs	148	148	148
University of Michigan – RSQE	148	148	148
Ford Motor Corporation	146	143	146
Oxford Economics	146	146	146
Macroeconomic Advisors	144	143	143
Morgan Stanley	142	142	142
Georgia State University	135	135	135

Abbreviation: GDP, gross domestic product.

TABLE 2a Results for real GDP growth, 2000M01–2013M12

	OLS		TSLS		*p* value
	α	β	α	β	H_0 : Exogeneity
DuPont	−0.692 (0.396)	0.957 (0.141)	−1.457 (0.449)	1.242 (0.162)	0.000
JP Morgan	−0.654 (0.414)	0.897 (0.142)	−1.579 (0.451)	1.229 (0.156)	0.000
Eaton	−0.234 (0.478)	0.694 (0.151)	−1.830 (0.550)	1.218 (0.176)	0.000
National Association of Home Builders	−1.284 (0.454)	1.069 (0.151)	−1.586 (0.473)	1.173 (0.158)	0.025
The Conference Board	−0.213 (0.330)	0.740 (0.110)	−0.996 (0.381)	1.105 (0.129)	0.000
Fannie Mae	−0.695 (0.374)	0.901 (0.126)	−0.981 (0.387)	1.003 (0.131)	0.004
General Motors	−0.976 (0.496)	1.013 (0.168)	−1.818 (0.526)	1.306 (0.179)	0.000
Wells Capital Management	−1.516 (0.421)	1.126 (0.143)	−2.180 (0.454)	1.361 (0.155)	0.000
Goldman Sachs	−0.310 (0.371)	0.899 (0.139)	−0.795 (0.413)	1.091 (0.157)	0.008
University of Michigan – RSQE	−0.809 (0.448)	0.923 (0.152)	−1.912 (0.486)	1.315 (0.166)	0.000
Ford Motor Corporation	−1.652 (0.458)	1.235 (0.155)	−2.405 (0.511)	1.497 (0.174)	0.000
Oxford Economics	1.154 (0.503)	0.303 (0.171)	−0.968 (0.583)	1.051 (0.200)	0.000
Macroeconomic Advisors	−2.434 (0.429)	1.296 (0.131)	−2.695 (0.461)	1.379 (0.141)	0.124
Morgan Stanley	−1.136 (0.309)	1.021 (0.099)	−1.035 (0.323)	0.987 (0.104)	0.274
Georgia State University	0.024 (0.307)	0.779 (0.120)	−0.802 (0.337)	1.140 (0.135)	0.000

Note. All variables are in levels. Standard errors are in parentheses.
Abbreviations: GDP, gross domestic product; OLS, ordinary least squares; TSLS, two stage least squares.

TABLE 2b Results for real GDP growth, 2000M01–2013M12

	OLS		TSLS		*p* value
	α	β	α	β	H_0 : Exogeneity
DuPont	−0.014 (0.040)	0.431 (0.071)	−0.013 (0.041)	0.625 (0.085)	0.000
JP Morgan	−0.011 (0.040)	0.619 (0.078)	−0.010 (0.040)	0.678 (0.087)	0.138
Eaton	−0.019 (0.043)	0.708 (0.103)	−0.017 (0.044)	0.948 (0.125)	0.000
National Association of Home Builders	−0.030 (0.043)	0.645 (0.085)	−0.031 (0.043)	0.711 (0.092)	0.060
The Conference Board	−0.014 (0.043)	0.641 (0.084)	−0.009 (0.043)	0.812 (0.107)	0.005
Fannie Mae	−0.022 (0.040)	0.672 (0.078)	−0.024 (0.040)	0.746 (0.083)	0.008
General Motors	−0.025 (0.045)	0.685 (0.090)	−0.026 (0.045)	0.718 (0.097)	0.371
Wells Capital Management	0.003 (0.041)	0.584 (0.091)	0.003 (0.042)	0.763 (0.103)	0.000
Goldman Sachs	−0.057 (0.035)	0.296 (0.081)	−0.056 (0.035)	0.127 (0.099)	0.002
University of Michigan – RSQE	−0.029 (0.048)	0.578 (0.090)	−0.033 (0.049)	0.784 (0.112)	0.001
Ford Motor Corporation	−0.015 (0.025)	0.051 (0.072)	−0.010 (0.026)	0.193 (0.099)	0.032
Oxford Economics	−0.039 (0.032)	0.192 (0.071)	−0.039 (0.032)	0.188 (0.092)	0.948
Macroeconomic Advisors	0.011 (0.037)	0.642 (0.065)	0.011 (0.037)	0.646 (0.070)	0.861
Morgan Stanley	−0.025 (0.051)	0.498 (0.086)	−0.024 (0.052)	0.647 (0.094)	0.000
Georgia State University	−0.010 (0.048)	0.381 (0.090)	−0.020 (0.052)	0.783 (0.120)	0.000

