Towson University
Department of Economics
Working Paper Series

Working Paper No. 2020-05

The Nordhaus Test with Many Zeros

By Kajal Lahiri and Yongchen Zhao

June 2020

The Nordhaus Test with Many Zeros

Kajal Lahiri[1] and Yongchen Zhao[2]

May 2020

Abstract: We reformulate the Nordhaus test as a friction model where the large number of zero revisions are treated as censored, i.e., unknown values inside a small region of "imperceptibility." Using Blue Chip individual forecasts of U.S. real GDP growth, inflation, and unemployment over 1985-2020, we find pervasive over-reaction to news at most of the monthly forecast horizons from 24 to 1, but the degree of inefficiency is very small. The updaters, i.e., those who make non-zero revisions, are not found to perform better than their "inattentive" peers do.

Keywords: Nordhaus test; Expectations updating; Forecast efficiency; Fixed-event forecasts; Inattentive forecasters.

JEL Codes: C53, E27, E37

Highlights:

- Zero revisions in fixed-event forecasts treated as censored observations
- The Nordhaus test of forecast efficiency extended to account for heaping at zero
- Blue Chip individual forecasts exhibit over-reaction to news
- Zero revision is not a good indicator of forecaster inattention

[1] Department of Economics, University at Albany, SUNY, klahiri@albany.edu.
[2] Department of Economics, Towson University, yzhao@towson.edu.

The Nordhaus Test with Many Zeros

1. Introduction

Revisions in fixed-target forecasts are often used to study information frictions, and how economic agents update their expectations. In particular, researchers have applied the Nordhaus (1987) test of forecast efficiency to check whether forecasters use all relevant information available to them at the time of making a prediction.

One common feature of fixed-event forecasts is that they often remain unchanged from one period to the next. For example, 37% of the revisions to the Blue Chip forecasts of U.S. real GDP growth rate are zero. At short horizons, the proportion of zero revisions may reach as high as 80% (see Figure 1). In recent years, many authors have used zero revisions as a non-parametric measure of forecaster inattention to news, generating measures of information rigidities at the macro level.[3] However, forecasters may choose not to revise their forecasts for many reasons. Most importantly, since many of the established professional survey institutions collect forecasts rounded to the first decimal point, small revisions may become zero after rounding. At long horizons, forecasters may be uncertain about how much to revise because of possibly offsetting news in the future, and thus decide to stick to their priors. At short horizons, zero revisions may arise when the news is too small to affect the outturn much during the few remaining months of the target year. The frequency of revisions may also be limited by the significant material and cognitive costs associated with acquiring and processing information. The forecasters may be inefficient in that non-zero revisions are made only when enough information is accrued over time to justify "big" revisions.

[3] See, for instance, Drager and Lamla (2012), Pfajfar and Santoro (2013), Dovern (2013) and Giacomini et al. (2020).

Using data on household expectations, Binder (2017a, 2017b) highlighted the possibility of obtaining misleading results on information frictions from rounded and aggregated data. More recently, Giacomini, Skreta and Turen (2020) examined Bloomberg's survey of professional forecasters and found that, compared with the "inattentive" forecasters who do not update their forecasts every month, the updaters produce more accurate forecasts that are also more rational and informationally efficient.

In this paper, we take a novel approach treating zero revisions as censored observations for which the true values fall into a "region of imperceptibility" around zero, and reformulate the Nordhaus test to accommodate forecast revisions with many zeros.

2. Data and Methods

We use the individual forecasts of the U.S. real GDP growth rate, CPI inflation rate, and unemployment rate from the *Blue Chip Economic Indicators*, reported each month targeting the current and the following calendar year. The survey collects responses to the nearest 0.1%, which seems reasonable for macro forecasts. Our sample covers a 35-year period from Jan 1985 to Feb 2020, giving us a large number of terminal years and horizons over which forecast efficiency is tested. Considering that there are a maximum of 24 monthly forecasts for a calendar year, we keep only the forecasters with at least 240 non-missing forecasts. The resulting sample contains 33,589 observations from 72 forecasters. The heaping at zero revisions makes the cross-sectional distributions highly leptokurtic (see Figure 2).

