Freie Universität Berlin

Data Revisions to German National Accounts: Are Initial Releases Good Nowcasts?

Till Strohsal
Elias Wolf

School of Business & Economics
Discussion Paper

Economics

2019/11

Data Revisions to German National Accounts: Are Initial Releases Good Nowcasts?*

Till Strohsal †, [a], [b]
Elias Wolf‡, [b]
[a] *German Federal Ministry for Economic Affairs and Energy*
[b] *Freie Universität Berlin*

Data revisions to national accounts pose a serious challenge to policy decision making. Well-behaved revisions should be unbiased, small and unpredictable. This paper shows that revisions to German national accounts are biased, large and predictable. Moreover, using filtering techniques designed to process data subject to revisions, the real-time forecasting performance of initial releases can be increased by up to 17%. For total real GDP growth, however, the initial release is an optimal forecast. Yet, given the results for disaggregated variables, the averaging-out of biases and inefficiencies at the aggregate GDP level appears to be good luck rather than good forecasting.

Keywords: Revisions, Real-Time Data, German National Accounts, Nowcasting

JEL classification: C22, C53, C82, E66

*This paper has benefited from valuable comments and suggestions received from Dieter Nautz, Helmut Lütkepohl, Charlotte Senftleben, Mathias Kesting as well as from participants of the 38th International Symposium on Forecasting, Boulder, Colorado, the Freie Universität Berlin's School of Business & Economics faculty council meeting during the presentation of Till Strohsal's habilitation and the BMWi Brown Bag Lunch Seminar. The opinions in this paper are those of the authors and do not necessarily reflect the views of the German Federal Ministry for Economic Affairs and Energy.

†Author information: Till Strohsal, German Federal Ministry for Economic Affairs and Energy, Macroeconomic Projections, Freie Universität Berlin, Department of Economics, Email: till.strohsal@bmwi.bund.de, phone: +49 (0)30 18615 6659.

‡Author information: Elias Wolf, Freie Universität Berlin, Department of Econometrics, Email: eliaswolf@zedat.fu-berlin.de, phone: +49 30 838 66788.

1 Introduction

Economic policy decisions have to be made on the basis of the latest official statistics on the current economic situation. In business cycle policy, assessing the current state of the economy as weak may result in stimulating policy measures. In financial policy, tax revenue estimation and budgetary planning build on the latest data on GDP, private consumption, wages and salaries and further variables from the national accounts. Similarly, interest rate decisions by monetary policymakers depend on the current rate of inflation and the current output gap estimate. In all of these areas, it is crucial to have data that reflect the current economic situation as precisely as possible. Therefore, the ideal state would be knowing the final data values, or, in other words, the truth.

Unfortunately, this is not realistic. Instead, the latest national accounts data represent only a first estimate of the final value. In subsequent releases, the data are necessarily revised and sometimes substantially so. Reasons for revisions include lagged deliveries of source data that enter the national accounts or the fact that the source data themselves are revised.

Existing literature on US data documents that revisions can matter for policy decisions, see e.g. the survey of Croushore (2011). Orphanides (2001) shows that the optimal interest rate implied by a Taylor Rule varies drastically with different data vintages of inflation and real activity.[1] Ghysels et al. (2018) find the forecasting power of macroeconomic variables for returns to decline strongly when their unrevised real-time values are used instead of latest vintage data. Similarly, Diebold and Rudebusch (1991) show that the US composite leading index's forecasting performance for industrial production substantially deteriorates when using real-time data.

A comprehensive study of revisions to US national accounts is provided by Aruoba (2008). The results indicate that revisions do not satisfy simple desirable statistical properties. Significant biases and predictability of revisions imply that initial releases are not rational forecasts of the final values.

This paper makes two contributions. The first one is to test whether revisions to German national accounts are well-behaved. In line with Aruoba's results for US data, it turns out that revisions are often biased, are fairly large and significantly correlated with information available at the time when the initial release was published. The correlation with time t information implies predictability. Therefore, the second contribution is to go one step further by analyzing whether the real-time forecasting performance of initial releases can be improved empirically. This requires models which are designed to process data that are subject to revisions. Using the filtering approach of Kishor and Koenig (2012) and a number of restricted versions of it, the root-mean-square forecast error is reduced by up to 17% relative to the initial release. For a few cases, notably real GDP growth, the initial release is found to be an optimal forecast. However, given the results for disaggregated variables, the averaging-out of biases and inefficiencies at the aggregate GDP level appears to be good luck rather than good forecasting.

