Statistics Norway
Statistisk sentralbyrå

Terje Skjerpen and
Anders Rygh Swensen

Forecasting Manufacturing Investment Using Survey Information

Terje Skjerpen and
Anders Rygh Swensen

Forecasting Manufacturing Investment Using Survey Information

Statistisk sentralbyrå • Statistics Norway
Oslo–Kongsvinger 1997

Standardtegn i tabeller	**Symbols in tables**	**Symbol**
Tall kan ikke forekomme	Category not applicable	.
Oppgave mangler	Data not available	..
Oppgave mangler foreløpig	Data not yet available	...
Tall kan ikke offentliggjøres	Not for publication	:
Null	Nil	-
Mindre enn 0,5 av den brukte enheten	Less than 0.5 of unit employed	0
Mindre enn 0,05 av den brukte enheten	Less than 0.05 of unit employed	0,0
Foreløpige tall	Provisional or preliminary figure	*
Brudd i den loddrette serien	Break in the homogeneity of a vertical series	—
Brudd i den vannrette serien	Break in the homogeneity of a horizontal series	|
Rettet siden forrige utgave	Revised since the previous issue	r

ISBN 82-537-4374-2
ISSN 0806-2056

Emnegruppe
10.90 Metoder, modeller, dokumentasjon

Emneord
Industri
Investeringer
Prognosemetoder
Prognosemodeller
Prognoser

Design: Enzo Finger Design
Trykk: Statistisk sentralbyrå

Abstract

Terje Skjerpen and Anders Rygh Swensen

Forecasting Manufacturing Investment Using Survey Information

Reports 97/3 • Statistics Norway 1997

Forecasting of realized investments in the Norwegian manufacturing sector are conducted utilizing survey data. The first time realized investments for an arbitrary year is forecasted in the second quarter in the year preceding the investment year. One has then just obtained the first preliminary estimate from the survey. New preliminary estimates are obtained through the next six quarters and the forecasts for final investment are accordingly updated at these stages. Different forecasting methods/models are compared with regard to the precision of forecasts using criteria which are relative versions of the wellknown RMSE and MAPE measures. It turns out that some of the simpler methods/models outperform the more complicated ones.

Keywords: Forecasting, investment.

Contents

1. Introduction*

The aim of this paper is to find «optimal» forecasting models/methods for total investment in the manufacturing sector using data from the quarterly investment survey which is conducted by Statistics Norway. Investment in the manufacturing sector is an important macroeconomic variable which is hard to forecast and survey information may be an important data source and supplement toother data. The Norwegian investment survey contains seven preliminary estimates for the value of the investments in a given year together with the realized investments. Formally we will treat this as eight different variables. It should be emphasized that all the eight variables are given in nominal terms. An analysis in real terms requires that the investment survey data are combined with other data sources.

Final investments are forecasted at seven different stages. At the three first stages we are in the year before the investment year and at the four last stages we are in the investment year. The first preliminary estimates are more imprecise than the latter ones and besides they indicate a systematic underpredicition. Thus to obtain «optimal» forecasts at the early stages it is important to correct for this bias. At the end of the year preceding the investment year the firms are setting up their budgets and as a result of this the preliminary estimates given in the investment year itself are of a much better quality.

Altogether we consider 13 models/methods for forecasting of realized investments. These are labelled alphabetically from A to M. Some of these methods have been used for a long while and are very easy to implement whereas others to our knowledge has not been applied before, at any rate not on Norwegian data. Some of the methods are partial in the sense that they only utilize information from some of the surveys, while in others we use all surveys given that they are available. A second distinction is whether the forecast is based on an explicit parametric model or not. For the first two models, i.e. A and B, no explicit model is assumed. A third distinction is whether one allows for time varying parameters or not. In the methods labelled E and J we allow the constant terms to develop according to random walk processes. A third way to distinguish between the models is whether we assume a data generating process for realized investment itself or not. For instance in our two methods based on VAR-modelling, which we have labelled method L and M, realized investment is treated symmetrically with the preliminary estimates. Another example is the methods labelled J and K. These methods typically pool information from the surveys with information from an ARIMA-model of realized investments. For the pooling of the different forecasts we use the Kalman filter, which is known to be well suited for such a task.

The rest of the paper is organized as follows. Since the design of the data is a very important feature for this empirical analysis we give a description of the data already in the next section. In section 3 we make an overview of the different methods/models which we are applying for forecasting purposes. In order to compare the forecasting capabilities of the different methods we make, in section 4, use of two criterions which are relative versions of RMSE (Root Mean Square Errors) and MAPE (Mean Average Prediction Errors). The conclusions are presented in section 4.

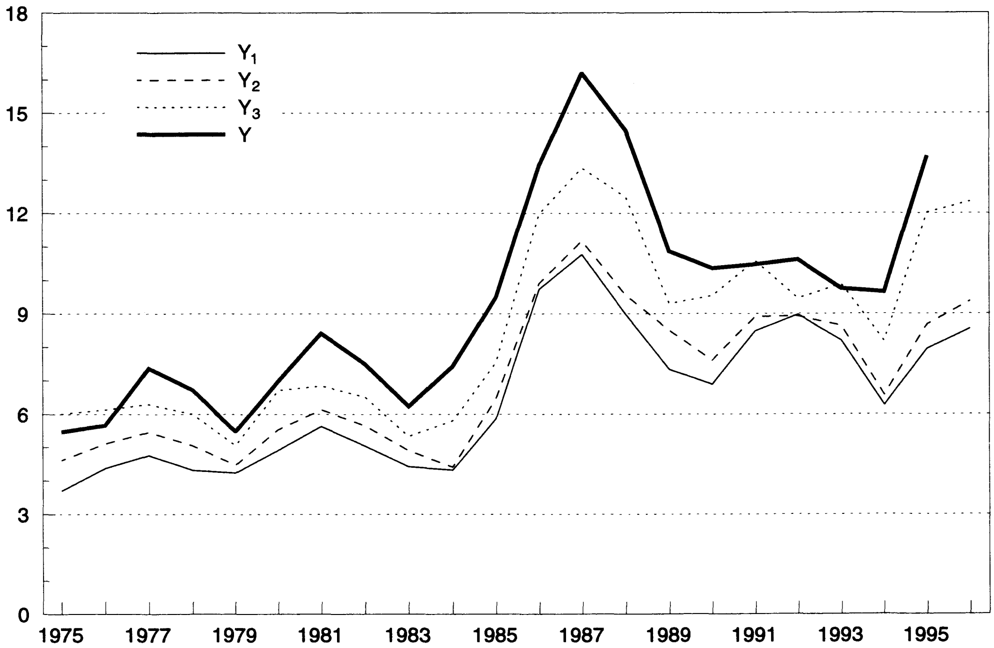
The main empirical result of this paper is that some of the simplest methods performs rather well compared to more complicated ones. Besides we find it hard to improve on the preliminary estimates obtained from the surveys in the investment year.

* The authors would like to thank Torbjørn Eika for comments and Marit Vågdal for the final typing of the report. The authors are solely responsible for any remaining errors or weaknesses.