The original Nordhaus (1987) test is based on a regression of forecast revisions on their lags. Let $r_{i,t,h}$ be the revision to individual i's forecast from horizon $h+1$ to h as observed in the data. If the forecasts are efficient, $r_{i,t,h+1}$ should be insignificant in the regression

$$r_{i,t,h} = \beta_0 + \beta_1 \, r_{i,t,h+1} + u_{i,t,h}, \tag{1}$$

where $u_{i,t,h}$ is a white noise. When using consensus forecasts, the equation is estimated by OLS or GLS. When using data pooled across individuals, horizons, and target years, we use weighted least squares and let β_0 to be individual specific. The weights are horizon-specific and proportional to the inverse of the variance of the OLS residuals. Following Giacomini et al. (2020), we also consider testing only with the forecasts that have been revised from the previous month (i.e., the updated forecasts). The results will differ if updaters are more (or less) efficient than non-updaters.

We extend the Nordhaus test by recognizing the difference between the actual ($r^*_{i,t,h}$) and the reported revisions ($r_{i,t,h}$). Our model is derived from the class of friction models discussed in Maddala (1983, ch. 6), but estimated using a 3-dimensional panel data. Specifically, we posit that a forecaster reports a non-zero revision $r_{i,t,h}$ only when the latent actual revision $r^*_{i,t,h}$ is large enough not to fall into the "region of imperceptibility." If the true revision is thought to be very small, a zero revision is reported. Thus,

$$r_{i,t,h} = \begin{cases} r^*_{i,t,h} \text{ if } r^*_{i,t,h} \notin [-\theta_{1,i,h}, \theta_{2,i,h}] \\ 0 \quad\; \text{if } r^*_{i,t,h} \in [-\theta_{1,i,h}, \theta_{2,i,h}] \end{cases}. \tag{2}$$

Our model differs from the standard two-limit Tobit model in that the values *inside* the interval $[-\theta_{1,i,h}, \theta_{2,i,h}]$ are censored in this application. The individual- and horizon-specific censoring thresholds ($\theta_{1,i,h}$ and $\theta_{2,i,h}$) are set to be the absolute value of the smallest observed negative and positive forecast revision, respectively.[4] We further assume that the relationship specified in the original Nordhaus test applies to the actual revisions $r^*_{i,t,h}$:

[4] About 8% to 12% of the thresholds are larger than 0.1 in magnitude. As robustness checks, we also experimented with several other ways of specifying these, including symmetric thresholds (-0.05/0.05, -0.1/0.1), as well as individual-specific but horizon-invariant thresholds based on observed maximum/minimum revisions. Our conclusions stay the same.

$$r_{i,t,h}^{*} = \beta_{0,i}' + \beta_1' \, r_{i,t,h+1}^{*} + u_{i,t,h}, Var(u_{i,t,h}) = \sigma_h^2, \tag{3}$$

where we allow the intercept to be individual specific and the variance horizon specific. The parameters are estimated by maximizing the likelihood function:

$$\begin{aligned}\mathcal{L} = &\prod_{r_{i,t,h} \neq 0} \frac{1}{\sigma_h} \phi\left(\frac{r_{i,t,h} - \beta_{0,i}' - \beta_1' \, r_{i,t,h+1}}{\sigma_h}\right) \\ &\times \prod_{r_{i,t,h} = 0} \left[\Phi\left(\frac{\theta_{2,t,h} - \beta_{0,i}' - \beta_1' \, r_{i,t,h+1}}{\sigma_h}\right) - \Phi\left(\frac{-\theta_{1,t,h} - \beta_{0,i}' - \beta_1' \, r_{i,t,h+1}}{\sigma_h}\right)\right],\end{aligned} \tag{4}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are normal density and distribution functions respectively. Note that apart from circumventing the heaping problem at zero revisions, in this formulation, there is no need to separately test the forecasts of the updaters – when there is no zero revisions in the data set, the model simply reverts to the standard test.