The rest of this paper is structured as follows. In Section 2, data revisions are defined and their desirable statistical properties are discussed. Section 3 briefly reviews the econometric models used to forecast the final data values. The real-time data are introduced in Section 4. Section 5 shows the main empirical results and Section 6 concludes.

[1] A contrasting result is given by Croushore and Stark (2001) who show that the identification of monetary policy shocks in structural vector autoregressions is quite robust with respect to different data vintages.

2 Defining Final Revisions and their Desirable Statistical Properties

Final revisions. The German Federal Statistical Office distinguishes between two types of revisions: ongoing revisions and benchmark revisions. *Ongoing revisions* are data-driven and can occur every quarter. Typical reasons for ongoing revisions are late data deliveries or revisions to the delivered data. For instance, industrial production figures are first published about 35 days after the reference period and are usually revised in the subsequent months. Another example is tax revenue statistics which in some cases become available to the statistical office only after several years. A specific characteristic of ongoing revisions to quarterly national accounts is that the final revision is carried out four years after the initial release, usually in August. Hence, the four quarterly growth rates of GDP in the year 2014 will undergo final revision in August 2018. Beyond four years, only *benchmark revisions* are undertaken. Benchmark revisions occur about every five years and, in contrast to ongoing revisions, are not data-driven. Instead, benchmark revisions include redefinitions of variables or the implementation of new calculation methods. Therefore, benchmark revisions can be interpreted as a redefinition of the truth, while ongoing revisions are an attempt to get closer to a given definition of the truth. Benchmark revisions cannot be anticipated and should not be considered when judging revisions. In order to limit the influence of benchmark revisions, we follow Aruoba (2008) and consider national accounts data as final after the last ongoing revision.

In the empirical analysis we also consider some monthly indicators. In contrast to the quarterly data, the final ongoing revision is often carried out after a period of less than 4 years. In particular, CPI undergoes final revision as early as 15 days after the initial release, industrial production in May next year and retail sales after 25 months and 15 days.[2]

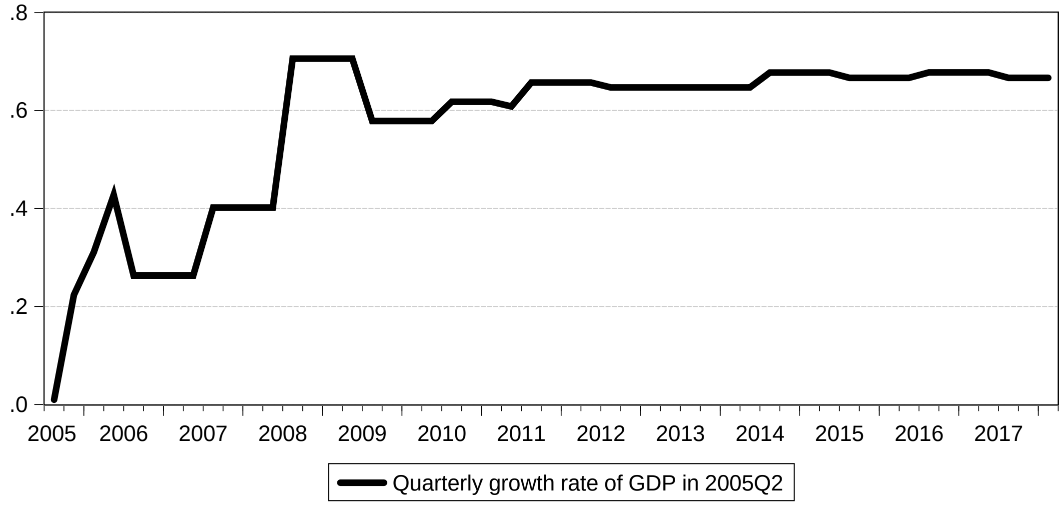
Throughout the paper, we will limit our attention to the properties of final revisions. The initial and final releases are denoted by x_t^{initial} and x_t^{final}, respectively. For quarterly national accounts data, x_t^{initial} is available 45 days after the reference quarter ("t+45 announcement"). Accordingly, final revisions are defined as

$$r_t^{\text{final}} = x_t^{\text{final}} - x_t^{\text{initial}} \tag{1}$$

Being aware of the existence of revisions is one thing. But do revisions actually matter? A first impression is provided in Figure 1 which documents the history of values assigned to the quarterly growth rate of real German GDP in 2005Q2. The first vintage is the initial release in 2005Q3, the latest vintage is February 14, 2018. To attach some meaning to the magnitudes, it is noted that the German economy is currently growing at a yearly rate of around 2%. Therefore, a typical quarterly growth rate may be around 0.5%. The first release suggested that the German economy was stagnating during the second quarter of 2005. However, the rate was substantially revised in subsequent periods. After about 4 years, the initial assessment has changed drastically from stagnation to very strong growth of around 0.7%. This example highlights the economic significance that revisions to national accounts can have. Figure 1 also shows that after around 4 years, revisions get very small and the time series somewhat converges to its true (final) value.