Note. All variables are in first differences. Standard errors are in parentheses.
Abbreviations: GDP, gross domestic product; OLS, ordinary least squares; TSLS, two stage least squares.

variance of the adjustment is 26.7% of the variance of the model forecasts, again using the median across the 15 cases.

Table 3a provides the results for the levels data for the test on exogeneity when the sample is split in a period before the Great Recession, and in a recession

TABLE 3a GDP growth results, when the data are in level

	2001M01–2007M12	2008M01–2013M012
DuPont	0.530	0.342
JP Morgan	0.009	0.010
Eaton	0.052	0.000
National Association of Home Builders	0.941	0.018
The Conference Board	0.534	0.000
Fannie Mae	0.083	0.000
General Motors	0.009	0.433
Wells Capital Management	0.715	0.086
Goldman Sachs	0.000	0.157
University of Michigan – RSQE	0.216	0.000
Ford Motor Corporation	0.982	0.618
Oxford Economics	0.002	0.000
Macroeconomic Advisors	0.000	0.840
Morgan Stanley	0.500	0.464
Georgia State University	0.008	0.000

Note. p values of the test for exogeneity in two samples, one before the Great Recession, one with and slightly after that recession.
Abbreviation: GDP, gross domestic product.

period with later years. For four forecasters, exogeneity is not rejected in the two samples, and for three forecasters, exogeneity is rejected in both samples. For four forecasters, it holds that the variable moves from exogenous to endogenous, whereas for three forecasters, it is the other way around. Hence, there is no clear tendency here into more (or less) adjustment, depending on a crisis. Table 3b provides similar results for the tests on exogeneity where now the MZ regression is considered for the differenced data.

TABLE 3b GDP growth results

	2001M01–2007M12	2008M01–2013M012
DuPont	0.003	0.001
JP Morgan	0.727	0.166
Eaton	0.128	0.001
National Association of Home Builders	0.001	0.738
The Conference Board	0.000	0.160
Fannie Mae	0.636	0.008
General Motors	0.107	0.926
Wells Capital Management	0.000	0.005
Goldman Sachs	0.885	0.000
University of Michigan – RSQE	0.103	0.028
Ford Motor Corporation	0.196	0.043
Oxford Economics	0.923	0.002
Macroeconomic Advisors	0.000	0.539
Morgan Stanley	0.035	0.001
Georgia State University	0.000	0.000

Note. p values of the test for exogeneity in two samples, one before the Great Recession, one with and slightly after that recession. All variables are in first differences.
Abbreviation: GDP, gross domestic product.

Table 3c summarizes the results in two tables. It can be seen that the forecasts seem to shift most from endogenous (before the recession) to exogenous afterwards. One may tentatively conclude that more trust is given to econometric model forecasts after the recession.