3. Results and Discussions

Table 1 contains the results from the standard Nordhaus tests and the friction model. When using data from all the forecasters including zero revisions, both the standard test and the maximum likelihood estimates of the friction model give largely similar results for all three variables – they are negative, statistically significant, but close to zero. Thus, these estimates suggest some but limited overreaction to news. However, with consensus data, we find the coefficients to be all positive and highly significant, cf. Coibion and Gorodnichenko (2015). With similar quantitative estimates as ours, Bordalo et al. (2018) reported this puzzling result on over-reaction at the individual level but under-reaction with consensus forecasts. Isiklar (2005) showed that substantial aggregation biases could result if forecasters are heterogeneous in their utilization of news. Clements (1997) provided a cogent explanation of the negative autocorrelation in forecast revisions in terms of random perception errors when weakly efficient forecasters process a series of small news. Using New York Fed's Survey of Consumer Expectations, Zhao (2019) found that some of these revisions are not justifiable by the latest available information, and are uninformative due to measurement errors in the absence of news.

The efficiency tests using only updated forecasts suggest somewhat higher degrees of inefficiency and over-reaction for all there variables. Giacomini et al. (2020) hypothesized that attentive forecasters are informationally more efficient than the inattentive ones. In Figure 3, using forecasts from 1985-2020 and the preliminary GDP values as the outturns, we see no perceptible difference in the root mean squared errors (RMSE) of the real GDP growth rate forecasts for the updated and all forecasts over each of the horizons. Though not reported, the results from inflation and unemployment forecasts were also the same. Thus, we do not find the updaters to outperform their "inattentive" peers in terms of both forecast efficiency and accuracy.

The friction model did not produce results significantly different from the conventional Nordhaus test. It is explained by the fact that the near efficient forecasts ($\beta_1' \simeq 0$) together with small monthly forecast revisions just prior to the month of zero revisions make $\beta_1' \, r_{i,t,h+1}$very small, which in turn, make the second component in (4) nearly invariant over individuals to have any impact on the parameter estimation. This result is consistent with the fact that the quality of the updated forecasts is almost the same as all forecasts, and the evidence in Dräger and Lamla (2012) that quantitative changes in expectations tend to be significantly smaller when the corresponding qualitative responses stay the same.

Note that the coefficient estimates in Table 1 are based on all 24 horizons. In order to check if over-reaction to news is the norm for all horizons, we report the test results for our three target variables in Table 2, separately for each horizon.[5] Note that unpredictable forecast revisions imply

[5] Bordalo et al. (2018) reported only for 12-month horizon using Blue Chip and SPF forecasts for a number of macro variables at quarterly frequency.

that fixed-target forecasts should look like a random walk – revisions should be spiky and not smooth, cf. Nordhaus (1987). Indeed, we find that many of the coefficients of lagged revisions are small, change their signs, and statistically insignificant. Even though there are interesting nuances in the sign and significance patterns in the horizon-specific regressions of the three variables, the degree of inefficiency is uniformly very small – suggesting adjustment to news in weeks rather than in months. Even though overreaction to news at each horizon is more common, the over-reaction to individual news is not as ubiquitous as suggested in Bordalo et al. (2018). For example, with GDP forecasts at horizons 9-15, the estimated coefficients are all persistently positive and many of them are statistically significant, suggesting under-reactions to news. A number of factors including the underlying data generating process, the timing of data releases, and the forecast horizon possibly mediate the precise response dynamics over horizons (cf. Lahiri (2012)).