[2]More details on the data are provided in Section 4.

Figure 1   Quarterly Growth Rate of Real GDP



Note: This figure shows historical data vintages for the German real quarterly GDP growth rate of 2005Q2. The initial release was published in 2005Q3, the latest vintage included is February 14, 2018.

Three properties revision should have. The first desirable statistical characteristic of revisions is unbiasedness. Second, revisions should not be too large in terms of their variance. Third, revisions should be unpredictable using information that was available at the time when the initial release was published (I_{initial}). If revisions were predictable, the initial release would not be an optimal forecast of the final value. In that case, a more precise forecast must exist that has a smaller root-mean-square error. Formally, the three properties can be summarized as

P1: Unbiasedness: $\mathrm{E}[r_t^{\text{final}}] = 0$

P2: Variance: $\mathrm{V}[r_t^{\text{final}}]$ should be small

P3: Unpredictability: $\mathrm{E}[r_t^{\text{final}} | I_{\text{initial}}] = 0$

There are several possibilities for empirically measuring the properties of revisions, particularly with respect to P2 and P3. Regarding P1, for revisions to be *unbiased*, the sum of the upward and downward revisions of the statistical agency should amount to zero. This null hypothesis can be tested, if necessary, using robust standard error estimates. The desire to have a *small variance* is more difficult to formalize and needs a reference value to judge what small actually means. In the following analysis, we therefore consider the ratio of the variance of revisions to the variance of the underlying time series, termed noise-to-signal ratio. A value of 1, for instance, can be reasonably considered as large and economically relevant since it implies that the revisions vary by the same size as the actual time series. *Unpredictability* can be analyzed by the correlation between the revisions and the initial releases. The idea is that the initial release is a proxy for the information available at the time when the data was published, even though the information may have not been optimally exploited. A second measure for predictability which we use is first order autocorrelation in revisions.[3]

To check whether first releases provide a good picture of the current economic situation, the performance of alternative forecasts of the final value needs to be evaluated. In fact, it is more accurate to term an initial release a nowcast than to call it a forecast. Therefore, we will often refer to nowcasts in the remainder of the paper.

3 Models for Nowcasting German National Accounts

A number of approaches exist which are designed to process data that are subject to revisions, see Kishor and Koenig (2012). Many of them have a similar basic framework. Suppose the true, or final, value of the time series is generated by an AR(1) process

$$x_t = \beta x_{t-1} + \nu_t \quad . \tag{2}$$

The statistical agency observes the true value only with measurement error η_t

$$w_t = x_t + \eta_t \quad . \tag{3}$$

The statistical agency may now use its experience, expert knowledge and possibly also a model to evaluate w_t. Thereafter, the value y_t is published.

Naive Nowcast For the users of the official data, the first option is to ignore the possible occurrence of revisions and to naively set the published initial release at T equal to the nowcast $\hat{x}_T$ of the final value.

$$\hat{x}_T = y_T \tag{4}$$

Kalman Nowcast The alternative is to explicitly model data revisions, being aware that the published value represents a noisy signal. In the simplest case, the statistical agency provides just the value that it has observed itself

$$y_t = x_t + \eta_t \quad . \tag{5}$$

If a forecaster receives a surprisingly large y_T on the current edge of the sample, it is unclear how to interpret that value. Either the economic variable of interest, x_T, has actually taken on a high value due to a large shock ν_T in (2) or the large y_T is only a result of a sizable measurement error η_T in (5). Therefore, forecasters optimally update their expectations $\hat{x}_T = \hat{\beta} x_{T-1}$ which they had, doing so just the second before the initial release became available.