2. Data

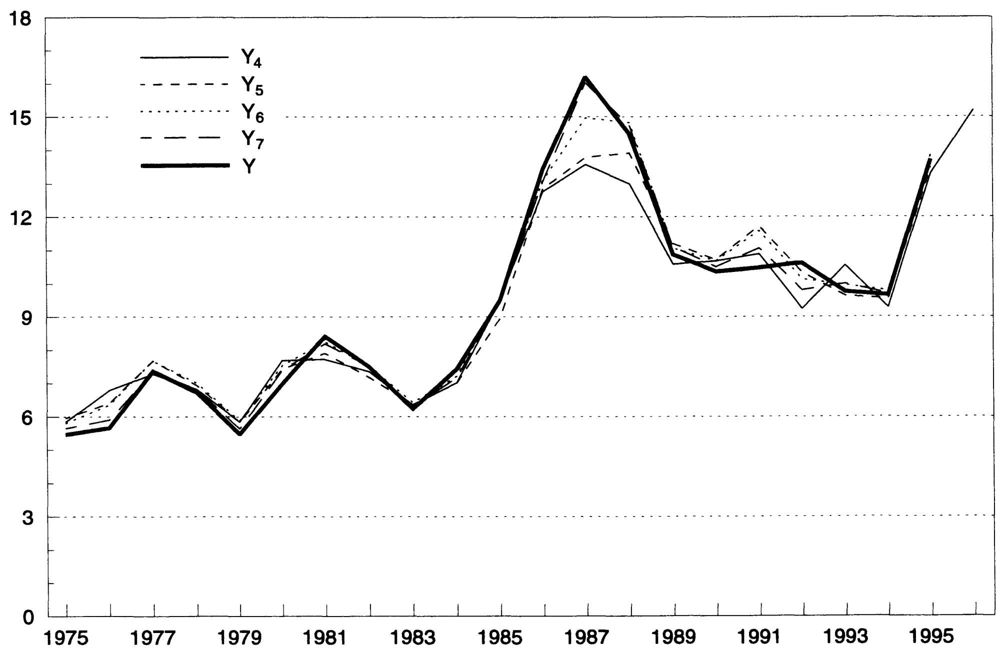
The data from the investment survey consists of eight variables. Let the variable Y_t denote final investment for the calendar year t. Values on this variable are obtained in the first quarter in the year succeeding year t. Before this point of time seven preliminary estimates are supplied by the firms. The three first preliminary estimates are obtained in the year preceding the investment year. The preliminary estimates obtained in the second, third and fourth quarter of year t-1 are labelled $Y_{1,t}$, $Y_{2,t}$ and $Y_{3,t}$ respectively. During the investment year four new estimates of the investment for year t become available, one in each quarter. Values for these variables are labelled $Y_{4,t}$, $Y_{5,t}$, $Y_{6,t}$ and $Y_{7,t}$ respectively. The values $Y_{5,t}$, $Y_{6,t}$ and $Y_{7,t}$ can be viewed as beeing the sum of two components, i.e. realized investments for the first part of the year and an estimate of the investment for the remaining part of the year. Another way of looking at the data which may be more fruitful from a time series analysis point of view is to focus on which values of the variables that become available through an arbitrary calendar year. During year t one obtains $(Y_{t-1}, Y_{4,t})$, $(Y_{5,t}, Y_{1,t+1})$, $(Y_{6,t}, Y_{2,t+1})$ and $(Y_{7,t}, Y_{3,t+1})$ in the first, second, third and fourth quarter respectively. This way of structuring the data is utilized in some of our forecasting methods. The historical time series for the eight series are reported in table 3 in Appendix A and in figures 1a and 1b. More data is now available, but an extension of the data set to cover the last observations requires that definitional changes has to be dealt with. From the second quarter of 1996 the data are published according to the new Standard of Industrial Classification (cf Statistics Norway (1994)). However for the aggregate manufactural sector these problems are not very serious since only minor changes have been undertaken.

Figure 1a. Preliminary estimates of investments and realized investments in the manufacturing sector[1]. Milliard Norwegian kroner



[1] Y_1 : Preliminary estimate of investments given in May in the year preceding the investment year
Y_2 : Preliminary estimate of investments given in August in the year preceding the investment year
Y_3 : Preliminary estimate of investments given in November in the year preceding the investment year
Y : Realized investments from February the year succeeding the investment year

Figure 1b. Preliminary estimates of investments and realized investments in the manufacturing sector[1]. Milliard Norwegian kroner



[1] Y_4 : Preliminary estimate of investments given in February in the year preceding the investment year
Y_5 : Preliminary estimate of investments given in May in the investment year
Y_6 : Preliminary estimate of investments given in August in the investment year
Y_7 : Preliminary estimate of investments given in November in the investment year
Y : Realized investments from February the year succeeding the investment year

3. Some forecasting models for final investments in the manufacturing sectors

METHOD A: Direct use of the investment surveys

If we generally adopt the convention that $\hat{Y}_t^N$ denotes the forecast of Y_t by method N, we get the following simple forecasting expressions:

$$(3.1) \quad \hat{Y}_{t|j}^A = Y_{j,t}, \; j = 1,\ldots,7.$$

The subscript of the variable at the left hand side indicates that the forecast of Y_t is not based on information later than at the time of the j measurement of investments in year t. Of course one cannot expect this method to do well under the «forecast competion» since no correction is made for systematic measurement errors which evidently are present in the data, but it serves as a natural reference for the other forecasting methods.

METHOD B: The «English method»

This method is not based on an explicit model and is therefore somewhat ad hoc. The main assumption is that the growth in final investment equals the growth in the preliminary measurements. This method utilizes only recent data and thus do not rely upon a stable relation lasting over time.

The forecasting equations may be written as:

$$(3.2a) \quad \hat{Y}_{t|j}^B = Y_{t-2} + \left(\frac{Y_{j,t} - Y_{j,t-2}}{Y_{j,t-2}} \right) Y_{t-2}, \; j = 1,\ldots,3$$

and

$$(3.2b) \quad \hat{Y}_{t|j}^B = Y_{t-1} + \left(\frac{Y_{j,t} - Y_{j,t-1}}{Y_{j,t-1}} \right) Y_{t-1}, \; j = 4,\ldots,7.$$

The reason that the equations are unequal for j=1,...,3 and j=4,...,7 is the following:

Let us take j=1 as the point of departure and that the period of observation for instance is the second quarter of 1994. In this quarter we receive the first measurement of the investments for 1995. The latest available observation of final investments is for 1993. Thus in order to use the «English method» one has to look at changes over two years. This is also valid for j=2 and j=3. The fourth measurement of the investments in 1995 is obtained in the first quarter of 1995, but now the final investments for 1994 have become known and one can therefore switch over to use changes over one year. The same is also valid when j=5, 6 and 7.

METHOD C: Ordinary least squares in levels (one regressor)

In a way this method is closely related to method A since it seeks to correct for the systematic measurement errors in the preliminary estimates of final investments. This method is based on regressions between final investments and preliminary estimates. The regression equations are given by:

$$(3.3) \quad Y_t = k_i + l_i Y_{i,t} + v_{i,t}; \; i = 1,\ldots,7; \; t = 1975,\ldots,S,$$

where k_i and l_i are unknown parameters and v_{it} is an error term. The relations contained in (3.3) are estimated by ordinary least squares. Final investments can now be forecasted by using the OLS-predictors of Y_t. Assume that we are in the second quarter of calendar year T which implies that $Y_{1,T+1}$ is observed. The last observed value of Y_t is Y_{T-1}, which became available in the preceding quarter. Thus (3.3) with i=1 may now be estimated by OLS by using data up to $(Y_{T-1},Y_{1,T-1})$. The predictor of Y_{T+1} is accordingly based on using these estimates together with $Y_{1,T+1}$. In the third and fourth quarter of calendar year T we are making analogous forecast of Y_{T+1}, but using Y_2 and Y_3 as regressors instead of Y_1 in the third and fourth quarter respectively. When we arrive at the first quarter of calendar year T+1, Y_T has become available and the coefficients k_4 and l_4 may therefore be estimated on data up to $(Y_T,Y_{4,T})$, but otherwise the procedure is analog to those utilized for the three preceding quarters. For the three remaining quarters of the calendar year T+1 the predictions of Y_{T+1} are based on using Y_5, Y_6 and Y_7 as regressors in the second, third and fourth quarter respectively.

METHOD D: Ordinary least squares in ratios (one regressor)
Forecast methods based on regressions on relative form have also been considered. As was the case for method C, method D is only utilizing one regressor. This variable is the ratio between the most recent observed variable and the same variable lagged one period. The following type of regressions are estimated by ordinary least squares.