To reiterate, non-zero revisions in the Blue Chip forecasts do not seem to be a good indicator of attentive forecasts, and as a result, revised fixed-target forecasts do not outperform the forecasts that often stay unrevised. This has the effect that treating zero revisions as censored observations has no impact on the outcome of the standard Nordhaus efficiency test.

References

BINDER, C. C. (2017a). "Consumer forecast revisions: Is information really so sticky?," *Economics Letters,* 161, 112–115.

——— (2017b). "Measuring uncertainty based on rounding," *Journal of monetary Economics,* 90, 1–12.

CLEMENTS, M. P. (1997): "Evaluating the rationality of fixed-event forecasts," *Journal of Forecasting,* 16, 225–239.

COIBOIN, O, and Y. GORODNICHENKO (2015). "Information Rigidity and the Expectations Formatin Process: A Simple Framework and New Facts", *American Economic Review,* 105 (8), 2644-2678.

DOVERN, J. (2013). When Are GDP Forecasts Updated: Evidence From a Large International Panel?" *Economics Letters*, 120 (3), 521-524.

DRAGER, L. AND M. J. LAMLA (2012). "Updating Inflation Expectations: Evidence from Micro-Data," *Economics Letters*, 117, 807-810.

GIACOMINI, R., V. SKRETA, AND J. TUREN (2020). "Heterogeneity, Inattention, and Bayesian Updates," *American Economic Journal: Macroeconomics,* 12, 282–309.

ISIKLAR, G. (2005). "On aggregation bias in fixed-event forecast efficiency tests," *Economics Letters,* 89, 312–316.

LAHIRI, K. (2012). "Comment", *Journal of Business and Economic Statistics*, 30 9(1), 20-25.

MADDALA, G. S. (1983). "Limited-Dependent and Qualitative Variables in Econometrics," *Cambridge University Press*.

NORDHAUS, W. D. (1987). "Forecasting Efficiency," *The Review of Economics and Statistics,* 69, 667.

PFAJFAR, D and E. SANTORO (2013). "News on Inflation and the Epidemiology of Inflation Expectations", *Journal of Money, Credit and Banking*, 45, (6), 977-1209.

ZHAO, Y. (2019). "Updates to household inflation expectations," *Economics Letters,* 181, 95–98.

Table 1. Efficiency of Blue Chip forecasts, all horizons pooled

Sample/Model	Real GDP	Inflation	Unemployment
Standard Nordhaus test, all forecasters	-0.024***	-0.059***	-0.107***
	(0.005)	(0.006)	(0.005)
Standard Nordhaus test, updaters only	-0.034***	-0.096***	-0.222***
	(0.008)	(0.009)	(0.010)
Friction model, all forecasters	-0.025***	-0.072***	-0.103***
	(0.005)	(0.006)	(0.006)
Standard Nordhaus test, consensus forecasts	0.417***	0.605***	0.396***
	(0.033)	(0.029)	(0.032)

* p<.05; ** p<.01; *** p<.001; standard errors are given in parenthesis below the coefficient of past revisions.

Table 2. OLS horizon-by-horizon efficiency tests.

Horizon	Real GDP	Inflation	Unemployment	Horizon	Real GDP	Inflation	Unemployment
1	-0.042	0.010	-0.135 *	12	0.201 *	0.208 *	-0.013
2	0.048	-0.101	-0.084 *	13	0.061	0.137 *	-0.049
3	-0.035	-0.058	-0.097 *	14	0.165 *	0.181 *	0.036
4	0.032	-0.124	-0.221 *	15	0.060	-0.094	0.022
5	-0.314 *	-0.033	0.090 *	16	0.012	-0.102 *	-0.069
6	0.064	0.054 *	-0.043	17	-0.033	-0.208 *	0.127 *
7	0.041	-0.013	0.020	18	-0.163 *	-0.099 *	-0.088 *
8	-0.097	-0.175 *	-0.574 *	19	-0.159 *	-0.096 *	-0.206 *
9	0.385 *	-0.015	0.518 *	20	-0.090	-0.236 *	0.043
10	0.101 *	-0.057 *	0.103 *	21	0.006	-0.161 *	0.182
11	0.047	-0.052	-0.001	22	-0.057 *	-0.069	0.071 *

The table reports the coefficient of the lagged revisions. * p<.05. The model includes individual dummies (omitted from the table). Each horizon is estimated separately with about 1300 observations.