[3] As argued in Mankiw and Shapiro (1986) autocorrelation in final revisions is not always a direct indication of predictability because the final value of the $t-1$ period may not yet be available at t. However, an AR(1) process also implies correlation at lags $t-i$ with $i > 1$. Thus, it is likely that the finding of significant first order autocorrelation implies correlation with information that actually is available at t.

$$\hat{x}_T = \hat{\beta} x_{T-1} + \hat{\gamma}_K (y_T - \hat{\beta} x_{T-1}) \tag{6}$$

The idea of (6) is that the forecasters only incorporate a fraction of the surprise that they experience. The parameter γ_K can be estimated as $\hat{\gamma}_K = \hat{\sigma}_\nu^2 / (\hat{\sigma}_\nu^2 + \hat{\sigma}_\eta^2)$. The estimate is optimal in the sense that it minimizes $\mathrm{E}[x_T - \{\beta x_{T-1} + \gamma_N (y_T - \beta x_{T-1})\}]^2$.

Howrey Nowcast A common empirical property of revisions is that they are autocorrelated; see Aruoba (2008) for US data. As an extension of (6), Howrey (1978) proposes

$$y_t - x_t = h(y_{t-1} - x_{t-1}) + v_t \quad . \tag{7}$$

Similar to the Kalman model, the nowcast is

$$\hat{x}_T = \hat{\beta} x_{T-1} + \hat{\gamma}_H [y_T - \hat{h}(y_{T-1} - x_{T-1}) - \hat{\beta} x_{T-1}] \quad . \tag{8}$$

In the Howrey model γ_H is estimated as $\hat{\gamma}_H = \hat{\sigma}_\nu^2 / (\hat{\sigma}_\nu^2 + \hat{\sigma}_v^2)$.

Sargent Nowcast The approach of Sargent (1989) is to assume that the statistical agency itself filters its noisy observation w_t. However, Sargent allows for a white noise filtering error ξ_t which could be a technical mistake, a typo or something alike. In fact, only because of the filtering error does it make sense to the forecaster to filter a second time. In the Sargent model, the initial release is assumed to be of the form:

$$y_t = \hat{\beta} x_{t-1} + g(w_t - \hat{\beta} x_{t-1}) + \xi_t \quad . \tag{9}$$

The nowcast then becomes

$$\hat{x}_T = \hat{\beta} x_{T-1} + \hat{\gamma}_S (y_T - \hat{\beta} x_{T-1}) \quad . \tag{10}$$

The parameter γ_S can be estimated consistently as $\hat{\gamma}_S = (\hat{\sigma}_\nu^2 + \widehat{\mathrm{Cov}}(\nu, y - x)) / (\hat{\sigma}_\nu^2 + 2\widehat{\mathrm{Cov}}(\nu, y - x) + \hat{\sigma}_{y-x}^2)$.

The Howrey model does not allow for correlation between the shocks in (2) and (7) but for autocorrelation in the revisions. In the Sargent model, there is no autocorrelation in revisions but typically correlation between the revisions and the shock ν_t that drives the true value of x_t.[4]

[4]Intuitively, since the statistical agency uses (9) to filter the source data, the first release y_t will always include some of the noise η_t and some of the news ν_t. When the true data point becomes available, the revision will remove the noise which has been included in the initial release and will add the part of the news which has not been included. Hence, ν_t and the revision are correlated. More formally, the Sargent model assumes that the user of the official data cannot observe the source data w_t of the statistical agency. Rewriting (9) in terms of observable data, the analyst perceives the pre-filtering process as $y_t = \hat{\beta} x_{t-1} + g(x_t - \hat{\beta} x_{t-1}) + (\xi_t + g\eta_t)$. Together with (2), this implies that the revision has the form $y_t - x_t = (\xi_t + g\eta_t) - (1-g)\nu_t$. Note that the revision includes ν_t.

Kishor-Koenig Nowcast Kishor and Koenig (2012) propose an encompassing model which includes the Howrey model and the Sargent model as special cases. The revision process is defined as $y_t - x_t = h(y_{t-1} - x_{t-1}) + \epsilon_t - (1-g)\nu_t$, where $\epsilon_t = (\xi_t + g\eta_t)$ and the nowcast follows the equation

$$\hat{x}_T = \hat{\beta} x_{T-1} + \hat{\gamma}_K [y_T - \hat{k}(y_{T-1} - x_{T-1}) - \hat{\beta} x_{T-1}] \quad . \tag{11}$$

Note that for $g = 1$ the model becomes the Howrey approach and for $k = 0$ the model simplifies to the Sargent nowcast. Consistent estimates of the true parameter $\gamma_K \equiv g\sigma_\nu^2/(g^2\sigma_\nu^2 + \sigma_\epsilon^2)$ can be obtained as in the Sargent model. However, efficiency requires application of the SUR estimator.