(3.4)
$$\frac{Y_t}{Y_{t-1}} = rk_i + rl_i \frac{Y_{i,t}}{Y_{i,t-1}} + rv_{i,t}, i = 1,\ldots,7; t = 1976,\ldots,S.$$

In equation (3.4) rk_i and rl_i are unknown coefficients to be estimated, whereas $rv_{i,t}$ denotes an error term. In the presence of the lagged variable t now begin in 1976. Based on (3.4) we employ the following predictors for final investments

$$(3.5) \quad \hat{Y}^D_{t|j} = \left[\overline{rk_i} + \overline{rl_i}\frac{Y_{i,t}}{Y_{i,t-1}}\right] Y_{t-1}, j = 1,\ldots,7.$$

In equation (3.5) a bar denotes an OLS-estimate. In order to forecast Y_T we use data ending in period (T-2), when $j \le 3$, to estimate the coefficients in (3.4). When $j \ge 4$ we use data up to and included Y_{T-1} in order to estimate the coefficients. This is parallell to what is done under method C.

METHOD E: Linear model in levels with a time-varying constant term
This method can be viewed as a sophistication of the one presented under method C. The following model is now being assumed:

$$(3.6a) \quad Y_t = f_{i,t} + g_i Y_{i,t} + \eta_{i,t}; i = 1,\ldots,7; t = 1975,\ldots,S$$

and

$$(3.6b) \quad f_{i,t} = f_{i,t-1} + \kappa_{i,t}; i = 1,\ldots,7; t = 1975,\ldots,S.$$

The difference between model C and model E is that that we in the latter allows the constant terms in equation (3.3) to develop according to random walk processes. For the disturbances terms in (3.6a) and (3.6b) we now state the following assumptions:

$$(3.7a) \quad \eta_{i,t} = NIID(0,\sigma^2_{\eta_i}); i = 1,\ldots,7; t = 1975,\ldots,S,$$

$$(3.7b) \quad \kappa_{i,t} = NIID(0,\sigma^2_{\kappa_i}); i = 1,\ldots,7; t = 1975,\ldots,S$$

and

$$(3.7c) \quad E(\eta_{i,t}\kappa_{j,s}) = 0; i,j = 1,\ldots,7; t,s = 1975,\ldots,S.$$

The time-varying parameter model defined by (3.6a)-(3.6b) and (3.7a)-(3.7c) may be estimated by maximum likelihood utilizing the state space form and the prediction error decomposition. The computer programme STAMP 5.0 (cf. Koopman et al. (1995)) has been utilized for this purpose. Let us again assume that the period of time is the second quarter of calendar year T and that we are interested in forecasting Y_{T+1}. As for model C we utilize data for Y_t and $Y_{1,t}$ up to $(Y_{T-1}, Y_{1,T-1})$. From maximum likelihood estimation and component extraction we obtain an estimate of g_1 and an estimate of $f_{1,t}$ (for all $t \le T-1$). Since $f_{1,t}$ follows a random walk process the optimal forecast of f_{T+1} is the estimate of f_{T-1}. Thus the prediction equation for Y_{T+1} will depend on the estimate of g_1 and $f_{1,T-1}$ using data up to $(Y_{T-1}, Y_{1,T-1})$ and $Y_{1,T-1}$. Except that one have to take account of the random parameter feature the update of the forecast in the six next quarters is analogous to what was described for model C. If $\sigma^2_{\kappa_i} = 0$ for all i=1,...,7 we are back to model C again.

METHOD F: Ordinary least squares using levels (several regressors)
One characteristic of the forecasting methods A, B and C is that eventual additional information in the previous surveys are disregarded. Only information for the variable most recently observed is utilized.

In forecasting method F we instead utilize the «accumulative» information. This means that the number of regressors will increase by one from one quarter to the next. If we are in the second quarter of calendar year T and want to forecast Y_{T+1} this method degenerates to method C since only the first survey is available. However, when we arrive at the third quarter of the calendar year T realized investments are forecasted by regressing on both Y_1 and Y_2. The last observation point used is $(Y_{T-1}, Y_{1,T-1}, Y_{2,T-1})$. In the fourth quarter Y_3 is also added as a regressor. In the first quarter of calendar year T+1 we have four right hand side variables Y_1, Y_2, Y_3 and Y_4. Since Y_T now has become available the last complete observation point will be $(Y_T, Y_{1,T}, Y_{2,T}, Y_{3,T}, Y_{4,T})$. During the rest of the years the information from the remaining surveys will be added as soon as they become available. In the last quarter of the calendar year T+1 we will use all the seven surveys.

Formally we thus have the following regression equations:

(3.8)

$$Y_t = m_i + \sum_{j=1}^{i} n_{i,j} Y_{j,t} + \delta_{i,t};\ i = 1,\ldots,7;\ t = 1975,\ldots,S.$$

The symbols $\delta_{i,t}$ in (3.8) are error terms. In principle we could also in this context allow the constant terms to vary over time, but because of the low number of degrees of freedom we have not followed this route. It should also be mentioned that method C easily can be obtained from method F by introducing the appropriate exclusion restrictions in (3.8).

METHOD G: Ordinary least squares in ratios (several regressors)

Forecasting method G may be viewed as an augmentation of forecasting method D. In method D only one regressor was utilized in forecasting final investments. This variable was the one containing the most recent estimate of final investment. In forecasting method G the ratios from previous measurements are also included as regressors. Thus we are now estimating the following type of regressions by ordinary least squares

(3.9)

$$\frac{Y_t}{Y_{t-1}} = rm_i + \sum_{j=1}^{i} rn_{i,j} \frac{Y_{j,t}}{Y_{j,t-1}} + r\delta_{i,t}, i = 1,\ldots,7; t = 1975,\ldots,S.$$

In equation (3.9) the symbols rm_i and $rn_{i,j}$ are coefficients to be estimated and $r\delta_{i,t}$ is an error term. Final investments can now be forecasted by using the following type of predictors

$$(3.10)\quad \hat{Y}^{G}_{t|i} = \left[\overline{rm_i} + \sum_{j=1}^{i} \overline{rn_{i,j}} \frac{Y_{j,t}}{Y_{j,t-1}}\right] Y_{t-1}, i = 1,\ldots 7.$$

Again a bar denotes an OLS-estimate. Let us again assume that we want to predict Y_T. When i≤3, data up to period T-2 is utilized, whereas data up to T-1 is utilized when i≥4.

METHOD H: Combining "corrected" preliminary estimates using a diagonal weightmatrix

Method H is a simple method for combining preliminary estimates for the purpose of forecasting. In this case we employ a reverse regression (relative to method C) in which we have final investment on the right hand side of the equation. Thus we are looking at the following regressions:

(3.11)

$$Y_{i,t} = \theta_{0,i} + \theta_{1,i} Y_t + \xi_{i,t};\ i = 1,\ldots,7;\ t = 1975,\ldots,S,$$

where $\xi_{i,t}$ is an error term. Our aim is to forecast the value of Y_t utilizing the information in the already available preliminary measurements. The implementation of our procedure can be shown through a simple example. Assume we are in the third quarter of 1994. Two preliminary estimates of the final investments of 1995 are then available. This implies that we are looking at (3.11) when i=1,2. At the third quarter of 1994 the latest available value of Y_t is Y_{1993}. This was obtained in the first quarter of 1994. Thus when estimating (3.11) we only use data up to 1993. If we denote OLS estimates with an ∧, we may calculate preliminary estimates of final investments corrected for systematic measurement errors. In our example this is given by:

$$(3.12)\quad Yc^{93}_{i,t} = \left(\frac{Y_{i,t} - \hat{\theta}^{93}_{0,i}}{\hat{\theta}^{93}_{1,i}}\right);\ i = 1,2;\ t = 75,\ldots,95.$$