Table 4a provides the results for the inflation rate for the data in levels. For eight of the 15 forecasters, exogeneity cannot be rejected, which in words means that they basically make use of an econometric model and do not adjust. When using OLS, the number of times $\alpha = 0$ is in the 95% confidence interval is 3, and the number of times $\beta = 1$ is in the 95% confidence interval is only 2. In contrast, when using TSLS, the number of times $\alpha = 0$ is in the 95% confidence interval is 6, and the number of times $\beta = 1$ is in the 95% confidence interval is 8. So, the application of TSLS gives more evidence in favor of unbiased forecasts. The median value of the 15 values as in 4 is 0.126.

Table 4b provides the results for the inflation rate for the data in differences. For 13 of the 15 forecasters, exogeneity is not rejected, which in words now suggests that many forecasters adjust an econometric model forecast. When using OLS, the number of times $\alpha = 0$ is in the 95% confidence interval is 15, and the number of times $\beta = 1$ is in the 95% confidence interval is zero. When using TSLS, the number of times $\alpha = 0$ is in the 95% confidence interval is 15, and the number of times $\beta = 1$ is in the 95% confidence interval is 5. So, again application of TSLS gives more evidence in favor of unbiased forecasts. The median value of the 15 values as in 4 is 0.933, which is much larger than for the levels case.

TABLE 3c Summary of results on before and after recession tests for exogeneity, based on a 5% cut-off point for the *p* values

Levels		**With/after recession**	
		Exogenous	Endogenous
Before recession	Exogenous	3	3
	Endogenous	5	4
Differences			
		With/after recession	
		Exogenous	Endogenous
Before recession	Exogenous	4	3
	Endogenous	6	2

TABLE 4a Results for the inflation rate, 2000M01–2013M12

	OLS		**TSLS**		***p* value**
	α	β	α	β	H_0 **: Exogeneity**
DuPont	2.141 (0.430)	0.132 (0.189)	1.046 (0.592)	0.622 (0.263)	0.007
JP Morgan	1.290 (0.286)	0.583 (0.143)	1.654 (0.333)	0.393 (0.168)	0.032
Eaton	−0.856 (0.586)	1.443 (0.253)	−0.290 (0.949)	1.197 (0.412)	0.449
National Association of Home Builders	1.758 (0.333)	0.325 (0.154)	1.588 (0.372)	0.407 (0.174)	0.305
The Conference Board	2.007 (0.307)	0.163 (0.111)	1.540 (0.375)	0.338 (0.137)	0.030
Fannie Mae	2.545 (0.358)	−0.064 (0.160)	1.418 (0.410)	0.457 (0.185)	0.000
General Motors	1.064 (0.329)	0.610 (0.141)	1.029 (0.424)	0.625 (0.184)	0.898
Wells Capital Management	0.887 (0.563)	0.555 (0.210)	0.700 (0.922)	0.625 (0.346)	0.798
Goldman S	2.306 (0.246)	0.115 (0.117)	2.081 (0.284)	0.228(0.137)	0.112
University of Michigan – RSQE	3.319 (0.447)	−0.362 (0.184)	1.325 (0.616)	0.477 (0.256)	0.000
Ford Motor Corporation	1.782 (0.474)	0.300 (0.223)	0.569 (0.924)	0.882 (0.441)	0.126
Oxford Economics	3.416 (0.422)	−0.407 (0.185)	2.000 (0.583)	0.224 (0.258)	0.000
Macroeconomic Advisors	2.237 (0.291)	0.066 (0.139)	1.908 (0.319)	0.232 (0.154)	0.012
Morgan Stanley	1.551 (0.394)	0.378 (0.169)	1.098 (0.554)	0.578 (0.241)	0.245
Georgia State University	0.738 (0.369)	0.846 (0.181)	1.252 (0.601)	0.586 (0.301)	0.279

Note. All variables are in levels. Standard errors are in parentheses.
Abbreviations: OLS, ordinary least squares; TSLS, two stage least squares.