4 Data: Deutsche Bundesbank's Real-Time Database

In order to compare alternative nowcast models, real-time data is needed. The Bundesbank provides a comprehensive database with vintages from 1995 onwards for quarterly data. For monthly economic indicators, data are available from 2005 onwards. We have selected a number of variables which receive particular attention in the business cycle analysis of the German government which is executed by the Federal Ministry for Economic Affairs and Energy. Specifically, we investigate the 9 variables in Table 1. GDP, private consumption, public consumption, investment in equipment and machinery and exports are real quarter-on-quarter growth rates. The German CPI, measured at a monthly frequency, is commonly analyzed as yearly growth rate. Industrial production and retail sales are month-on-month growth rates. Employment refers to the monthly absolute change in the number of people employed. All variables except the CPI are seasonally adjusted (SA) and most of the series are calendar adjusted (CA).

Table 1 Economic Indicators in Real-Time

	Sample period	Frequency	Unit	# obs.	SA	CA	Publication lag
GDP	1995Q2 2017Q4	quarterly	QonQ growth	75	yes	yes	45 days
Private Cons.	1995Q2 2017Q4	quarterly	QonQ growth	75	yes	yes	45 days
Public Cons.	1995Q2 2017Q4	quarterly	QonQ growth	75	yes	yes	45 days
Investment	1995Q2 2017Q4	quarterly	QonQ growth	75	yes	yes	45 days
Exports	1995Q2 2017Q4	quarterly	QonQ growth	75	yes	yes	45 days
CPI	2005M10 2018M02	monthly	YonY growth	99	no	no	no lag
Production	1995M02 2017M12	monthly	MonM growth	227	yes	yes	ca. 35 days
Retail Sales	2005M09 2017M12	monthly	MonM growth	100	yes	yes	ca. 30 days
Employment	1995M05 2018M01	monthly	MonM abs. changes	224	yes	no	ca. 66 days

Note: The sample period is predefined by the Bundesbank's real-time database. Publication lags reflect the institutional conventions of the German Federal Statistical Office. SA and CA stand for seasonal adjustment and calendar adjustment, respectively.

5 Results

5.1 Actual Properties of Revisions

It is fair to say that unbiasedness, small size and unpredictability are desirable properties of revisions. Yet, taking a closer look at the data, Table 2 shows that in many cases these properties are not satisfied. While quarterly national accounts data appear unbiased, the bias of the monthly economic indicators is statistically significant. This is shown in column 2, where bold numbers indicate significance at least at the 10%-level. With the exception of the CPI, the biases are also of economic significance. For instance, the monthly growth rate of German industrial production usually does 0.14 percentage points better than first thought.

What about the size of revisions? The noise-to-signal ratio, depicted in column 6, documents that revisions are large. Apart from CPI, the noise-to-signal ratio varies from about 0.5 to more than 1, implying that in some cases, the variation of revisions is as large as the variation of the underlying time series.

Generally, sizable revisions may well be unpredictable. If this is the case, they should be uncorrelated with the information available at the time when the initial value was released. Using the initial release itself as a proxy for this information, it can be seen from column 7 in Table 2 that for many variables the null hypothesis of unpredictability is significantly rejected. The negative sign of the correlations is in line with the noise interpretation of revisions; see Mankiw and Shapiro (1986). If the true value of a data point is observed only with noise, the revision will subtract the noise from the initial release and thereby induce the negative correlation. Note that, with one exception, the sign is also negative for insignificant correlations. Moreover, for almost all variables, there exists significant first order autocorrelation which is another indication of predictability of revisions, see the last column of the table.

Table 2 Revision Properties: Unbiasedness, Size and Predictability

	# obs.	Mean	Min	Max	SD	Noise/signal	Corr. with initial	AC(1)
Real GDP	75	0.03	-0.91	1.02	0.36	0.44	-0.07	0.12
Private cons.	75	0.08	-1.48	1.36	0.51	0.74	**-0.22**	-0.17
Public cons.	75	-0.01	-3.58	2.60	1.02	1.24	**-0.65**	**-0.24**
Investment	75	0.02	-5.18	4.72	2.01	0.60	-0.12	**-0.29**
Exports	75	-0.01	-2.12	2.48	1.09	0.43	**-0.32**	**-0.28**
CPI	147	**0.01**	-0.30	0.64	0.08	0.09	0.04	**-0.14**
Production	263	**0.14**	-2.78	2.22	0.80	0.56	**-0.57**	**-0.16**
Retail Sales	123	**0.30**	-4.32	6.57	1.47	1.09	**-0.66**	**-0.32**
Employment	224	**9.56**	-91.00	166.00	33.42	0.78	**-0.18**	**0.42**

Note: The table shows estimated statistics of final revisions. Numbers in bold indicate statistical significance at least at the 10%-level. SD refers to the standard deviation. The term noise/signal is defined as the variance of the revisions divided by the variance of the underlying time series and is thus a relative measure of the size of revisions. Corr. with initial measures the correlation of the revisions with the initial release, the latter being a proxy for the information available at the time of the initial release. AC(1) is the first order autocorrelation coefficient.

