Journal of Official Statistics, Vol. 24, No. 4, 2008, pp. 613–622

Seasonality in Investment Plans and Their Revisions

Alex Teterukovsky[1]

This article deals with application of seasonal adjustment methods to short-term forecasting of investments in the Business Enterprise sector. The approach is to forecast *the revisions*, i.e., the differences between the preliminary statistics based largely on planned investments, and the final ones. The magnitude of the revisions by the enterprises of planned quarterly investments depends not only on their economic activity and the economic situation, but also on timing of the investment surveys. These factors result in seasonal effects in revisions. We correct for the seasonality in revisions, so that the preliminary investment figures will not have to be substantially and systematically revised when the final figures arrive.

Key words: Official statistics; seasonality; forecasting; revisions.

1. Introduction

Suppose we ask someone on Thursday how much he/she is planning to spend during the weekend (let us say the answer is A), repeat this question on Saturday (answer: B) and ask on Monday (C) how much he/she in fact spent during the weekend. If we were probing on a regular basis, we would be able to construct a time series. The question is whether we can forecast $C - A$ (and therefore the "final" weekend spending C) as early as Thursday and $C - B$ on Saturday. If we know that the respondent is likely to underestimate his/her forthcoming weekend spendings, which sadly is often the case, the answer is yes. In other words, given the sequence of forecasts, preliminary and final observations, the early estimates of the final spending should be facilitated by taking into account the (possibly stochastic) pattern present in the revisions $C - A$, $B - A$ and $C - B$.

This article deals with forecasting a particular kind of revisions in statistics based on survey data, namely the revisions by enterprises of their own planned and realized economic activity. The set-up is similar to that in the example above, but instead of "spendings" we use a more appropriate term: "investments."

Investments represent the gross value spent by an enterprise on acquisition of tangible assets with an estimated service life of at least one year, reconstruction and improvement work. Investments are known to dominate the economic growth, both for individual industries and the economy as a whole. Three times a year Statistics Sweden collects the

[1] Email: statistika@telia.com

Acknowledgments: An early version of the article was presented at the Conference on Seasonality, Seasonal Adjustment and their Implications for Short-term Analysis and Forecasting, May 10–12, 2005, in Luxemburg. The author is grateful to Sven Öhlen, Stefan Svanberg and the participants for in-depth discussions. Special thanks go to Gunila Nockhammar, product manager for investment statistics at Statistics Sweden, for valuable comments.

data on realized and planned investments in the Business Enterprise sector for each of the four quarters. That is, for each quarter, three estimates are available. Differences between successive estimates are what I shall call *revisions*. Statistics are published three times a year and used by the National Accounts for calculating the GDP, by the Ministry of Finance for planning the national budget, by the Institute of Economic Research and by the industrial unions for business cycle and market analysis.

Much of the collected data is investment plans, and the differences between them and the outcomes can be significant. Triannual publication and users' demands for earlier statistics also mean that one is regularly forced to rely on plans when quarterly outcomes are not yet available. The combination of these factors renders it necessary to make reliable short-term forecasts.

There exist elaborate models for forecasting and modelling particular types of investments (e.g., Colwell and Park 1990; Assarsson, Berg, and Jansson 2004; Chacra and Kichian 2004). In our case the forecasts are needed on a rather detailed level regarding industries and types of investments. We are assisted by the fact that data on both realized and planned investments (enterprisewise forecasts) are collected. Our goal is to make use of all the data and produce a forecasting model, which could be relatively easily integrated into the production process of the official statistics.

Investments are known to be seasonal. In addition to that, because of the timing of surveys, not only the *levels* of realized and planned investments, but also their *revisions* by the enterprises, exhibit strong seasonal behavior. When the enterprises revise their investments plans and report the revised figures, the differences turn out to be of different magnitudes for different quarters. Why this is so will become clear in the next section where data are described. By accounting for the seasonality in revisions one is able to make a better use of preliminary figures, so that the estimates will not have to be substantially revised later. In other words, I will correct the investment estimates by forecasting the revisions.

Although there exist differences between industries and between types of investments, the problems and the proposed solutions are common. I shall use the investments in machinery and equipment in extracting and manufacturing industries as an example.

