Expectation Formation with Coincident Indicators

Johannes Leitner
Karl-Franzens-University of Graz, Department of Statistics and Operations Research, Universitätsstraße 15/E3, 8010 Graz, Austria

Abstract

Judgmental forecasting of real time series is performed in complex information environments. The complexity is caused by the number, quality and temporal characteristics of the sources of information. In a laboratory experiment we analyze the impact of these three factors on the forecasting behavior of the participants. The subjects are presented coincident indicators with a constant lead period of zero. The results are compared to prior experimental settings including leading indicators (lead period one) and no indicators. We find that forecasts with coincident time series hardly differ from information sets without indicators, but differ significantly from the leading indicator settings. The subjects are able to identify the temporal structure of the additional information. We present a model for the explanation of the average forecasts of the subjects to support this result.

Key words: Expectation Formation, Experimental Economics, Heuristics, Judgemental Forecasting, Time Series.

1. Introduction

Decision-makers who forecast an economic variable are often faced with a large amount of additional information. Not only that several sources of information are available, but also their temporal dimensions add to the complexity of the forecasting task. Such time series can be classified as leading, coincident or lagging indicators. Leading indicators tend to change before the economic variable, for instance consumer confidence for future consumption levels. Examples for lagging indicators are unemployment and corporate profits for total economic activity.

Coincident indicators describe the current state of an economic variable. The National Bureau of Economic Research introduced composite indices of coincident and leading economic indicators for macroeconomic activity of the USA (Burns & Mitchell, 1946). The methods for aggregating time series to such indices have been discussed (Forni et al., 2001) and indices for other national economies (Rua & Nunes, 2005; Fukuda & Onodera, 2001) or sub-national levels (Megna & Xu, 2003) have been created.

Coincident indicators do not provide information appropriate for the improvement of ex ante forecasts. However, decision-makers can use these time series for ex post forecasts not necessarily of macroeconomic data: For instance forecasting a share price based on the past development of the share index.

In our laboratory experiment participants judgementally forecast a time series. The information set contains past realizations of the time series and coincident indicators. The number and the quality of the indicators is systematically varied in three versions of the experiment.

Despite the relevance of indicators for the decision-making of economic agents in reality, experimental studies on judgmental forecasting behavior have mostly been limited to one time series to be fore-

Email address: johannesleitner@gmx.de (Johannes Leitner).

casted. The presentation of additional information in an experimental forecasting task has only been studied by a few authors. In Lim & O'Connor (1996) subjects were given low-correlated and high-correlated causal information for the judgmental forecast of sales figures. Subjects were able to judge the reliability of the information and to improve their forecasting accuracy by using it. However, their use was inefficient. Similar results were also found in more recent studies by Sanders (1997) and Goodwin & Fildes (1999).

In the context of testing expectation hypotheses, Garner (1982) and Brennscheidt (1993) reported on experiments with more than a single time series. Both authors found biases in individual and group forecasts which caused the rejection of the rational expectations hypothesis. Becker et al. (2007) conducted an experiment in seven different versions. The subjects were provided with 0 to 4 leading indicators of different quality for the forecast of a time series. The authors present a simple heuristic which is better able to explain average forecasting behavior consistently better than the rational expectations hypothesis. It is also found that the efficiency of forecasts decreases with an increasing number of indicators.

In these studies it is explored how the number and the quality of sources of information affect decisions. The temporal aspect is not addressed. The paper at hand is concerned with closing this gap. We will test the effects of coincident indicators on the forecasting behavior of subjects in a laboratory experiment. The experimental design is based on the study by Becker et al. (2007) with leading indicators, i.e. time series with a constant lead of one period. In our experiment, subjects forecast the same time series and are given the same indicators but the lead period is eliminated. This design allows the direct comparison of forecasting situations with leading indicators to coincident indicators and situations without any additional information. The comparison will be carried out under two different focuses: Average forecasts and forecasting accuracy. Although most forecasts are made by individuals, many economically relevant forecasts are aggregated measurements, i.e. statistically combined individual forecasts. The Delphi method was one of the first conjectural forecasting techniques. Delphi is a multi round (several times repeated) forecasting procedure based on independent inputs of selected experts (Sackman, 1974). It was developed in the 1950's at the Rand corporation in order to forecast future developments in areas lacking theoretical background and/or sufficient data for the application of quantitative forecasting models. In such situations aggregated judgment provides valuable information. Other examples of such procedures comprise the annual economic surveys by Livingston and the publication of consensus forecasts of financial analysts or economists for future earnings of corporations or macroeconomic data. Consensus forecasts are popular as an information source and research shows that this popularity is justified. By combining individual forecasts accuracy can be increased (Clemen, 1989; Granger, 1989).