The top script 93 means that 1993 is the last year which is utilized in the OLS-regressions. Let the two latest received transformed observations be collected in a vector (which is written with an underscore):

$$(3.13)\quad \underline{Yc}^{93}_{95} = \begin{pmatrix} Yc^{93}_{1,95} \\ Yc^{93}_{2,95} \end{pmatrix}.$$

Analogous to the corrected variables defined by (3.12) we can also obtain the transformed residuals given by:

$$(3.14)\quad \xi c^{93}_{i,t} = \left(\frac{\hat{\xi}^{93}_{i,t}}{\hat{\theta}^{93}_{1,i}}\right);\ t = 75,\ldots,93;\ i = 1,2.$$

Let us furthermore define the following data matrix of transformed residuals:

$$(3.15)\quad \xi c2^{93} = \begin{bmatrix} \xi c^{93}_{1,75} & \xi c^{93}_{2,75} \\ \xi c^{93}_{1,76} & \xi c^{93}_{2,76} \\ \cdot & \cdot \\ \cdot & \cdot \\ \xi c^{93}_{1,93} & \xi c^{93}_{2,93} \end{bmatrix}$$

From (3.16) we may now calculate the matrix of the second order moments:

$$(3.16)\quad \Psi c2^{93} = (1/17)(\underline{\xi c2^{93}})'(\underline{\xi c2^{93}}).$$

We also need a further matrix which we label $\Psi c2D^{93}$. This has diagonal elements common with $\Psi c2^{93}$, but all the off-diagonal elements are zero. The value of Y_{95} can now be predicted using a minimum distance approach and a diagonal weightmatrix. The following quadratic form is minimized with regard to μ_{95} (where $\underline{\iota 2}$ is a 2×1-vector of ones):

$$(3.17) \quad (\underline{Yc2}_{93}^{95} - \underline{\iota 2}\mu_{95})'(\Psi c2D^{93})^{-1}(\underline{Yc2}_{93}^{95} - \underline{\iota 2}\mu_{95}).$$

The first order condition for minimization of (3.17) gives the following solution value for μ_{95} which again is taken as a forecast of Y_{95}:

(3.18)

$$\hat{Y}_{95|2}^{F} = \hat{\mu}_{95} = (\underline{\iota 2}'(\Psi c2D^{93})^{-1}\underline{\iota 2})^{-1}\underline{\iota 2}'(\Psi c2D^{93})^{-1}\underline{Yc2}_{93}^{95}.$$

The general problem, i.e. when one combine more than two preliminary measurements is not substantially different compared to the problem outlined above. When one has four or more preliminary estimates the coefficients and the transformed variables will be based on one more observation. Since the weighting matrix is diagonal it can be shown that the predictor is a linear combination of the corrected preliminary estimates in the weights are in the interval [0,1].

METHOD I: Combining corrected preliminary estimates using a full weightmatrix

This method works as method H, but with one exception. Instead of using the weighting matrix Ψc2D as in (3.17) and (3.18), one uses the full covariance matrix of the residuals as given by (3.16). In this case the predictor is again a linear combination of preliminary estimates, but in this case one may have weights that are negative or exceed unity.

METHOD J: State space model pooling information from preliminary measurement of final investments with information from an ARIMA(3,1,0) for final investments

In this section we will describe a forecasting model where we pool information from the preliminary estimates of investments together with information from an ARIMA-model for final investments. A building block in this model is the equations which are already given in connection with the presentation of method C. Several ARIMA-models for final investments were investigated and at the end an ARI(3,1) model was chosen. The two above mentioned information sources are pooled within a state space framework. Our state space model is inspired by among others Patterson (1994) which however works with a different situation since he is modelling a revision process for national accounts data. In addition we can also refer to Harvey et al. (1983) and the pages 337-341 in Harvey (1989). Our measurement equations for the state space form can be written as:

$$(3.19) \quad \underline{X}_{\tau} = V_{\tau}\underline{\alpha}_{\tau} + \underline{u}_{\tau}$$

In (3.19) the footscript τ has been used in order to emphasize that the observation vector X_{τ} is defined such as it has quarterly observations. As a convention we will always let $\tau=1$ mean the first quarter of a calendar year. The observation vector $\underline{X}_{\tau}$ is bivariate. How the $\underline{X}_{\tau}$ vector is defined can be shown by writing up the the four values of the vector during a calendar year. Assume for instance that $\tau=1$ correspond to the first quarter of 1994. Then the first four values of $\underline{X}_{\tau}$, namely $\underline{X}_1$, $\underline{X}_2$, $\underline{X}_3$ and $\underline{X}_4$ will be given by:

$$(3.20a) \quad \underline{X}_1 = [Y_{93} \quad Y_{4,94}]',$$

$$(3.20b) \quad \underline{X}_2 = [Y_{5,94} \quad Y_{1,95}]',$$

$$(3.20c) \quad \underline{X}_3 = [Y_{6,94} \quad Y_{2,95}]' \text{ and}$$

$$(3.20d) \quad \underline{X}_4 = [Y_{7,94} \quad Y_{3,95}]'.$$

The state vector $\underline{\alpha}_{\tau}$ is set up to contain 11-variables. It is possible under method H to write the model using fewer state variables, but it is convenient to view method H as a simplification of method I which will be commented on later. It seems natural to partion the vector $\underline{\alpha}_{\tau}$ in two subvectors:

$$(3.21) \quad \underline{\alpha}_{\tau} = \begin{bmatrix} \underline{\alpha 1}_{\tau} \\ \underline{\alpha 2}_{\tau} \end{bmatrix}.$$

These two subvectors are again defined as:

$$(3.22a) \quad \underline{\alpha 1}_{\tau}' = \begin{bmatrix} \alpha_{2,\tau} & \alpha_{1,\tau} & \alpha_{0,\tau} & \alpha_{-1,\tau} \end{bmatrix}$$

and

(3.22b)

$$\underline{\alpha 2}_{\tau}' = \begin{bmatrix} a_{1,\tau} & a_{2,\tau} & a_{3,\tau} & a_{4,\tau} & a_{5,\tau} & a_{6,\tau} & a_{7,\tau} \end{bmatrix}.$$

The state variables in (3.22a) are related to the ARIMA-part of the model, whereas the state variables in (3.22b) are related to the links between the preliminary measurements and final investments. The system matrix V_{τ} has a periodic structure. The equation (3.23a)-(3.23d) shows the specification of V_{τ} in the first, second, third and fourth quarter of a calendar year respectively:

$$(3.23a) \quad V_{\tau} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & g_4 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix},$$

$$(3.23b) \quad V_{\tau} = \begin{bmatrix} 0 & g_5 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ g_1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$(3.23c)\ V_\tau = \begin{bmatrix} 0 & g_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ g_2 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$(3.23d)\ V_\tau = \begin{bmatrix} 0 & g_7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ g_3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The measurement errors in two different time periods are assumed to be independent and normally distributed. Also the measurement errors within the same period is assumed to be independent. However the variances will not be constant, but will have a periodic structure according to the quarter. The bivariate measurement error vector in the first quarter of the calendar year will have a singular covariance matrix since the first element in the vector is measured without noise.