5.2 Nowcasting vs. Initial Data Release

Evidence-based economic and financial policy needs a precise idea of the current economic situation. The results on the statistical properties of revisions suggested that there may be some room to improve upon the initial releases. We verify this conjecture again by running the following Mincer-Zarnowitz regression

$$x_t^{\text{final}} = \alpha + \beta x_t^{\text{initial}} + u_t \tag{12}$$

where x_t refers to the final and the initial data value, respectively. Under the joint null hypothesis $\alpha = 0$ and $\beta = 1$ the forecast errors are purely random and thus unpredictable. Unbiasedness is given if $\alpha = 0$ and an efficient forecast would satisfy $\beta = 1$.

Given the predictability and bias of revisions in many cases, the results in Table 3 do not come as a great surprise. According to the p-values reported in the last column, the joint hypothesis of $\alpha = 0$ and $\beta = 1$ is rejected in most cases. In some cases, the rejection is caused by a bias ($\alpha \neq 0$), in other cases by inefficiency ($\beta \neq 1$) and sometimes by both. However, there are two notable exceptions: GDP and investment in equipment and machinery.

Table 3 Optimality Tests of Initial Releases: Mincer-Zarnowitz Regressions

		Unbiasedness		Efficiency	Optimal Forecast
	α	α=0	β	β=1	α=0, β=1
Real GDP	0.04	0.30	0.97	0.59	0.54
Private cons.	0.10	**0.04**	0.81	0.23	**0.10**
Public cons.	0.25	**0.00**	0.29	**0.00**	**0.00**
Investment	0.07	0.77	0.92	0.43	0.72
Exports	0.18	0.21	0.87	**0.04**	**0.10**
CPI	0.01	0.48	1.00	0.61	**0.07**
Production	0.14	**0.00**	0.74	**0.00**	**0.00**
Retail Sales	0.17	**0.00**	0.44	**0.00**	**0.00**
Employment	10.26	**0.00**	0.82	**0.03**	**0.00**

Note: This table shows parameter estimates of α and β and corresponding p-values for the single and joint hypotheses of the Mincer-Zarnowitz regression. Inference is based on heteroskedasticity and autocorrelation-consistent standard errors. Bold numbers indicate statistical significance at least at the 10%-level.

Concluding that the first release is often not an optimal prediction of the final value cannot be the end of the story. The more important question is whether superior alternative nowcasts can be found empirically. This exercise requires the use of real-time data and models which are able to handle data revisions. The candidate models we consider were described in Section 3: The Kalman nowcast, the Howrey nowcast, the Sargent nowcast and the Kishor-Koenig nowcast. The performance of the competing nowcast models is measured by the root-mean-square forecast errors. The forecast errors are defined as $\hat{x}_t - x_t^{\text{final}}$. The precision is analyzed in relative terms where the Naive nowcast, the initial release of the German Federal Statistical Office, serves as the benchmark model.

Table 4 summarizes the results. Values below 1 imply that the root-mean-square error (RMSE) of the alternative model is smaller than the one of the Naive nowcast. Values above 1 indicate that the initial release has a higher precision. In many cases, one or more alternative nowcast models outperform the initial release. The reduction in the RMSE is up to 17%, which is comparable to the gains in precision reported in Aruoba (2008) for US data. The improvement is generally higher for monthly data than for quarterly data.

Table 4 Nowcasting Competition: Initial Release vs. Alternative Models

	Real GDP	Private cons.	Public cons.	Investment	Exports	CPI	Production	Retail sales	Employment
Naive (StBA)	1	1	1	1	1	1	1	1	1
Kalman	1.17	0.94	0.83	1.17	1.00	1.02	0.86	0.84	1.01
Howrey	1.18	0.95	0.85	1.18	1.00	1.01	0.85	0.83	0.99
Sargent	1.02	0.96	0.84	1.13	0.96	1.02	0.88	0.85	1.02
Kishor-Koenig	1.14	1.08	0.95	1.06	1.11	1.02	0.88	1.07	0.99
Sample start	1995Q2	1995Q2	1995Q2	1995Q2	1995Q2	2005M10	1995M06	2005M09	1995M12
First nowcasted value	2000Q1	2000Q1	2000Q1	2000Q1	2000Q1	2007M01	2000M01	2007M01	2000M01
# of nowcasts	56	56	56	56	56	121	204	107	168
Last nowcasted value	2013Q4	2013Q4	2013Q4	2013Q4	2013Q4	2017M01	2016M12	2015M11	2013M12
Publication lag	1Q	1Q	1Q	1Q	1Q	1D	2M	2M	3M / 2M