We are interested in explaining such average forecasts and we present the bounds & likelihood heuristic as a simple model for this purpose. This heuristic was developed for information environments without indicators (Becker & Leopold-Wildburger, 1996) and with leading indicators (Becker et al., 2007) and proved to model forecasts very well. The model will be extended for an application to coincident indicators and its ability to explain the average forecasts of the subjects will be tested.

Second, we analyze the effects of providing coincident information on the quality of the forecasts. Unsurprisingly the results by Becker et al. (2007) show that high correlated leading indicators improve the forecasting accuracy and low correlated indicators worsen it. The stable lead has a significant effect. However, it can be assumed that additional information has an influence on forecasts if subjects think they recognize a lead although none exists. It could be hypothesized that the accuracy decreases when subjects are confronted with too much information. We will answer this question by comparing our results with Becker et al. (2007).

2. The experiment

2.1. *The experimental design*

In the experiment, subjects make judgmental forecasts of a time series x_t. The time series x_t (hereafter called "base series") is a realization of the stochastic difference equation

$$x_t = x_{t-1} - int(\frac{1}{2} \cdot x_{t-2}) + u_t \qquad (1)$$

with $x_1 = 7, x_2 = 12$, the endogenous variable x_t and the white noise u_t. The uniformly distributed variable u_t represents the realizations drawn from a six-sided dice, i.e. $u_t \in \{1,2,3,4,5,6\}$ and $f_u(u) = \frac{1}{6}$. The integer function *int* ensures that all values of the base series and the indicators are integer. Two coincident indicators were generated as follows:

$$y_t = x_t + 15 + v_t \quad (2)$$

$$z_t = x_t + 15 + w_t \quad (3)$$

The random variables v_t and w_t are discretely, triangularly distributed, i.e. $v_t \in \{-3,-2,\ldots,3\}$ with $f_v(v) = \frac{1}{41}(11 - 3|v_t|)$ and $w_t \in \{-5,-4,\ldots,5\}$ with $f_w(w) = \frac{1}{36}(6 - |w_t|)$ for all v_t and w_t respectively. This characteristic of the random variables makes y_t more reliable. The correlation of y_t/z_t and x_t is 0.85/0.67. Without the white noise, the coincident indicators would equal the base series plus 15 units. All time series are stationary. Figure 1 shows these three time series.

The definitions (1) to (3) yield that the coincident indicators are the base series with additional random components. They are generated by adding noise to the base series and cannot improve forecasting accuracy. With these characteristics the time series y_t and z_t of our experiment conform to the definition of coincident indicators.

The subjects were told in the instructions that the indicators might give them a certain insight into the change of the base series but no contextual information was given to the subjects and the time series were unlabelled in order to avoid a priori biases of forecasts (Chapman & Chapman, 1969). All subjects were presented the same realizations drawn from (1), (2) and (3).

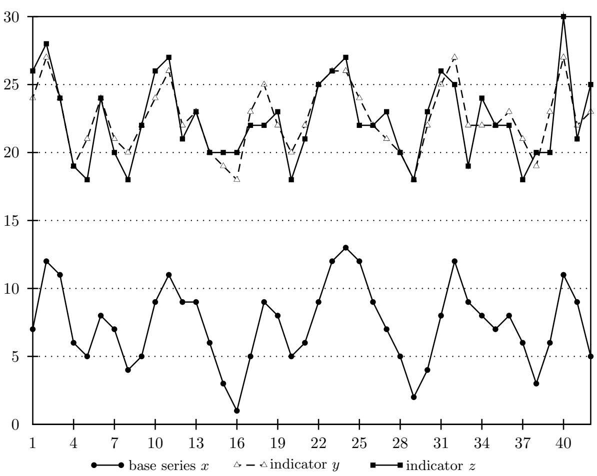

Figure 1. The base series and the indicators

The experiment was carried out with computers. After a forecast for period t was entered, the true realization of x_t was graphed first and each value of the coincindent indicators for period t was graphed in a fixed sequence afterwards. The most recent values of the base series and the indicators were also presented as figures. The time series were graphed in different colors. The subjects were allowed to proceed at their own pace but graphing the time series on the chart was performed slowly, which caused a minimum decision-making time of about 10 seconds per period. The subjects did not see a history of past values at period one of the experiment.