The transition equations in our state space model can be written as:

$$(3.24)\quad \underline{\alpha}_\tau = T_\tau \underline{\alpha}_{\tau-1} + R_\tau \underline{\varepsilon}_\tau.$$

Again both T_τ and R_τ have a periodic structure according to the quarter of the calendar year. The vector equation (3.24) can be partioned in the following way, if we again focus on the two types of informations being pooled:

(3.25)

$$\begin{pmatrix} \underline{\alpha 1}_\tau \\ \underline{\alpha 2}_\tau \end{pmatrix} = \begin{pmatrix} T_{11,\tau} & 0_{4\times 7} \\ 0_{7\times 4} & I_{7\times 7} \end{pmatrix} \begin{pmatrix} \underline{\alpha 1}_{\tau-1} \\ \underline{\alpha 2}_{\tau-1} \end{pmatrix} + \begin{pmatrix} \underline{R}_{11} & 0_{4\times 7} \\ 0_{7\times 1} & R_{22,\tau} \end{pmatrix} \begin{pmatrix} \varepsilon_{1\tau} \\ \underline{\varepsilon}_{2\tau} \end{pmatrix}.$$

In equation (3.25) $T_{11,\tau}$ is a 4×4 matrix which in the second, third and fourth quarter equals the identity matrix, whereas it in the first quarter is given by the following matrix:

$$(3.26)\quad T_{11,\tau} = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \phi_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The parameters ϕ_i are from the ARIMA-model for final investments. Further $\underline{R}_{11}$ is a 4×1 vector in which the top element is 1, and all other elements are zero. To state the periodic structure of the submatrix $R_{22,\tau}$ which is of dimension 7×7, it is convenient to define the following symbols: Let $\underline{0}$ denote a 7×1 zero-vector and let $\underline{e}_i$ be a 7×1 vector which has the property that its i'th element equals one and all the others zero (i runs from 1 to 7). With these symbols we are now ready to give the periodic structure of $R_{22,\tau}$ in the first, second, third and fourth quarter of a calendar year respectively:

$$(3.27a)\ R_{22,\tau} = \begin{bmatrix} \underline{0} & \underline{0} & \underline{0} & \underline{e}_4 & \underline{0} & \underline{0} & \underline{0} \end{bmatrix},$$

$$(3.27b)\ R_{22,\tau} = \begin{bmatrix} \underline{e}_1 & \underline{0} & \underline{0} & \underline{0} & \underline{e}_5 & \underline{0} & \underline{0} \end{bmatrix},$$

$$(3.27c)\ R_{22,\tau} = \begin{bmatrix} \underline{0}_1 & \underline{e}_2 & \underline{0} & \underline{0} & \underline{0} & \underline{e}_6 & \underline{0} \end{bmatrix}$$

and

$$(3.27d)\ R_{22,\tau} = \begin{bmatrix} \underline{0} & \underline{0} & \underline{e}_3 & \underline{0} & \underline{0} & \underline{0} & \underline{e}_7 \end{bmatrix}.$$

The eight dimensional disturbance vector $\underline{\varepsilon}_\tau$ is partioned into a scalar $\varepsilon_{1\tau}$ and a 7x1-vector $\underline{\varepsilon}_{2\tau}$. The disturbance vector $\underline{\varepsilon}_\tau$ is assumed to be stochastic independent and to have a diagonal covariance matrix. In method J we assume that $\underline{\alpha 2}_\tau$ is a constant and henceforth the covariance matrix of $\underline{\varepsilon}_\tau$ will be singular, i.e. all elements but the one in the upper left corner will be zero. This corresponds to the situation in which the preliminary estimates is linked to final investments through a linear relation with constant parameters. It should be emphasized that the state space model only is used for prediction of components. Estimates of the coefficients are obtained from the partial estimation of different blocks.

METHOD K: State space model pooling information from preliminary measurements with information from an ARIMA(3,1,0) model for final investments allowing for time varying constant terms

The difference between this method and method J is that none of the variances of $\underline{\varepsilon}_t$ are restricted to zero a priori. This method combines the elements of method E with an ARIMA-model for final investments.

To show how to utilize the state space specification for forecasting purposes we give an example. Assume that the point of time is the first quarter of 1995 (i.e. $\tau=1$), such that information dated later is still unavailable. In the first quarter of 1995 we have just obtained Y_{94} and we now estimate the ARI(3,1) model for this variable using data from 1975 to 1994. Likewise we estimate the linear relations between the seven preliminary measurements over the same sample assuming that the constant terms follow random walk processes. As a result of this we obtain estimated values of $f_{i,94}$, g_i, $\sigma_{\kappa i}^2$ and $\sigma_{\eta i}^2$ for all i=1,...,7. We specify the following estimate of the initial state vector $\alpha_0=(\alpha_{01}',\alpha_{02}')'$ as:

$$(3.28)\quad \underline{a}_{1,0} = \begin{pmatrix} \hat{Y}_{95} & Y_{94} & Y_{93} & Y_{92} \end{pmatrix}'$$

and

(3.29)

$$\underline{a}_{2,0} = \left(\tilde{f}_{1,94} \quad \tilde{f}_{2,94} \quad \tilde{f}_{3,94} \quad \tilde{f}_{4,94} \quad \tilde{f}_{5,94} \quad \tilde{f}_{6,94} \quad \tilde{f}_{7,94}\right)'.$$

Except for the first element in $\underline{a}_{1,0}$ we treat the entire initial vector as non-stochastic. The symbol $\hat{Y}_{95}$ denotes the forecast from the ARI(3,1) model for Y_{95}. If we let P_0 denote the estimate of the covariance matrix of the initial state vector, so will under our assumptions all the elements of it but the first element on the main diagonal be zero. At the upper left corner of the matrix we insert the estimated variance of the error term from the ARI(3,1) model. This ends the specification of the initial conditions. In the system matrices V_1, V_2, V_3 and V_4 we insert the estimated values of $g_1,...,g_7$ mentioned above. Likewise the estimates of the parameters $\sigma_{\eta 1}^2, ..., \sigma_{\eta 7}^2$ enter $E(\underline{u}_1\underline{u}_1')$, $E(\underline{u}_2\underline{u}_2')$, $E(\underline{u}_3\underline{u}_3')$ and $E(u_4\underline{u}_4')$. In the system matrix $T_{11,1}$ we insert the estimates of the coefficients in the systematic part of the ARI(3,1) model, i. e. the ϕ_i-parameters. The last matrix to be specified is the one which contains the variances of the disturbances in the transition equations. The first element along the main diagonal is the same elements as specified in the covariance matrix of the initial state vector. Thereafter follow the estimated values of $\sigma_{\kappa 1}^2, \ldots, \sigma_{\kappa 7}^2$.

We run the Kalman-filter (cf. for instance Harvey (1989, p. 100 and thereafter)) and the forecast for 1996 in the first, second, third and fourth quarter of 1995 are obtained as the Kalman-filter estimates of the first element of the state vectors $\underline{\alpha}_1$, $\underline{\alpha}_2$, $\underline{\alpha}_3$ and $\underline{\alpha}_4$ respectively. When we reach at the first quarter of 1996, Y_{95} has been available. We now reestimate the parameters entering the system-matrices, and we also insert a new initial state vector which is in accordance with these new estimates. We run the Kalman-filter for four new quarters and the forecasts of Y_{96} in the first, second, third and fourth quarter of 1996 are now obtained as the Kalman-filter estimate of the second element of the state vectors $\underline{\alpha}_5$, $\underline{\alpha}_6$, $\underline{\alpha}_7$ and $\underline{\alpha}_8$. Following a procedure like this we obtain 8 estimates of investments for an arbitrary year and the last seven of these are used when we are comparing the accuracy of forecasts done by different methods.