In Section 2 I describe the data material, its quality and timing. In Section 3 a forecasting problem, which is the concern of the article, is stated. The solution is presented and evaluated in Section 4. The approach utilizes TRAMO/SEATS (T/S), the standard seasonal adjustment software at Statistics Sweden (e.g., Gomez and Maravall 2001). A significant part of the T/S output, including values of estimated parameters and the results of diagnostic tests, is not included into this article, and can be obtained from the author. Section 5 concludes.

2. Data

Statistics Sweden collects data on investments by sending out the questionnaires to a sample of enterprises three times a year: in May, October and February. Each time, the enterprises are asked to provide *quarterly* figures (realized or planned) for all quarters of the year (Statistics Sweden 2007).

In the following notations, I use superscripts $p1$, $p2$ and $p3$ for the first preliminary figures, the revised figures and the final figures, respectively. Superscripts $f0$, $f1$ and $f2$ stand for zero, one and two-quarters ahead forecasts (plans), respectively, where by the zero-quarter ahead forecast we mean the "nowcast" for the ongoing quarter.

In May year T, the respondents are asked to provide preliminary figures $Q1_T^{p1}$ for the first quarter of T, and forecasts $Q2_T^{f0}$, $Q3_T^{f1}$ and $Q4_T^{f2}$ for the remaining quarters. In October, the respondents send in the revised figures $Q1_T^{p2}$ and $Q2_T^{p2}$ for the first two quarters, a preliminary figure $Q3_T^{p1}$ for the third quarter and a forecast $Q4_T^{f0}$ for the fourth quarter. In February year $T+1$, the respondents are asked to provide the investment figures $Q1_T^{p3}$, $Q2_T^{p3}$, $Q3_T^{p3}$ and $Q4_T^{p3}$, i.e., the "final" figures for each quarter of year T. This set-up is displayed in Figure 1.

Although the data collected on any of the three occasions are actually a mixture of forecasts, preliminary figures, and final figures, they do exhibit stable, albeit different, seasonal patterns (see Figures 2–4).

3. Problem: Correcting the Nowcasts

The investment figures are needed by the Ministry of Finance for preliminary national budget planning. This planning is done in August when the most recent data available are those from the May questionnaires (Figure 5). It necessitates the use of forecasts $Q2_T^{f0}$, $Q3_T^{f1}$ and $Q4_T^{f2}$ when estimating the second, third and fourth quarter's investments. It has been noted, however, that in May enterprises tend to misjudge the planned investments, even when making nowcasts for the ongoing second quarter. Figure 6 displays the discrepancies for each quarter, between the investments reported in May, those reported in October, and their final values reported in February next year. Observe that, for example, the second quarter is systematically significantly overestimated in May. Clearly, a correction is necessary, one which would prevent large errors in the national budget, which is finalized in October. Hence I put forward a problem:

"*How are we to correct the nowcasts and thus produce early estimates of the investments, which will not deviate too much from the revised and final figures?*"

Data on investments are used for the calculation of the GDP. The EU requirement to publish an early "flash" estimate of the GDP gives rise to similar complications, as can be seen from Figure 5. In order to produce better "flash" estimates of the GDP for the second and the fourth quarters, one again has to solve a forecasting problem.