Note: This table shows root-mean-square errors (RMSE) of alternative nowcast models relative to the initial release of the German Federal Statistical Office (Naive nowcast) which serves as the benchmark model. If the alternative model has a smaller RMSE than the benchmark model, the relative RMSE is below 1. Similarly, if the benchmark model performs better, the relative RMSE is above 1. The forecast error is defined as $\hat{x}_t - x_t^{\text{final}}$.

For those nowcasts where the alternative models deliver lower RMSEs, the smaller models with less parameters show the best performance. The simple Kalman nowcast has the lowest RMSE for private and public consumption while the same holds for the Howrey model nowcasts of industrial production, retail sales and employment. The greater flexibility of the Kishor-Koenig model does not seem to pay off for German national accounts data. Instead, the additional parameter uncertainty is likely to cause the lower forecast precision.

For some variables, the initial release appears fairly hard to beat. GDP and investment in equipment and machinery represent two very important cases where the alternative nowcast models show a lower precision than the Naive nowcast. Note that the finding is in line with the Mincer-Zarnowitz results from Table 3 where the initial releases already turned out to be optimal forecasts. Particularly due to GDP, the most important economic activity indicator, this result could be seen as quite encouraging. However, since many initial releases of the other variables were found not to be optimal forecasts, it may only be good luck that the biases and inefficiencies average out at the aggregate level of total GDP.

The results have shown that even simple models can improve the nowcasts of German national accounts. Yet there also exists evidence that the forecasting performance may be increased further. The lagged value used in equation (2) from Section 3 represents the second release of the second to last available quarter. Running the Mincer-Zarnowitz regressions for the second release provides more evidence for a rational forecast than was obtained for the first release.[5] Hence, if forecasts are not rational for some variables, it could increase the forecasting performance further to filter the second release as well.

The use of real-time data is crucial to compare the performance of different forecast mod-

[5]In fact, the βs are 1.01 (GDP), 0.92 (private consumption), 0.47 public (consumption), 0.92 (investment), 0.86 (exports), 1.00 (CPI), 0.84 (industrial production), 0.72 (retail sales) and 0.8 (employment).

els. Ignoring data revisions by estimating models with more recent vintages means to use information that was not actually available when the original forecast was conducted. Clearly, the additional information produces artificially high forecasting power, see e.g. Diebold and Rudebusch (1991) and Ghysels et al. (2018). In order to emphasize the importance of real-time data and revisions, we apply the flexible Kishor-Koenig model to (pseudo) nowcast all variables with vintages that became available only after the initial release. For quarterly data, we use vintages from 1, 2, 3 and 4 years after the initial release. For monthly variables, the datasets that were available 3, 6 and 12 months later as well as the final release are employed. As can be seen from Table 5, the gains in forecasting power are tremendous. The precision increases when a larger amount of information is put into the model. The reduction of the RMSE reaches values of up to 70%.

Table 5 Pseudo-Nowcasting with more Information

	GDP	Private cons.	Public cons.	Investment	Exports		CPI	Production	Retail sales	Employment
Naive (StBA)	1	1	1	1	1	**Naive (StBA)**	1	1	1	1
Vintage 1 year later	0.88	0.59	0.83	0.68	0.93	**3 months later**	0.30	0.54	0.85	0.92
Vintage 2 years later	0.78	0.65	0.68	0.61	0.73	**6 months later**	0.30	0.46	0.86	0.82
Vintage 3 years later	0.55	0.57	0.58	0.47	0.50	**9 months later**	0.30	0.40	0.79	0.79
Final Release	0.54	0.31	0.42	0.47	0.41	**Final release**	0.30	0.31	0.61	0.61

Note: This Table shows mean-squared forecast errors of the Kishor-Koenig model relative to the Naive nowcast (initial release). The Kishor-Koenig model is applied to data vintages which became available only after the initial release was published.