METHOD L: VAR-models of order 2

We also give forecasts based on VAR-models. Because of few observations we do not consider VAR-models which exceeds 3 in dimension. Actually, we are considering four VAR-models where we select different time series. Which VAR-model that is used depends on the quarter of the calendar year. Three of the VAR-models are of dimension three, whereas the fourth is of dimension 2. The general VAR-model, in error-correction form, is given as:

(3.30) $$\Delta\underline{Z}_t = \underline{\mu} + \Pi\underline{Z}_{t-1} + \sum_{i=1}^{l-1}\Gamma_i\Delta\underline{Z}_{t-i} + \underline{\varepsilon}_t.$$

In (3.30) the symbol Δ denotes the difference operator. The data vector $\underline{Z}_t$ is of dimension p. The parameter matrices Π, Γ_1, ..., Γ_{l-1} are all of dimension $p \times p$. The symbol $\underline{\mu}$ denotes a vector of constant terms of dimension $p \times 1$. Finally $\underline{\varepsilon}_t$ is supposed to be a p-dimensional innovation vector which follows the following distribution:

(3.31) $$\underline{\varepsilon}_t = \text{NIID}(\underline{0}_p, \Omega_{p\times p}).$$

In this section we assume that the VAR-models are of order 2, i.e. l=2. In the VAR-model used in the second quarter of the calendar year we are setting $\underline{Z}_t=(Y_{1,t+1}, Y_{5,t}, Y_{t-1})'$. In the second quarter we have just obtained the first measure of investment for year t+1, the fifth measure for investment in the current year, which is t. From the first quarter of the same year one has the final observed value for investment in year t-1. We estimate the VAR-model under the condition that the matrix Π is of reduced rank, namely that it has rank 2. This corresponds to assume the existence of one common trend, which seems natural in our context since the variables virtually are measuring the same phenomenon. Our objective is to forecast Y_t and Y_{t+1}. This involves one period ahead and two period ahead forecasting. When we arrive at the third quarter of a calendar year t we switch to a VAR-model in which the observation vector is given as $\underline{Z}_t=(Y_{2,t+1}, Y_{6,t}, Y_{t-1})'$. The comments with regard to estimation and forecasting are as for the first VAR-model. When we reach the fourth quarter of year t we utilize our third trivariate VAR-model. The observation vector is now given as $\underline{Z}_t=(Y_{3,t+1}, Y_{7,t}, Y_{t-1})'$. Again with regard to estimation and forecasting we refer to the comments given for the model used in the second quarter. The bivariate VAR-model is used when we are in the first quarter of the calendar year. We have then just obtained the final value for investment in the preceeding year together with the fourth estimate of investments for the investment in the current year. Thus if we stick to the earlier convention and define t as the current year we now get the following bivariate observation vector $\underline{Z}_t=(Y_{4,t}, Y_{t-1})'$. For this model we concentrate on the one step-ahead prediction, since we yet have no indicators of the final investments for period t+1. Both for estimation and forecasting we are applying the computer programme PcFIML 8.1 (cf Doornik and Hendry (1994)). The reduced rank VAR-models are estimated by utilizing the Johansen method (cf Johansen(1995)).

There is an interesting connection between method L and method C. When the long-run matrix Π has reduced rank equal to p-1, we may write Π as the product of two matrices α and β' which are of dimensions pxp-1 and p-1xp respectively. Thus we may now write:

$$(3.32)\quad \Delta \underline{Z}_t = \underline{\mu} + \alpha\beta'\underline{Z}_{t-1} + \sum_{i=1}^{l-1} \Gamma_i \Delta \underline{Z}_{t-i} + \underline{\varepsilon}_t .$$

It is well known that α and β are not uniquely identfied such that some (arbitrary) normalization rule must be imposed. A convenient normalization for us is to parameterize β, when Z_t is of dimension 3, as:

$$(3.33)\quad \beta = \begin{bmatrix} \beta_{11} & 0 & 1 \\ 0 & \beta_{22} & 1 \end{bmatrix}.$$

This gives, if we for instance look at the VAR-model used in the 2. quarter the following stationary combinations:

$$(3.34a)\quad Y_{t-2} + \beta_{11} Y_{1,t}$$

and

$$(3.34b)\quad Y_{t-2} + \beta_{22} Y_{5,t-1}.$$

Since different lags in the levels are of no importance in the long run, we reach at the type of equations used for method C by equalizing the terms in (3.34a) and (3.34b) to zero after having subtracted the expectation of the stationary combinations, which are given by formula (4.9) in Johansen (1991) . Thus method C can be interpreted as only using the long-run information contained in the method L.

Method M: VAR-models of order 1

This model is a simplification of the one presented for method L above. The difference is that the matrix Γ_1 a priori is set to zero in all the four VAR-models that is used within a year. The main reason for looking at this model is that it requires fewer parameters to be estimated. This seems important given the relative small number of observations.

4. A comparison of forecasts made by different methods

In order to compare forecasts using different methods we utilize two measures of predictive accuracy which are similar to RMSE (Root Mean Square Error) and MAPE (Mean Square Prediction Error) but which are defined on relative deviations. These two measures we refer to as RMSPE (Root Mean Square Percentage Error) and MAPPE (Mean Average Percentage Prediction Error) respectively. Let $\hat{Y}^p_{t|j}$ be the forecast of Y_t at stage j using method p. If we let t=1 indicate the first year we forecast final investments and t=T the last one, RMSPE and MAPPE at stage j using forecasting method p is defined as

$$(4.1) \qquad RMSPE^p_j = 100\sqrt{\frac{1}{T}\sum_{t=1}^{T}(\frac{\hat{Y}^p_{t|j}-Y_t}{Y_t})^2}, j=1,\ldots,7$$

and

$$(4.2) \qquad MAPPE^p_j = 100\frac{1}{T}\sum_{t=1}^{T}\left|\frac{\hat{Y}^p_{t|j}-Y_t}{Y_t}\right|, j=1,\ldots,7.$$

Note that all forecasts are based only on information which is available at the time of forecasting. Besides the measures defined in (4.1) and (4.2) we also define the pooled measures of predicitive accuracy which takes into account forecasts at all seven stages. These pooled measures are given in equations (4.3) and (4.4) below

$$(4.3) \qquad RMSPE^p_{1-7} = 100\sqrt{\frac{1}{7T}\sum_{t=1}^{T}\sum_{j=1}^{7}(\frac{\hat{Y}^p_{t|j}-Y_t}{Y_t})^2}$$

and

$$(4.4) \qquad MAPPE^p_{1-7} = 100\frac{1}{7T}\sum_{t=1}^{T}\sum_{j=1}^{7}\left|\frac{\hat{Y}^p_{t|j}-Y_t}{Y_t}\right|.$$

In our empirical investigation t=1 represents 1991 and T=5 represents 1995. Table 1 contains a survey of the forecasting methods employed in this analysis and desribed in section 2. In table 2 the main focus is on the forecasts made in the year before the investment year. The six first columns with figures are connected to the forecasts done at the first three stages and the calculations have been carried through according to (4.1) and (4.2) for j=1, 2 and 3. The last two columns contain the pooled measures of predicitive accuracy. To calculate the figures occuring there we have applied the formulas given in formulas (4.3) and (4.4). In table 2 the different forecasting methods have been ranked according to their predicitive accuracy. The model with the lowest value of RMSPE or MAPPE at a stage has been given the highest rank, i.e. 1.

The ranking of the different methods does not seem to depend much on the choice of forecasting comparison criteria. At the first two stages all forecasting methods improve upon the pure forecasts from the investment surveys which is contained in method A, whereas at the third stage the forecasts based on method A is only outperformed by two of the methods according to the RMSPE measure and by three methods according to the MAPPE measure. For the forecasts made at the two first stages we find that methods D and G give the best predictions since they have the lowest value of RMSPE and MAPPE. Method D and G, which coincides at the first stage, has an RMSPE which is only half of the RMSPE associated with method A. The third best method at the first stage is M, which is based on using VAR-models of order 1. Method D and G which coincides at the first stage are ranked four.

Method E, which is based on linear relations in levels with allowance for time-varying constant terms, is the best method at the third stage according to both measures. The only other method which outperforms method A according to the the RMSPE criterion is method M, which is based on using four VAR-models of order 1. According to the MAPPE criterion method B, the socalled «English method» is also better than method A. Method C, which is the forecasting method based on linear regression (in levels), has the lowest RMSPE and MAPPE when pooling forecasts at all stages. A general feature of table 2 is that the simpler models seem to perform rather well compared to the more complicated ones. The method in which one pool an ARIMA-model for final investments with the information contained in the preliminary estimates for final investments and besides allowing for time-varying constant terms, i.e. method K, have a low rank at all the stages. The forecasting performance improves somewhat when one goes from method K to method J, i.e. when one imposes the constant parameter restrictions. At the first stage method J is ranked fourth according to MAPPE. 1. At the first three stages, method M has a rank about three. The method is also ranked three according to the pooled measures of predicitive accuracy. Methods H and I performs rather poor and are never ranked higher than 7th and 5th, respectively, in table 2.