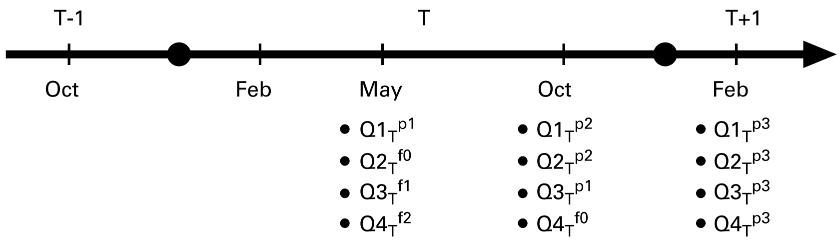

Fig. 1. Investment surveys for year T

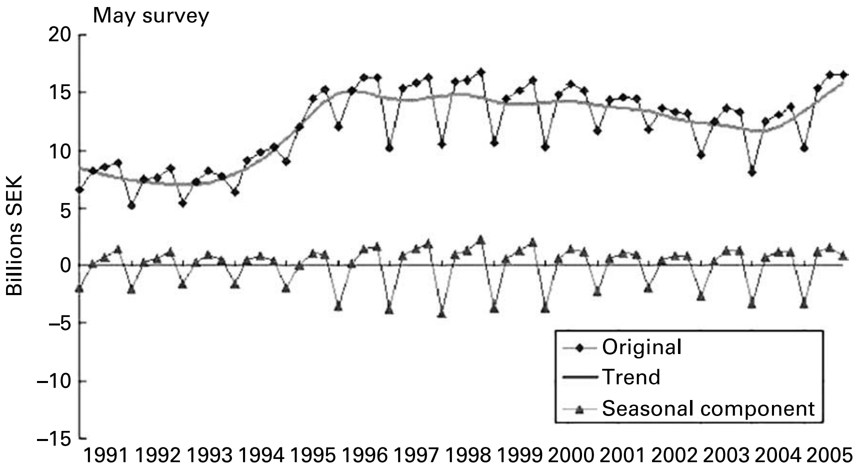

Fig. 2. Investments. Q1 – preliminary, Q2, Q3 and Q4 – planned

Below I describe a solution based on seasonal decomposition. I start with a simplified presentation of model-based seasonal adjustment by T/S.

4. Solution

4.1. Model-Based Seasonal Adjustment by TRAMO/SEATS

T/S is a freely distributed software for extracting (estimating) trend and seasonal components from a time series. In order for the estimates to be reliable, the original time series are first brought into a stationary form. This transformation (also called preadjustment) is achieved in TRAMO by taking logarithms, differencing the series at various lags, estimating and removing effects of extreme observations (outliers), calendar

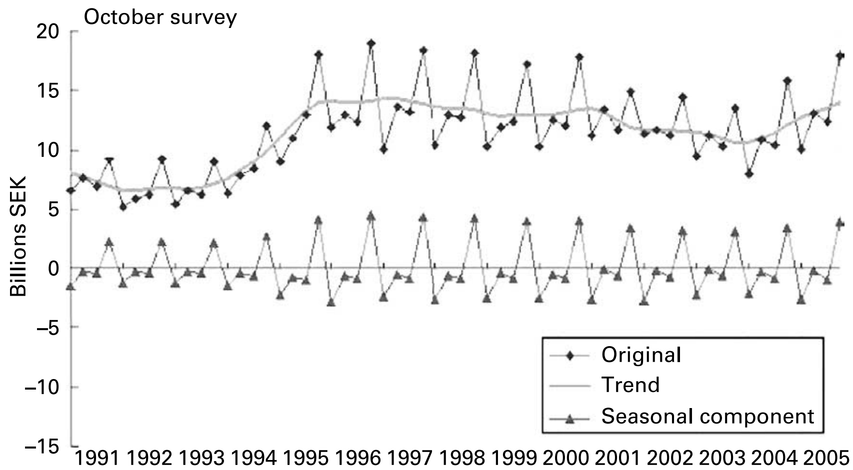

Fig. 3. Investments. Q1 and Q2 – revised, Q3 – preliminary, Q4 – planned

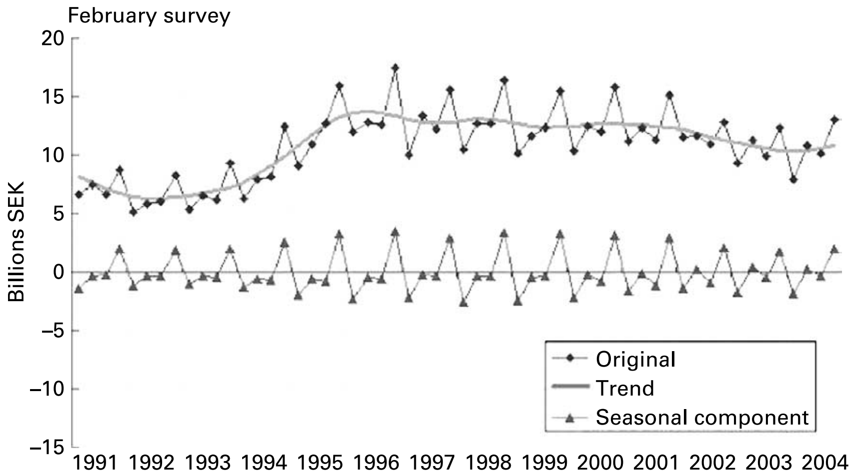

Fig. 4. Investments. Q1, Q2, Q3 and Q4 – "final"

effects and effects of specified external variables. After the preadjustment is done, a stationary Autoregressive Moving Average (ARMA) model with a small number of parameters is fitted. In the next step, SEATS decomposes the estimated ARMA model into seasonal and nonseasonal components and makes componentwise forecasts. A number of statistical tests are also performed, the results of which serve to diagnose the quality of estimation. T/S is a standard software at Statistics Sweden for seasonally adjusting economic time series. It is therefore desirable to use the same software for related problems, such as the one considered here.