6 Conclusions

Data revisions are a natural phenomenon. They arise as a matter of course when official data are produced. One reason is that not all data needed to calculate a data point of a variable are available to the statistical office at the date of the initial release. These data gaps have to be filled by estimated values in one way or another. Moreover, the data delivered to the statistical office by third parties are usually revised, too. Also, statistical offices produce large amounts of different data and have to operate under budget restrictions. However, official data are used for policy decision making. Therefore, it is fair to require the initial releases to be unbiased and unpredictable.

In this paper, we have provided some evidence that revisions to German official data are often biased and predictable. Using a number of simple univariate models designed for data revisions, we show that the forecast power of the initial release can be increased. The effort of estimating theses models is limited and reasonable. Future research may well find more powerful models that can improve the forecasting performance further. For instance, Kishor and Koenig (2012) found multivariate versions of their model to perform well. The model of Jacobs and Van Norden (2011) represents another promising alternative.

At least two policy conclusions emerge from the current paper. First, the users of data should consider simple filtering of the initial release. On average, they will have a better idea on the actual situation for many important economic variables. Second, the providers of data may want to consider eliminating the predictability of revisions. Since this might require extensive institutional reform efforts, there exists an alternative: the data production process could be left unchanged. Instead, just a second before releasing the data point, the time series should be put through a simple filter.

References

Aruoba, S. B. (2008). Data revisions are not well behaved. *Journal of Money, Credit and Banking*, 40(2-3):319–340.

Croushore, D. (2011). Frontiers of real-time data analysis. *Journal of Economic Literature*, 49(1):72–100.

Croushore, D. and Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, 105(1):111–130.

Diebold, F. X. and Rudebusch, G. D. (1991). Forecasting output with the composite leading index: A real-time analysis. *Journal of the American Statistical Association*, 86(415):603–610.

Ghysels, E., Horan, C., and Moench, E. (2018). Forecasting through the rearview mirror: Data revisions and bond return predictability. *Review of Financial Studies*, 31(2):678–714.

Howrey, E. P. (1978). The use of preliminary data in econometric forecasting. *The Review of Economics and Statistics*, 60(2):193–200.

Jacobs, J. P. and Van Norden, S. (2011). Modeling data revisions: Measurement error and dynamics of true values. *Journal of Econometrics*, 161(2):101–109.

Kishor, N. K. and Koenig, E. F. (2012). VAR estimation and forecasting when data are subject to revision. *Journal of Business & Economic Statistics*, 30(2):181–190.

Mankiw, N. G. and Shapiro, M. D. (1986). News or noise: An analysis of GNP revisions. *Survey of Current Business*, 66:20–25.

Orphanides, A. (2001). Monetary policy rules based on real-time data. *American Economic Review*, 91(4):964–985.

Sargent, T. J. (1989). Two models of measurements and the investment accelerator. *Journal of Political Economy*, 97(2):251–287.

Diskussionsbeiträge - Fachbereich Wirtschaftswissenschaft - Freie Universität Berlin
Discussion Paper - School of Business and Economics - Freie Universität Berlin

2019 erschienen:

2019/1	WANGENHEIM, Jonas von English versus Vickrey Auctions with Loss Averse Bidders *Economics*
2019/2	GÖRLITZ, Katja; Merlin PENNY und Marcus TAMM The long-term effect of age at school entry on competencies in adulthood *Economics*
2019/3	BEZNOSKA, Martin Do Couples Pool Their Income? Evidence from Demand System Estimation for Germany *Economics*
2019/4	BÖNKE, Timm; Astrid HARNACK und Miriam WETTER Wer gewinnt? Wer verliert? Die Entwicklung auf dem deutschen Arbeitsmarkt seit den frühen Jahren der Bundesrepublik bis heute *Economics*
2019/5	WICKBOLDT, Clemens Benchmarking a Blockchain-based Certification Storage System *Information Systems*
2019/6	STANN, Carsten M. und Theocharis N. Grigoriadis Monetary Policy Transmission to Russia & Eastern Europe *Economics*
2019/7	PEEVA, Aleksandra Did sanctions help Putin? *Economics*
2019/8	ADAM, Marc Christopher Return of the Tariffs: The Interwar Trade Collapse Revisited *Economics*
2019/9	BRILL, Maximilian; Dieter NAUTZ und Lea SIECKMANN Divisia Monetary Aggregates for a Heterogeneous Euro Area *Economics*
2019/10	FISCHER, Benjamin; Robin JESSEN und Viktor STEINER Work Incentives and the Cost of Redistribution via Tax-transfer Reforms under Constrained Labor Supply *Economics*