Table 5a provides the results for the MZ regression in levels for the USA unemployment rate. Exogeneity is rejected in 10 of the 15 cases, so this means that, also for this variable, model-based forecasts are often adjusted. When using OLS, the number of times $\alpha = 0$ is in the 95% confidence interval is 1, and the number of times $\beta = 1$ is in the 95% confidence interval is also one. In contrast, when using TSLS, the number of times $\alpha = 0$ is in

TABLE 4b Results for the inflation rate, 2000M01–2013M12

	OLS		TSLS		p value
	α	β	α	β	H_0 : Exogeneity
DuPont	−0.010 (0.031)	0.318 (0.075)	−0.010 (0.031)	0.490 (0.092)	0.002
JP Morgan	−0.010 (0.034)	0.476 (0.118)	−0.009 (0.034)	0.815 (0.159)	0.002
Eaton	−0.014 (0.037)	0.373 (0.201)	0.022 (0.038)	4.208 (0.838)	0.000
National Association of Home Builders	−0.004 (0.036)	0.339 (0.115)	−0.002 (0.036)	0.784 (0.164)	0.000
The Conference Board	0.000 (0.034)	0.237 (0.061)	0.002 (0.034)	0.414 (0.080)	0.001
Fannie Mae	0.012 (0.025)	0.219 (0.083)	0.008 (0.025)	0.643 (0.123)	0.000
General Motors	−0.005 (0.040)	0.065 (0.128)	0.025 (0.040)	1.700 (0.334)	0.000
Wells Capital Management	−0.005 (0.037)	0.316 (0.094)	−0.001 (0.037)	0.698 (0.143)	0.000
Goldman S	−0.031 (0.034)	0.431 (0.105)	−0.043 (0.035)	0.833 (0.215)	0.032
University of Michigan – RSQE	−0.014 (0.039)	0.345 (0.110)	−0.014 (0.039)	0.774 (0.161)	0.000
Ford Motor Corporation	−0.022 (0.014)	0.040 (0.040)	−0.024 (0.014)	−0.062 (0.200)	0.600
Oxford Economics	0.007 (0.018)	0.261 (0.063)	0.007 (0.018)	0.288 (0.093)	0.696
Macroeconomic Advisors	−0.028 (0.036)	0.485 (0.107)	−0.028 (0.036)	0.715 (0.137)	0.008
Morgan Stanley	−0.006 (0.039)	0.360 (0.090)	−0.006 (0.039)	0.557 (0.117)	0.008
Georgia State University	−0.000 (0.040)	0.274 (0.116)	0.002 (0.040)	1.236 (0.270)	0.000

Note. Standard errors are in parentheses. All variables are in first differences.
Abbreviations: OLS, ordinary least squares; TSLS, two stage least squares.

TABLE 5a Results for the unemployment rate, 2000M01–2013M12

	OLS		TSLS		p value
	α	β	α	β	H_0 : Exogeneity
DuPont	1.085 (0.304)	0.856 (0.047)	0.815 (0.306)	0.899 (0.048)	0.000
JP Morgan	1.492 (0.277)	0.803 (0.043)	1.353 (0.279)	0.825 (0.043)	0.000
Eaton	1.490 (0.281)	0.807 (0.044)	1.295 (0.285)	0.839 (0.045)	0.000
National Association of Home Builders	0.991 (0.312)	0.875 (0.048)	0.850 (0.314)	0.898 (0.049)	0.000
The Conference Board	1.727 (0.278)	0.754 (0.042)	1.571 (0.280)	0.779 (0.042)	0.000
Fannie Mae	1.090 (0.279)	0.852 (0.043)	1.036 (0.280)	0.860 (0.043)	0.065
General Motors	0.819 (0.305)	0.883 (0.047)	0.632 (0.307)	0.913 (0.048)	0.000
Wells Capital Management	1.490 (0.318)	0.817 (0.049)	1.254 (0.321)	0.855 (0.049)	0.000
Goldman S	1.255 (0.229)	0.791 (0.035)	1.360 (0.232)	0.774 (0.035)	0.006
University of Michigan – RSQE	1.238 (0.296)	0.834 (0.048)	1.251 (0.298)	0.832 (0.048)	0.691
Ford Motor Corporation	0.508 (0.300)	0.916 (0.045)	0.483 (0.304)	0.920 (0.046)	0.623
Oxford Economics	0.919 (0.321)	0.883 (0.051)	0.610 (0.324)	0.934 (0.052)	0.000
Macroeconomic Advisors	1.393 (0.319)	0.833 (0.049)	1.272 (0.321)	0.852 (0.049)	0.000
Morgan Stanley	1.016 (0.299)	0.865 (0.045)	1.019 (0.302)	0.865 (0.045)	0.934
Georgia State University	1.392 (0.324)	0.784 (0.046)	1.363 (0.326)	0.788 (0.046)	0.442