For the forecasts made in the investment year itself we have not been able to improve upon the pure forecasts from method A. Methods D and G which performed rather well for forecasts made in the year before the investment year, are not so good within the investment year. Closest to method A in predicitive accuracy are method C and the two methods based on VAR-modelling, i. e. method L and M.

Table 1. A survey of the different forecasting methods

Method	
A	Direct use of the investment survey.
B	The «English method».
C	Ordinary regressions using levels. Only the variable for the most recent measurement is utilized as regressor .
D	Ordinary regression using relative ratios. Only the variable for the most recent measurement is utilized as regressor.
E	Linear model in levels with a timevarying constant term. Only the variable for the most recent measurement is utilized as a regressor.
F	Ordinary regressions using levels. Both the variable for the most recent measurement and the variable corresponding to earlier obtained measurement are utilized as regressors.
G	Ordinary regressions using relative ratios. Both the variable corresponding to the most recent measurement and the variables corresponding to earlier obtained measurements are utilized as regressors.
H	Combination of «corrected» preliminary estimates using a diagonal weight matrix. The weights are based on regression.
I	Combination of «corrected» preliminary estimates using a full weight matrix. The weights are based on regression.
J	State space model pooling information from preliminary measurements of final investments with information from an ARIMA(3,1,0) model for final investments.
K	As method J, but with some allowance for time-varying parameters.
L	Forecasts from different VAR-models of order 2.
M	Forecasts from different VAR-models of order 1.

Table 2. Predictive accuracy of different forecasting methods according to RMSPE[1] and MAPPE[1]. Ranks in parentheses

Forecast method[2]	Point of time for forecast: 2. quarter in year t-1		3. quarter in year t-1		4. quarter in year t-1		All (7) stages	
	RMSE	MAPPE	RMSE	MAPPE	RMSE	MAPPE	RMSE	MAPPE
A	27.716 (13)	25.460 (13)	24.347 (13)	22.139 (12)	10.039 (3)	8.169 (4)	15.090 (12)	10.431 (7)
B	23.058 (10)	21.065 (10)	17.199 (4)	14.983 (3)	10.387 (5)	7.985 (2)	13.597 (6)	10.178 (6)
C	20.302 (4)	19.033 (5)	17.249 (5)	16.761 (6)	10.700 (7)	8.191 (5)	12.178 (4)	9.170 (4)
D	13.085 (1)	11.411 (1)	11.496 (1)	10.104 (1)	10.261 (4)	9.262 (7)	10.848 (2)	8.751 (1)
E	25.277 (12)	25.037 (12)	19.881 (10)	19.396 (10)	9.212 (1)	6.597 (1)	14.188 (9)	10.956 (10)
F[3]	20.302 (4)	19.033 (5)	19.724 (9)	18.833 (9)	14.681 (9)	12.600 (9)	13.762 (8)	10.834 (9)
G[3]	13.085 (1)	11.411 (1)	12.364 (2)	10.921 (2)	10.530 (6)	9.782 (8)	10.571 (1)	8.980 (2)
H	21.139 (8)	19.630 (8)	17.810 (7)	16.857 (7)	19.262 (13)	16.877 (13)	14.906 (11)	11.989 (12)
I[3]	21.139 (8)	19.630 (8)	20.591 (11)	19.449 (11)	15.512 (10)	12.831 (10)	14.360 (10)	11.225 (11)
J	20.908 (7)	18.880 (4)	19.426 (8)	18.312 (8)	15.779 (11)	14.645 (11)	13.739 (7)	10.791 (8)
K	24.592 (11)	23.986 (11)	23.166 (12)	22.738 (13)	17.111 (12)	16.709 (12)	15.708 (13)	12.766 (13)
L	20.488 (6)	19.239 (7)	17.587 (6)	16.571 (5)	11.546 (9)	8.358 (6)	12.282 (5)	9.433 (5)
M	19.438 (3)	18.020 (3)	16.261 (3)	15.861 (4)	9.526 (2)	8.018 (3)	11.690 (3)	8.989 (3)

[1] For a definition of these measures of predicitive accuracy cf equations (4.1)-(4.4).
[2] Table 1 gives a survey of these methods.
[3] At the first stage, i.e. at the second quarter of year t-1, method F degenerates to C, G to D and I to H.

5. Conclusions and further possibilities

At the first two stages, i. e. for forecasts made at the second and third quarter in the year preceding the investment year, a simple linear regression (one explanatory variable) model using relative ratios was found to perform best in the sense of giving the lowest value for RMSPE and MAPPE. For the forecasts made at the fourth quarter in the year before the investment year a simple linear model in levels (one regressor) which allowed the constant term to follow a random walk process was found to be preferable. For forecasts made in the investment years itself we were unable to outperform the pure forecasts produced by the surveys. A general conclusion seems to be that simple methods are doing rather well compared to more complicated methods. An natural interpretation of this is that the methods which requires many unknown parameters to be estimated are to complicated given the relative scarcity of data.

An important issue, which deliberately has been neglected in this report, is the question of forecast uncertainty. Other things equal a forecast method with a low variance of the forecast error is preferable. In order to look deeper into this a more explicit treatment of the distribution of the error terms are necessary. Likewise the order of integration of the variables is crucial. If the involved time series follow nonstationary processes, which for instance is the explicit position taken when using method L and M, the variance of the forecast error will grow fast with the number of forecast steps. When calculating the variance of the forecast errors it is possible to take both a partial view, concentrating solely on the contribution from the noise, or one may take a broader approach in which coefficient uncertainty also is taken into account.

It may be valueable to consider time series techniques which can combine data of different periodicities. In this report we have only considered data which cover a whole calendar year. This is the case both for realized investments and for the preliminary estimates of investment. The firms always report information about investment already undertaken in the calendar year, an estimate of investment for the remaining of the year and an estimate of investment for the next whole calendar year. It is well known that information from data given at different frequencies can be combined within a state space framework. For a recent contribution see Shen (1996).

This paper applies data in nominal terms. If one for instance want to utilize survey information in a large scale macroeconometric model with structural real investment equations a transformation of the survey data to fixed prices seems necessary. However such an analysis has been beyond the scope of this paper.

The fact that we are only using data from the investment survey does not mean that we assert that extraneous information is uninformative in view of forecasting investment, but rather that our modelling exercises aim at finding a benchmark (sub)model which may serve as a point of departure for further analysis. Besides, an obvious advantage of all the methods presented in this report is that it is rather easy to update the forecasts when new information becomes available. An analysis in real terms and with an extended data set has been done by Hagelund (1985).

In this paper we have only considered the aggregate manufacturing sector. It is also possible to conduct similiar analysis for different manufacturing sectors. Moreover we have also only considered total investment whereas data is available for different types of capital equipment. In the data section it was mentioned that the new Standard for Industrial Classification implicates a structural break in the time series. This feature will be much more significant using a disaggregated approach.

References

Doornik, J. A. and D. F. Hendry (1994): *PcFiml 8.0. Interactive Econometric Modelling of Dynamic Systems*, International Thompson Publishing.

Hagelund, K. (1985): Investeringstelingen som grunnlag for investeringsanslag, Arbeidsnotater nr. 11, Finansdepartementet (In Norwegian).

Harvey, A. C. (1989): *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press.

Harvey, A.C., C.R. Mc Kenzie, D.P.C. Blake and M.J. Desai (1983): «Irregular Data Revisions» in A. Zellner (ed.): *Applied Time Series Analysis of Economic Data*, U.S. Depratment of Commerce and Bureau of the Census, 329-343.

Johansen, S. (1991): Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Models, *Econometrica* **59**, 1551-1580.

Johansen, S. (1995): *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.

Koopmans, S. J., A. C. Harvey, J. A. Doornik and N. Shephard (1995): *STAMP 5.0. Structural Time Series Analyser, Modeller and Predictor*, Chapman & Hall.