4.2. Correcting the Nowcasts

Consider the second quarter and the May questionnaires. The solution is similar for the other quarters. As we can see in Figure 6, in May the investments for the second quarter are regularly overestimated. However, the magnitude of this error can be better assessed when taking into consideration the errors of the investment figures for the remaining quarters,

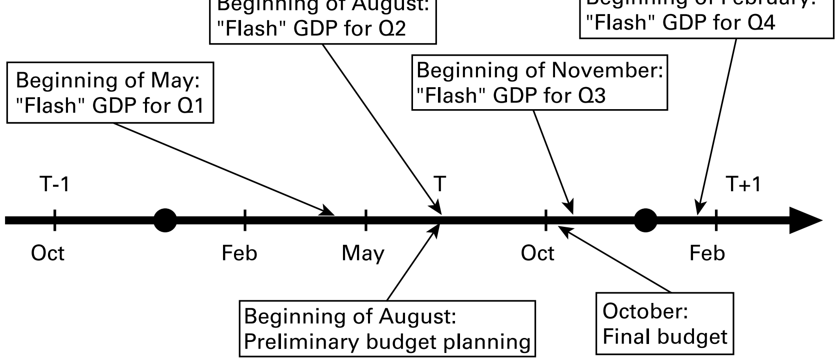

Fig. 5. National budget and "flash" GDP timing, year T

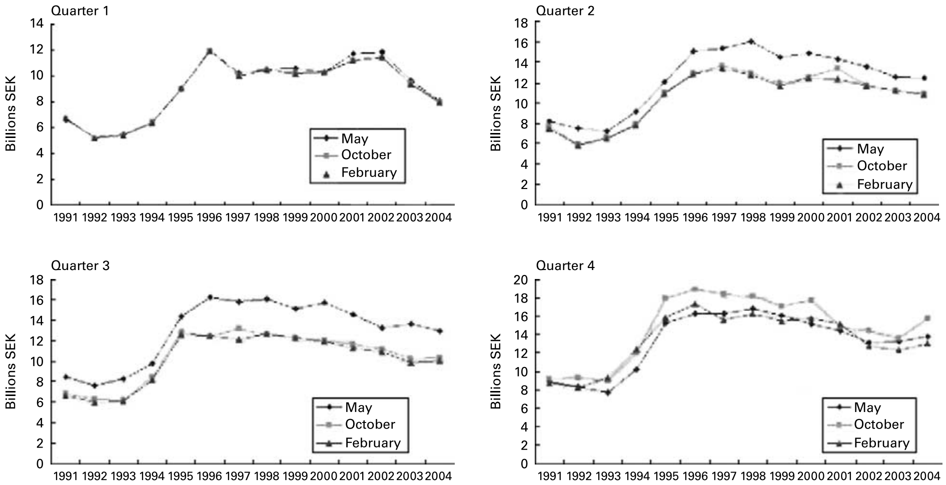

Fig. 6. Discrepancies between planned and realized investments

and the economic situation in industry. This becomes clear if one realizes that the errors are autocorrelated, e.g., a large underestimation of the investment for a certain quarter typically leads to overestimation for the remaining ones. It is also true that enterprises tend to overestimate planned investments when the economic situation is favourable and *vice versa*.

Consider the differences between the "final" February figures and the early ones from the May questionnaires:

$$\ldots, Q1_T^{p3} - Q1_T^{p1}, Q2_T^{p3} - Q2_T^{f0}, Q3_T^{p3} - Q3_T^{f1}, Q4_T^{p3} - Q4_T^{f2}, \ldots \quad (1)$$

These differences constitute a time series with an easily graspable seasonal component: the first quarter refers to the revisions of the preliminary figures, the second quarter refers to the errors of the nowcasts, and the third and fourth quarters refer to the forecast errors with different horizons. The trend of these series, if it exists, should be linked with the *level* of investments, since we expect (a) the magnitude of forecast and revision errors of the variable to be proportional to the variable in question and (b) the economic situation (expressed in, for example, growth of investments, capacity utilization or industrial production) to have an effect on forecast and revision errors. Our main approach is to model and forecast the series (1).