The number and the quality of the indicators was varied between three versions of the experiment.

Coincident 1 (C1): with x_{t+1} to be forecasted in period t and the coincident indicator y_t, i.e. information set $\Omega_t^1 = \{x_t, x_{t-1}, \ldots, x_1; , y_t, y_{t-1} \ldots, y_1\}$.

Coincident 2 (C2): with x_{t+1} to be forecasted in period t and the coincident indicator z_t, i.e. information set $\Omega_t^2 = \{x_t, x_{t-1}, \ldots, x_1; z_t, z_{t-1}, \ldots, z_1\}$.

Coincident 3 (C3): with x_{t+1} to be forecasted in period t and the coincident indicators y_t, z_t, i.e. information set $\Omega_t^3 = \{x_t, x_{t-1}, \ldots, x_1; y_t, y_{t-1}, \ldots, y_1; z_t, z_{t-1}, \ldots, z_1\}$.

The subjects were paid for their forecasts with the function $p_t = 20 \cdot max\{3 - |x_t - f_t|; 0\}$, i.e. in each period they received 60 cents for an exact prediction and 40 (20) cents for a deviation of one (two) unit(s). The average payment was 7 Euros for a duration of about 40 minutes. Altogether 70 subjects participated in the experiment: 30 in version C1, 22 in version C2 and 18 in version C3.

2.2. *Comparison to other versions*

The experimental setup described in subsection 2.1 is a modification of earlier experiments reported in Becker et al. (2007, in press). These experiments were conducted in several versions with leading indicators and with no indicators, i.e. only the base series. Four of these versions are of particular interest to us: A comparison of the three coincident information sets C1-C3 to the version without any indicators (version 0) will demonstrate the effect of coincident indicators on forecasting accuracy and forecasting behavior. Three of these experiments contained the time series y_t and z_t with a constant lead of 1. The information sets of these experiments with leading indicators correspond to the three coincident versions and therefore we call them L1, L2 and L3. In the table below the old datasets and the new experiments are summarized [1].

Table 1
Information sets and size of the total sample

Version	Subjects	Information Set
0	297	x_t
L1	150	x_t, y_{t+1}
L2	62	x_t, z_{t+1}
L3	30	x_t, y_{t+1}, z_{t+1}
C1	30	x_t, y_t
C2	22	x_t, z_t
C3	18	x_t, y_t, z_t

3. The bounds & likelihood heuristic

The bounds&likelihood heuristic (b&l heuristic) models average forecasts. The calculation is based on the same information that is available to the subjects, namely the time series' gestalt characteristics. Becker & Leopold-Wildburger (1996) presented the model for situations without indicators. In Becker et al. (2007) a description of an extended model for leading series can be found. The model suggests that subjects generate a predictor for each time series and combine these values linearly to a forecast of the base series. The base series predictor considers the absolute changes of x_t observed so far and the likelihood of x_t being a turning point determined by the relative position of x_t to the local extrema. All leading series in the experiment have a stable lead of one period, i.e. a certain change of each indicator is observed and has to be related to the expected change of the base series in the next period. This is performed by a factor that relates the average changes of the base series and the indicators. In the leading series case this model explains at least 90% of the variance of the subjects' average forecasts (Becker et al., 2007).
In our information sets including only coincident indicators and the base series all time series are presented simultaneously. Two questions arise: How do subjects generate predictors from each time series and how are these predictors combined? We assume that the subjects perform similarly to the versions L1 to L3. Since there is no lead, all time series are considered the same way as the base series. The combination of predictors is based on absolute prediction errors. We do not have a reason to believe that this works in the other versions but is done differently by the subjects in the coincident setup.
For the b&l heuristic it is assumed that two characteristics of the base series are essential for the forecasts: the average variation and the turning points. It is assumed that the average absolute variations of the time series $b_t^s = \frac{1}{t-1}\sum_{j=2}^{t}|s_j - s_{j-1}|$ are the bounds for the predicted change based on the actual time series value s_t (average variations and all other parameters are calculated not only for the base series, therefore we write e.g. b_t^s for $s = x, y$ and z). The subjects should on average predict changes which are in the interval [-b_t^s, b_t^s]. The actually predicted change depends on the likelihood that s_t is a turning point. For $s_t > s_{t-1}$, an upswing case, $l_{t(peak)}^s$ is the probability that s_t is a local maximum. The total number of local maxima observed so far (N_t^s) and the number of local maxima $\leq s_t$ (n_t^s) are considered. If all local maxima are below s_t, i.e. $n_t^s = N_t^s$, it is very likely to be a turning point. For a downswing case ($s_t < s_{t-1}$), the total number of local minima (M_t^s) and the number of local minima $\geq s_t$ (m_t^s) are considered for the calculation of $l_{t(trough)}^s$. This is shown in definition (4).