Figure 1. Proportion of zero revisions by horizon and target variable

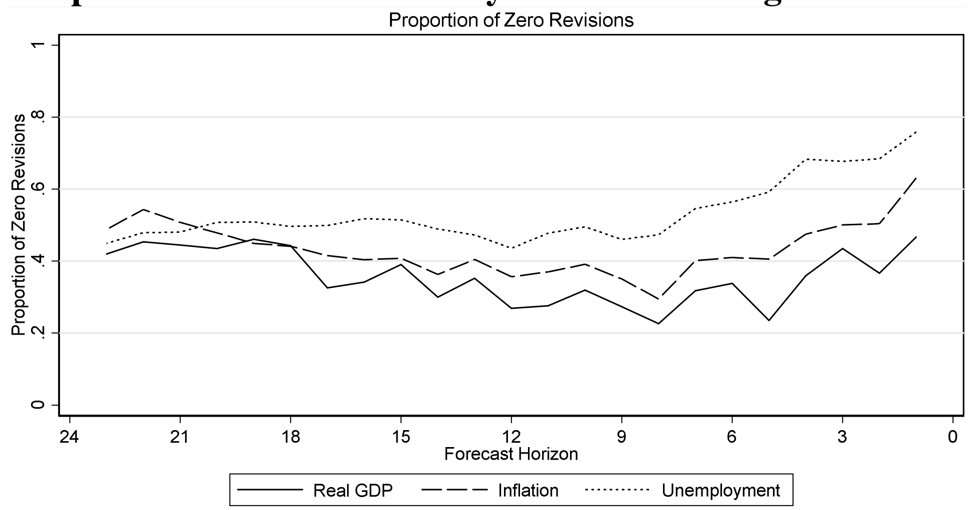

Figure 2. Distribution of forecast revisions

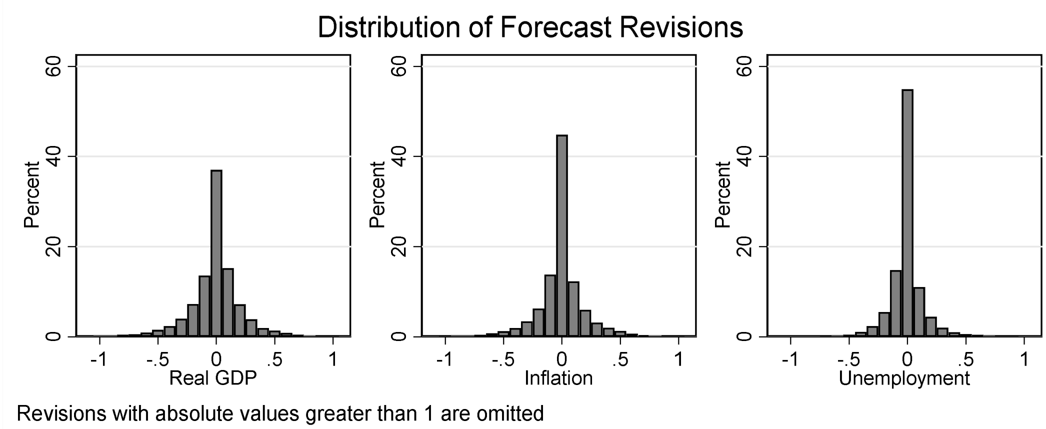

Figure 3. RMSE of GDP growth forecasts: all forecasts vs. updated