An alternative and in some sense a more general approach is to regress the final February figures $\ldots, Q1_T^{p3}, Q2_T^{p3}, Q3_T^{p3}, Q4_T^{p3}, \ldots$ on the preliminary May figures $\ldots, Q1_T^{p1}, Q2_T^{f0}, Q3_T^{f1}, Q4_T^{f2}, \ldots$ and to estimate the ARIMA model for the regression residuals. In fact, when modelling differences (1), we effectively enforce the regression coefficient to be equal to -1. The reason behind this restriction is that the revisions (1) themselves constitute a series which I believe it is meaningful to analyze. The seasonality in these series stems from and is controlled by the particular time schedule used for the surveys. By estimating the seasonal effects in revisions, we enable assessment of the timing of surveys and therefore of the quality of the data. Finally, focusing on revisions instead of on the original series gives us a possibility of using more observations, evading complications with possible structural breaks in the original data. Nevertheless, since the main goal is forecasting the original series, an alternative approach allowing for estimation of the regression function linking the preliminary figures with the outcomes is also evaluated.

Our first step is to decompose the series (1) in T/S to obtain the trend and the seasonal components (Figure 7). These components per quarter together with the original data are displayed in Figure 8.

Note that for the first quarter the estimated trend and the seasonal component of the differences practically take out each other. Clearly, the preliminary figures for the first quarter that we have in May are very close to the final figures. This is not the case for the other quarters. Our concern now is the second quarter, and we are going to forecast the revisions of the May nowcasts of Q2. One way to do this is to rely on the T/S forecast, which is the sum of a trend forecast and a forecast of the estimated seasonal component. The second quarter's May nowcasts are to be corrected by adding the forecasted seasonal component (this forecast is done automatically by SEATS) and then adding the forecasted trend.

Regarding trend, we choose between two possibilities: take the automatic SEATS forecast (as we do with the seasonal component) or forecast the trend separately. Depending on type of investments, it is sometimes possible to find external variables that explain well the planning of investments for a particular quarter. Trend extrapolating can then be done separately by regressing the estimated trend for a given quarter on these external variables. If we restrict ourselves to a particular quarter, the number of observations becomes small, and only the simplest models can be fitted. In Figure 9 we plot the by SEATS estimated trend of (1) together with planned investments and the capacity utilization for Q2.[2] The capacity utilization and the level of planned investments seem to have a causal relationship with the nowcast errors. The trend forecast is obtained by means of linear regression, with i.i.d. error terms, of the estimated trend (for Q2 only) on these two variables.[3]

The forecasts are evaluated with the aid of the normalized root mean squared error (RMSE):

$$\frac{\sqrt{\frac{1}{N}\sum_{T}\left(\widehat{Q2}_T - Q2_T^{p3}\right)^2}}{\overline{Q2^{p3}}} \times 100\%$$

where $\widehat{Q2}_{\mathrm{T}}$ is a forecast for Q2 year T, N is a number of years (= 10 in this example) and $\overline{Q2^{p3}}$ is the arithmetic average of $Q2_T^{p3}$ over the evaluation period.

The RMSE of the second quarter's investments estimate obtained after the correction (adding the forecast of revision to a preliminary May figure) is 4.6%, as compared to the initial RMSE of 18.4% for the preliminary figures themselves. Even more important is the

[2] All variables are brought to the the same scale for displaying purposes.
[3] We use the variables that correlate well with revisions for this particular type of investments. When such variables are not found (for other type of investments), we rely on trend forecasts produced by SEATS.