$$\begin{aligned} l_{t(peak)}^s &= \frac{1+n_t^s}{2+N_t^s} \\ l_{t(trough)}^s &= \frac{1+m_t^s}{2+M_t^s} \end{aligned} \quad \text{for s= x, y, z} \qquad (4)$$

If no extrema have occured so far, i.e. $m_t^s = M_t^s = 0$ and/or $n_t^s = N_t^s = 0$ the likelihoods of an upswing (downswing) are 0.5. In the case of no change ($s_t = s_{t-1}$) it is assumed that the upswing and downswing cases are combined linearly. At a high level of the time series, subjects will forecast a downswing; at a low level, an upswing.
In most cases the base series and the coincident indicators have different dimensions and average variations. In order to relate smaller (larger) average variations of the base series to an indicator we define

[1] The experiments in version 0 were conducted over several years and the experimental setup changed in two aspects: The first experiments were paper based and the participants were not paid. The most recent experiments were computerized and the subjects received rewards based on the function p_t. Statistical analyses, however, did not reveal any significant differences between the forecasts and therefore the sample is considered to be homogenous.

the relation factor r_t:

$$r_t^s = \frac{\sum_{i=1}^{t} |\Delta x_i|}{\sum_{i=1}^{t} |\Delta s_i|} \quad \text{for} \quad s = x, y, z \tag{5}$$

Based on these assumptions, the predictor of the series s_t, $f_{t+1,b\&l}^s$, is described by definition (6):

$$f_{t+1,b\&l}^s = \begin{cases} x_t + b_t^s \cdot (1 - 2l_{t(peak)}^s) \cdot r_t^s & \text{for } s_t > s_{t-1} \\ x_t + b_t^s \cdot (l_{t(trough)}^s - l_{t(peak)}^s) \cdot r_t^s & \text{for } s_t = s_{t-1} \\ x_t - b_t^s \cdot (1 - 2l_{t(trough)}^s) \cdot r_t^s & \text{for } s_t < s_{t-1} \end{cases}$$

In the experiment up to 3 predictors are available. If their values differ, they have to be combined to one forecast. The predictors are weighted with α_t reciprocally proportional to their absolute prediction error D_t for each time series, given in (7):

$$D_t^s = \sum_{i=2}^{t} |f_{t,b\&l}^s - x_i| \text{ for } s = x, y, z \tag{7}$$

The calculation of the weights is described in (8):

$$\alpha_t^s = \frac{\frac{1}{D_t^s}}{\frac{1}{D_t^x} + \frac{1}{D_t^y} + \frac{1}{D_t^z}} \qquad \text{for } s = x, y, z \tag{8}$$

Definition (8) implies that the relationships of the weighting factors are

$$\frac{\alpha_t^x}{\alpha_t^y} = \frac{D_t^y}{D_t^x}, \quad \frac{\alpha_t^x}{\alpha_t^z} = \frac{D_t^z}{D_t^x}, \text{ and } \quad \frac{\alpha_t^y}{\alpha_t^z} = \frac{D_t^z}{D_t^y} \tag{9}$$

Using these terms the combined forecast, $f_{t+1,b\&l}$, of the b&l heuristic is generated. The forecast values of the next period consist of a linear combination of all predictors:

$$f_{t+1,b\&l} = \alpha_t^x \cdot f_{t+1,b\&l}^x + \alpha_t^y \cdot f_{t+1,b\&l}^y + \alpha_t^z \cdot f_{t+1,b\&l}^z \tag{10}$$

The definitions were only presented for version 3 of the experiment but they are analogous if only one indicator is in the information set such as in versions 1 and 2 and it can be extended to an arbitrary number of coincident indicators. The values of (10) will be compared to the average forecasts of the subjects in the experiment in the next section.

4. Experimental results

Our main interest is whether the average forecasts of the subjects can be explained by the bounds & likelihood heuristic when coincident indicators are included in the information set. Based on the characteristics of the coincident indicators we assume that the model cannot explain the forecasting behavior better than the model of version 0, which considers the base series as the only information.