Patterson, K. D. (1994): A state space model for reducing the uncertainity associated with preliminary vintages of data with an appplication to aggregate consumption, *Economics Letters* **46**, 215-222.

Shen, C. H. (1996): Forecasting macroeconomic variables using data of different periodicities, *International Journal of Forecasting* **12**, 269-282.

Statistics Norway (1994): *Standard Industrial Classification*, NOS C 182.

Statistics Norway (1996): *Ukens statistikk* 11/96 (In Norwegian).

Appendix A

Preliminary estimates of investments and realized investment in the manufacturing sector

Table 3. Preliminary estimates of investments and realized investments in the manufacturing sectors. Mill. Norwegian kroner

Investment year	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y
1975	3695	4611	6002	5841	5957	5809	5647	5463
1976	4376	5110	6130	6784	6394	6354	5904	5657
1977	4761	5456	6295	7281	7673	7671	7344	7373
1978	4310	5046	5992	6816	6930	6999	6816	6721
1979	4240	4472	5057	5840	5839	5877	5639	5478
1980	4914	5537	6705	7686	7421	7530	7388	6986
1981	5630	6136	6847	7711	7900	8248	8187	8406
1982	5043	5649	6505	7348	7174	7510	7503	7491
1983	4422	4904	5330	6347	6350	6420	6311	6225
1984	4323	4415	5798	7017	7007	7216	7163	7423
1985	5858	6476	7544	9501	8943	9580	9468	9489
1986	9731	9903	11986	12750	12840	13083	13057	13421
1987	10769	11173	13351	13567	13771	14957	16032	16198
1988	8967	9553	12490	12982	13915	14823	14743	14483
1989	7330	8497	9297	10571	11193	11060	11095	10856
1990	6881	7606	9528	10676	10708	10665	10500	10349
1991	8481	8914	10580	10874	11709	11584	11056	10471
1992	8974	8925	9451	9232	10331	10130	9800	10607
1993	8203	8639	9890	10541	9623	9960	9991	9751
1994	6270	6566	8180	9270	9537	9778	9696	9649
1995	7949	8688	12026	13295	13475	13716	13843	13706
1996	8542	9373	12355	15196				

Source: Statistics Norway (1996)

De sist utgitte publikasjonene i serien Rapporter

Recent publications in the series Reports

96/2 A. Essilfie: Investeringer, kostnader og gebyrer i den kommunale avløpssektoren: Resultater fra undersøkelsen i 1995. 1996. 36s. 80 kr. ISBN 82-537-4239-8

96/3 Resultatkontroll jordbruk 1996: Gjennom-føring av tiltak mot forurensninger. 1996. 85s. 95 kr. ISBN 82-537-4244-4

96/4 Å. Osmunddalen og T. Kalve: Bofaste innvandreres bruk av sosialhjelp 1987-1993. 1996. 33s. 80 kr. ISBN 82-537-4245-2

96/5 S. Blom: Inn i samfunnet? Flyktningkull i arbeid, utdanning og på sosialhjelp. 1996. 84s. 95 kr. ISBN 82-537-4249-5

96/6 J.E. Finnvold: Kommunale helsetilbud: Organisering, ulikhet og kontinuitet. 1996. 70s. 95 kr. ISBN 82-537-4221-5

96/7 Offentlig sektor i Norge: Strukturelle hovedtrekk og utvikling i perioden 1988-1994. 1996. 43s. 80 kr. ISBN 82-537-4268-1

96/8 K.E. Rosendahl: Helseeffekter av luft-forurensning og virkninger på økonomisk aktivitet: Generelle relasjoner med anvendelse på Oslo. 1996. 40s. 80 kr. ISBN 82-537-4277-0

96/9 S.-E. Mamelund og J.-K. Borgan: Kohort- og periodedødelighet i Norge 1846-1994. 1996. 236s. 165 kr. ISBN 82-537-4278-9

96/10 A. Schjalm: Kvalitetsundersøkelsen for Folke- og boligtelling 1990. 1996. 36s. 80 kr. ISBN 82-537-4279-7

96/11 K. Skrede og M. Ryen: Levekår i støpeskjeen. Status og utvikling i ungdomsgenerasjonenes materielle levekår 1990-1995. 1996. 80s. 95 kr. ISBN 82-537-4284-3

96/12 K.H. Alfsen, P. Boug and D. Kolsrud: Energy Demand, Carbon Emissions and Acid Rain: Consequences of a Changing Western Europe. 1996. 26s. 80 kr. ISBN 82-537-4285-1

96/13 M.W. Arneberg: Theory and Practice in the World Bank and IMF Economic Policy Models: Case study Mozambique. 1996. 28s. 80 kr. ISBN 82-537-4296-7

96/14 O. Skorge, F. Foyn og G. Frengen: Forsknings- og utviklingsvirksomhet i norsk industri 1993. 1996. 57s. 95 kr. ISBN 82-537-4306-8

96/15 K.O. Oftedal: Framskriving av markeds-situasjonen for helse- og sosialpersonell fram mot år 2030. 1996. 66s. 95 kr. ISBN 82-537-4307-6

96/16 M.I. Hansen, T.A. Johnsen og J.Ø. Oftedal: Det norske kraftmarkedet til år 2020: Nasjonale og regionale fremskrivninger. 1996. 39s. 80 kr. ISBN 82-537-4316-5

96/17 K. Flugsrud og K. Rypdal: Utslipp til luft fra innenriks sjøfart, fiske og annen sjøtrafikk mellom norske havner. 1996. 52s. 95 kr. ISBN 82-537-4321-1

96/18 T. Fæhn og T. Hægeland: Effektive satser for næringsstøtte 1994. 1996. 79s. 95 kr. ISBN 82-537-4323-8

96/19 A. Bråten og L. Sandberg: Priser på jordbruks-varer: En analyse av statistiske kilder. 1996. 84s. 95 kr. ISBN 82-537-4325-4

96/20 E. Gulløy, S. Gåsemyr og A. Vedø: Forslag til et nytt system for norsk bistandsstatistikk. 1996. 50s. 95 kr. ISBN 82-537-4338-6

96/21 A. Thomassen og T. Tørstad: Prisstatistikk for næringseiendommer: Prøveundersøkelse for Oslo og Akershus. 1996. 31s. 80 kr. ISBN 82-537-4340-8

96/22 A.K. Essilfie: Investeringer, kostnader og gebyrer i den kommunale avløpssektoren: Resultater fra undersøkelsen i 1995. 1996. 44s. 80 kr. ISBN 82-537-4344-0

96/23 S. Glomsrød, A.C. Hansen og K.E. Rosendahl: Integrering av miljøkostnader i makroøkono-miske modeller. 1996. 46s. 95 kr. ISBN 82-537-4348-3

97/2 T. Eika og K.-G. Lindquist: Konjunktur-impulser fra utlandet. 1997. 28s. 80 kr. ISBN 82-537-4357-2

97/3 T. Skjerpen and A.R. Swensen: Forecasting Manufacturing Investment Using Survey Information. 1997. 23s. 80 kr. ISBN 82-537-4374-2

B

Returadresse:
Statistisk sentralbyrå
Postboks 8131 Dep.
N-0033 Oslo

Publikasjonen kan bestilles fra:

Statistisk sentralbyrå
Salg-og abonnementservice
Postboks 8131 Dep.
N-0033 Oslo

Telefon: 22 00 44 80
Telefaks: 22 86 49 76

eller:
Akademika – avdeling for
offentlige publikasjoner
Møllergt. 17
Postboks 8134 Dep.
N-0033 Oslo

Telefon: 22 11 67 70
Telefaks: 22 42 05 51

ISBN 82-537-4374-2
ISSN 0806-2056

Pris kr 80,00

Statistisk sentralbyrå
Statistics Norway

Falch Hurtigtrykk, Oslo