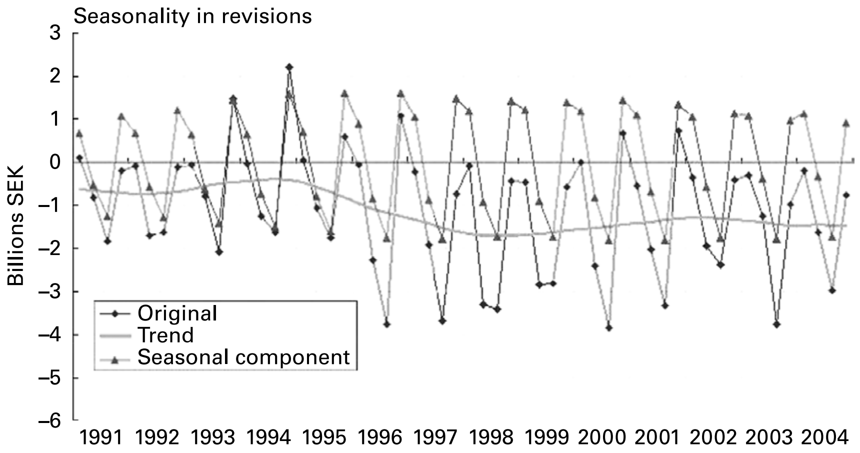

Fig. 7. Difference between May and February figures

fact that the systematic (i.e., seasonal) error component is removed. In Figure 10 we plot the original nowcasts, the corrected values and the outcomes.

For comparison, we evaluate an alternative direct approach to the problem, treating the preliminary May figures as a regression variable in T/S, hoping that differences between seasonality in the preliminary and the final series will be captured by the ARIMA model. This way, we move our attention from a revision study to solely forecasting. T/S was applied to the final figures, with preliminary figures included as regression variables preadjusted for by TRAMO. The default Airline model was enforced. No outliers were detected and the diagnostic tests were passed at the 99% level. The RMSE of SEATS forecasts for this model was 7.8%. The RMSE of simple T/S ARIMA forecasts not using the May figures as regression variables was 12.3%. We conclude that collecting the preliminary figures in May is worthwhile only if a correction is performed (otherwise

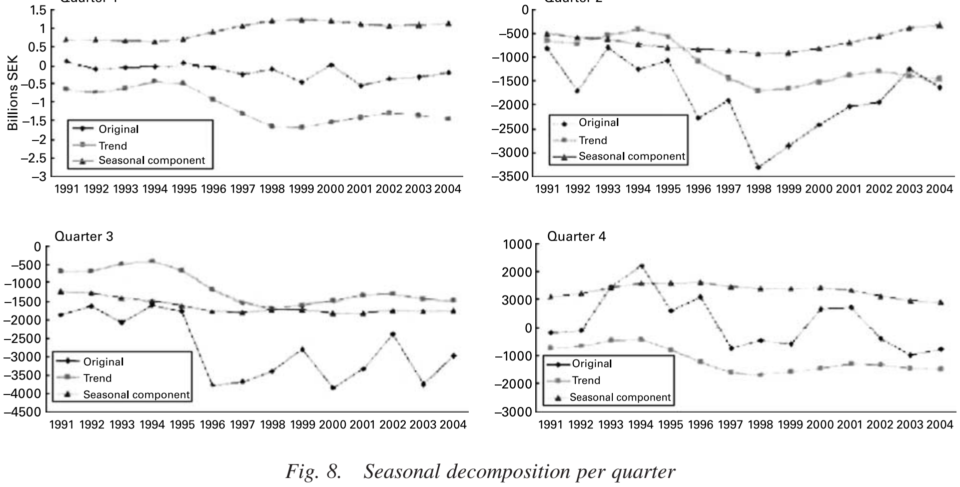

Fig. 8. Seasonal decomposition per quarter

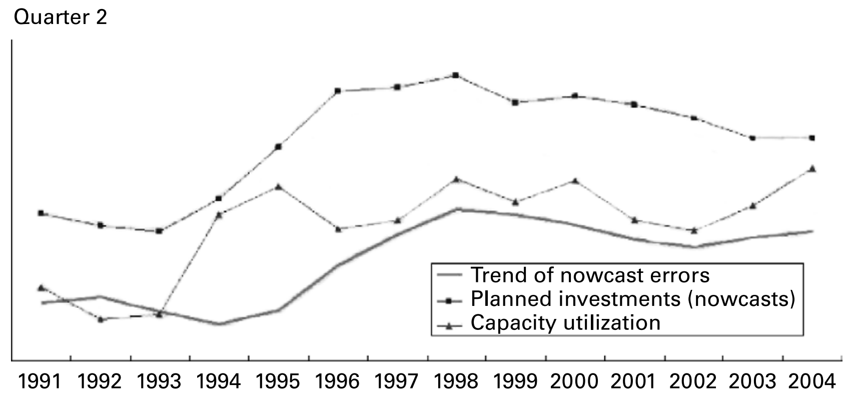

Fig. 9. Estimated trend of May nowcast errors and correlated indicators

simple ARIMA forecasts of investments would suffice). Our main approach based on modelling revisions has for our data outperformed the direct modeling of levels.