The coincident indicators in our experiment cannot improve forecasting performance. They are derived from the base series by adding random components(6). The forecasting performance of the subjects is supposed to be worse compared to version 0 (no indicators) and significantly worse than in the three versions with leading indicators (L1-L3). This effect is assumed to be largest in version C2 with the possibly most misleading coincidental indicator z_t. The forecasting accuracy of the subjects will be compared between the seven versions of the experiment. The first periods serve as phase for practice and familiarization. For this reason, periods 1-6 are excluded from the analysis. Only periods 7-42 are considered.

4.1. *Modeling average forecasts*

The bounds & likelihood heuristic models the average forecasts $f_{t,avg} = \sum_{i=1}^{N} f_t^i$ of N individuals. Before we test the ability of the heuristic to forecast the subjects we check the characteristics of the average forecasts in all versions of the experiment.

The forecasts in the three versions with coincident indicators (C1-C3) hardly differ from the average forecasts in the experiment without any indicators (version 0). The absolute deviations $\frac{1}{36}|f_{t,avg}^C - f_{t,avg}^0|$ of versions C1/C2/C3 from version 0 are 0.4/0.55/0.54. This is an important result. Independent of the number and the quality of coincidental indicators, average forecasting behavior is not significantly affected. This is even more obvious when the versions C1-C3 are compared to their corresponding leading series information sets L1-L3: The absolute average deviations are 1.42/1.36/1.61. Motivated by the similarity of average forecasts we also want to test for differences in the distributions of the forecasts. For this purpose Kolmogorov Smirnov tests are performed in each of the 36 periods pairwise, i.e. the distribution of forecasts in version 0 is tested for differences with C1, C2 and C3. The same tests are performed for comparisons with the leading indicator versions: C1 vs L1, C2 vs L2, C3 vs L3. In Table 2 the number of periods with significantly ($p < 0.1$) different distributions of forecasts are reported.

Table 2
Numbers of periods with different distributions of individual forecasts

C1 vs 0	C2 vs 0	C3 vs 0	C1 vs L1	C2 vs L2	C3 vs L3
4	3	2	26	25	27

The distributions of individual forecasts in version C1 (coincident indicator y_t) differ significantly from the distributions in version 0 in only 4 of 36 periods. A test for differences between the distributions of version C1 and L1 (leading indicator y) reveals that these differ in 26 periods. The results for the other versions are very similar. It can be concluded that neither average forecasts nor the distributions differ significantly between versions with coincident indicators and no indicators, but they differ from leading indicator versions.
As a consequence for the modeling of average forecasts it can be hypothesized that versions with coincident indicators will be explained well by the heuristic in version 0 although it does not include any of the additional time series. We test the performance of the heuristic with the estimation of simple linear time series regressions. In this model, the heuristic in version 0, C1, C2 and C3 are the independent variables and the average forecasts are the dependent variables:

$$f_{t,avg} = \alpha + \beta \cdot f_{t,b\&l} + \epsilon_t \tag{11}$$

The results of these regressions and the absolute average forecasting errors of the heuristic $\frac{1}{36}\sum_{t=7}^{42} |f_{t,b\&l} - f_{t,avg}|$ are presented in Table 3.

The b&l model in version 0 explains 92.4% of the variance of average forecasts in version C1 of the experiment. The modified heuristic including the coincident indicator y_t as described in section 3 explains 90%. All modified versions of the heuristic explain at least 87.1% (version C2) of the variance of average forecasts, which is a good result. However, the version 0 b&l model consistently outperforms the modified heuristics.
As was to be expected the incorporation of the coincident indicators worsens the performance of the heuristic. Coincident indicators cannot improve forecasting performance and by not including them in the information set, the behavior can be forecasted more accurately.

Table 3
Simple linear time series regressions

b&l	Experiment	α	β	R^2	MAE
0	0	1.325 (0.261)	0.891 (0.034)	0.952	0.642
0	C1	0.701 (0.364)	0.977 (0.048)	0.924	0.73
0	C2	0.221 (0.46)	1.038 (0.061)	0.896	0.838
0	C3	0.749 (0.352)	1.014 (0.047)	0.933	0.875
C1	C1	0.486 (0.429)	0.997 (0.056)	0.9	0.716
C2	C2	-0.015 (0.523)	1.06 (0.069)	0.871	0.836
C3	C3	0.454 (0.514)	1.034 (0.069)	0.872	0.894

Note: standard error in parentheses

4.2. *Forecasting accuracy*

We evaluate the forecasting accuracy of the subjects in terms of the mean error (ME), the mean absolute error (MAE), and Theil's U for a direct comparison of the subjects' and the naive model's mean squared errors. We calculate these error measures on the collective and individual level, for instance $\text{MAE}_{avg} = \frac{1}{36}\sum_{t=7}^{42} |f_{t,avg} - x_t|$ and $\text{MAE}_{ind} = \frac{1}{N}\frac{1}{36}\sum_{i=1}^{N}\sum_{t=7}^{42} |f_t^i - x_t|$. The results are summarized in Table 4.