Note. All variables are in levels. Standard errors are in parentheses.
Abbreviations: OLS, ordinary least squares; TSLS, two stage least squares.

TABLE 5b Results for the unemployment rate, 2000M01–2013M12

	OLS		TSLS		*p* value
	α	β	α	β	H_0 : Exogeneity
DuPont	0.017(0.025)	0.218 (0.096)	0.013 (0.025)	0.424 (0.129)	0.002
JP Morgan	0.018 (0.027)	0.237 (0.104)	0.014 (0.270)	0.442 (0.139)	0.026
Eaton	0.022 (0.029)	−0.056 (0.093)	0.015 (0.029)	0.724 (0.208)	0.000
National Association of Home Builders	0.013 (0.028)	0.412 (0.130)	0.013 (0.028)	0.509 (0.156)	0.253
The Conference Board	0.017 (0.028)	0.199 (0.109)	0.014 (0.028)	0.435 (0.165)	0.058
Fannie Mae	−0.000 (0.015)	0.358 (0.062)	−0.001 (0.015)	0.449 (0.080)	0.076
General Motors	0.029 (0.030)	0.134 (0.120)	0.026 (0.031)	0.435 (0.181)	0.026
Wells Capital Management	0.018 (0.029)	0.064 (0.100)	0.011 (0.030)	0.421 (0.139)	0.000
Goldman S	0.025 (0.026)	0.752 (0.109)	0.025 (0.026)	0.744 (0.192)	0.963
University of Michigan – RSQE	0.027 (0.030)	0.389 (0.120)	0.027 (0.030)	0.541 (0.161)	0.157
Ford Motor Corporation	0.002 (0.013)	0.088 (0.045)	0.003 (0.013)	0.352 (0.076)	0.000
Oxford Economics	0.013 (0.015)	0.174 (0.062)	0.011 (0.015)	0.246 (0.092)	0.293
Macroeconomic Advisors	0.019 (0.031)	0.245 (0.129)	0.014 (0.031)	0.522 (0.177)	0.022
Morgan Stanley	0.013 (0.031)	0.420 (0.122)	0.012 (0.031)	0.516 (0.162)	0.367
Georgia State University	0.010 (0.031)	0.228 (0.108)	0.007 (0.031)	0.475 (0.155)	0.027

Note. Standard errors are in parentheses. The variables are in first differences.
Abbreviations: OLS, ordinary least squares; TSLS, two stage least squares.

the 95% confidence interval is two, and the number of times $\beta = 1$ is in the 95% confidence interval is three. So, also here the application of TSLS gives slightly more evidence in favor of unbiased forecasts. The differences between the $\hat{\beta}_{\text{TSLS}}$ and the $\hat{\beta}_{\text{OLS}}$ are much smaller than for real GDP growth. The median value of the 15 values as in 4 is only 0.026.

Finally, Table 5b provides the results for unemployment in differences. When using OLS, the number of times $\alpha = 0$ is in the 95% confidence interval is 15, and the number of times $\beta = 1$ is in the 95% confidence interval is 0. When using TSLS, the number of times $\alpha = 0$ is in the 95% confidence interval is 15, and the number of times $\beta = 1$ is in the 95% confidence interval is two. Application of TSLS gives a little more evidence in favor of unbiased forecasts. The median value of the 15 values as in 4 is 0.865, which is again much larger than for the levels case. Exogeneity is rejected in eight cases.