5. Final Comments

The solution presented in this article is implemented with the aid of standard seasonal adjustment software. This makes it possible to easily integrate the proposed method into the regular process of official statistics production. Using standard software has obvious advantages but also certain drawbacks. Our data do exhibit, as should be expected from their construction, season heteroscedasticity: the variance is smallest for the first quarter, much larger for the other quarters, and largest for the fourth quarter. Season heteroscedasticity is not explicitly accounted for by T/S. How to properly deal with such series is an interesting question (see a paper by Tripodis and Penzer (2006)).

I have analyzed and forecasted, on an aggregated level, the revisions by enterprises of their investment plans. Forecasting was assisted by the presence of a seasonal component in the data – both in levels and in revisions. Seasonality in revisions is to some extent "artificial" and stems from the data collection procedures: the data are a mixture of plans

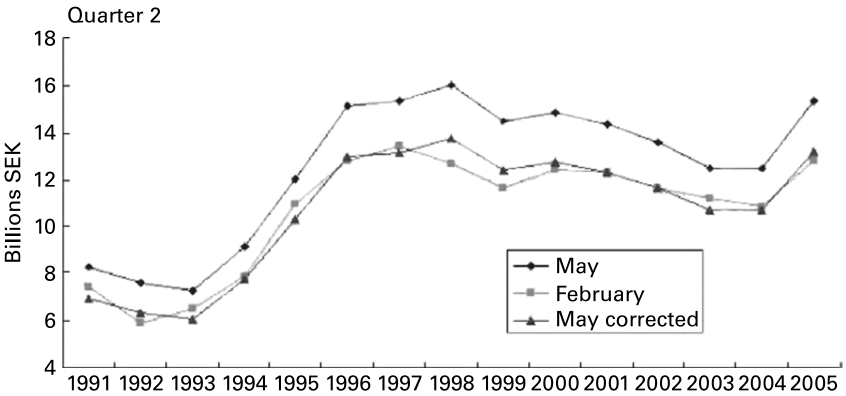

Fig. 10. Corrected nowcasts

with varying horizons and actual outcomes, and the revisions in the outcomes are naturally expected to be smaller than errors in planning. It is also human nature to underestimate future investments when the economic situation is unfavourable and overestimate them when the opposite is true.

In fact, the process of estimating any economic variable starts with forecasts, which subsequently turn into preliminary, revised and definite outcomes. Differences between all these figures contain valuable information. Analyzing the process in its entirety can provide an insight into the quality of the statistical variable in question. For an integrated quantitative assessment of the quality of the economic statistics, based on combining the forecast errors and the revisions, see Öller and Teterukovsky (2007).

The focus of the current article is not, however, on quality assessment but on forecasting, triggered by the users' demands for earlier statistics. The National Statistical Offices are often forced to make use of the, in many cases, uncertain preliminary figures. Forecasting revisions can therefore become an important subject in official statistics.

6. References

Assarsson, B., Berg, C., and Jansson, P. (2004). Investment in Swedish Manufacturing: Analysis and Forecasts. Empirical Economics, 29, 261–280.

Chacra, M. and Kichian, M. (2004). A Forecasting Model for Inventory Investments in Canada. Working Paper 39, Bank of Canada.

Colwell, P. and Park, H.Y. (1990). Seasonality and Size Effects: The Case of Real-Estate-Related Investment. The Journal of Real Estate Finance and Economics, 3, 251–259.

Gomez, V. and Maravall, A. (2001). Seasonal Adjustment and Signal Extraction in Economic Time Series. A Course in Time Series Analysis, D. Pena, G. Tiao, and R. Tsay (eds). New York: L. Wiley and Sons.

Statistics Sweden (2007). Business Investments. http://www.scb.se/ NV0801-EN.

Tripodis, Y. and Penzer, J. (2006). Single Season Heteroscedasticity in Time Series. Journal of Forecasting, 26, 189–202.

Öller, L.-E. and Teterukovsky, A. (2007). Quantifying the Quality of Macro-Economic Variables. International Journal of Forecasting, 23, 205–217.

Received August 2006
Revised June 2007