Table 6 concisely summarizes the main results on tests for exogeneity. Clearly, exogeneity is rejected for many forecasts, and hence, there is ample evidence of adjustment to model-based forecasts.

For the levels data, exogeneity is rejected for all three variables for Dupont, JP Morgan, TCB, and Oxford; for two variables for Eaton, NAHB, Fannie Mae, GM, Wells CM, Goldman Sachs, RSQE, and Macroeconomic Advisors; and for one variable for Ford, Georgia State University, whereas Morgan Stanley seems to uniquely rely on an econometric model. For the differences data, exogeneity is rejected for all three variables for Dupont, Eaton, Wells, and GSU, for two variables for JP Morgan, TCB, Fannie Mae, GM, Goldman Sachs, RSQE, Ford, Macroeconomic Advisors, and Morgan Stanley, for one variable NAHB, whereas now Oxford seems to rely on an econometric model only. Although the results are somewhat mixed, it rarely seems to happen that pure econometric model-based forecasts are used.

TABLE 6 A summary of the results, when the 5% significance level is taken

	GDP growth	Inflation	Unemployment
Number of times exogeneity is rejected			
Levels	13	7	10
Differences	10	13	8

Abbreviation: GDP, gross domestic product.

4 | CONCLUSION

It is assumed that professional forecasters can rely on the outcome of an econometric model and on their adjustment of the model forecast. If that is the case, this paper proposed to use TSLS for the familiar MZ regression when examining bias in professional forecasts. The instrument is the average forecast across all professional forecasts. As there can be autocorrelation, one may want to consider the MZ regression for levels and for differences data. An illustration for 15 forecasters with their quotes for three important macroeconomic variables for the United States showed the relevance of the TSLS estimation method.

The illustration learned that adjustment of model-based forecasts is very common. This follows from the rejection of the null hypothesis of exogeneity in many cases. On the size of the variance of adjustment relative to the variance in the model-based forecasts, results are mixed for inflation and unemployment, depending on whether one takes the levels or differences data, but for GDP growth, the results are consistent, suggesting that adjustment variance is about 27% of the variance of the model forecasts. Another conclusion is that the use of TSLS with the consensus forecast as an instrument provides more evidence of unbiasedness than OLS does. So, professional forecasters seem better that one would have thought. An advantage of this new method is that all MZ regressions have the same instrument, and this allows for comparison across forecasters. As such, a final insight is that solely relying on an econometric model to create forecasts seems very rare.

The analysis in this paper considers three variables for the United States with forecasts from 15 forecasters. Further experience with other countries, more forecasters and more variables should tell how relevant the new method is in other settings.

DATA AVAILABILITY STATEMENT

The data used in this paper can be obtained from the author.

ORCID

Philip Hans Franses https://orcid.org/0000-0002-2364-7777

ENDNOTES

[1] https://www.consensuseconomics.com/

[2] The literature summarized in Franses (2014) shows that many model-based forecasts are modified manually, for various reasons.

REFERENCES

Franses, P. H. (2014). *Expert adjustments of model forecasts*. Cambridge UK: Cambridge University Press.

Jeong, J., & Maddala, G. S. (1991). Measurement errors and tests for rationality. *Journal of Business & Economic Statistics*, *9*(4), 431–439.

Lovell, M. C. (1986). Tests of the rational expectations hypothesis. *American Economic Review*, *76*(1), 110–124.

Mincer, J., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In J. Mincer (Ed.), *Economic forecasts and expectations* (pp. 81–111). New York: National Bureau of Economic Research.

AUTHOR BIOGRAPHY

Philip Hans Franses (1963) is Professor of Applied Econometrics and Professor of Marketing Research, both at the Erasmus University Rotterdam.

How to cite this article: Franses PH. Testing bias in professional forecasts. *Journal of Forecasting*. 2021;1–9. https://doi.org/10.1002/for.2765