Table 4
Forecasting accuracy of the subjects in versions 0-3 of the experiment

Version	ME	Average forecasts		Individual forecasts		MAE-relation (%)
		MAE	Theil's U	MAE	Theil's U	
0	0.431	1.765	0.738	2.444	1.153	0.722
C1	0.423	1.96	0.826	2.338	1.092	0.838
C2	0.383	2.069	0.851	2.451	1.106	0.844
C3	0.739	2.005	0.85	2.505	1.166	0.8
L1	0.06	1.031	0.469	1.398	0.687	0.737
L2	0.352	1.771	0.772	2.422	1.157	0.731
L3	0.217	1.393	0.64	1.739	0.824	0.801

A comparison of the error measurement categories reveals very small differences between the co-

incident indicator versions (C1-C3) and very large differences when leading indicators are in the information set (L1-L3). This holds for average forecasts and individuals and all error measurements except the ME that is higher in version C3. The variation of the quality of leading indicators obviously has a significant impact on forecasting accuracy, but the analogous variation of coincident indicators does not affect the accuracy of the forecasts.
This analysis supports the findings of the preceding subsection. It can be concluded that coincidental indicators hardly affect the forecasting behavior of the subjects independent of their number and their correlation to the base series.
The results of Table 4 also demonstrate the improvement of forecasting accuracy caused by the combination of individual forecasts with the arithmetic mean. The forecasting errors of the average forecasts are lower than those of the average individuals in all versions. In table 4 the relation of the $\text{MAE}_{avg}/\text{MAE}_{ind}$ indicates a reduction of the MAE of about 20%. The largest effect occurs in the largest sample of version 0 with 297 subjects. An increasing sample size does not monotically reduce this relation and a theoretically infinite sample size will not generate perfect forecasts. However, we can conclude that averaging the forecasts always effectively improves the forecasting accuracy and that larger samples increase this effect.

5. Conclusion

A forecasting experiment with coincident indicators was presented. The experiment was conducted in three versions with a varied number and quality of coincident indicators. The bounds & likelihood heuristic was modified to incorporate coincident information for the modelling of average forecasts. The results were compared to versions with leading indicators and a version with no indicators.
We found that the forecasts of the subjects hardly differ from the information set including no indicators but they differ significantly from experiments with leading indicators. This holds for the individual forecasts, their arithmetic means and their distributions. It is known from earlier experiments (Becker et al., 2007, in press) that the the bounds & likelihood heuristic explains more than 95% of the variance of average forecasts in versions with leading indicators. Including coincident information in the calculation of the heuristic did not improve the forecasting performance.
It can be concluded that subjects were able to identify the time structure of the time series and that their behavior can be explained best by not considering the coincident indicators at all.

Acknowledgements
The author thanks Otwin Becker, Ulrike Leopold-Wildburger, Ulrich Pferschy and Hans Kellerer for helpful comments on draft versions of this paper.

The financial support of the Austrian Science Foundation (FWF) for the project *"Heuristics for the Subjective Prediction of Time Series"* P17156-N12 is gratefully acknowledged.

References

Becker, O., Leopold-Wildburger, U., 1996. The bounds and likelihood-procedure - A Simulation Study Concerning the Efficiency of Visual Forecasting Techniques. *Central European Journal of Operations Research and Economics*, 4, 223-229.

Becker, O., Leitner, J., Leopold-Wildburger, U., 2007. Heuristic modeling of expectation formation in a complex experimental information environment. *European Journal of Operational Research*, 176, 975-985.

Becker, O., Leitner, J., Leopold-Wildburger, U., in press. Modeling Expectation Formation Involving Several Sources of Information. *German Economic Review.*

Bolger, F., Harvey, N., 1993. Context-sensitive Heuristics in Statistical Reasoning. *The Quarterly Journal of Experimental Psychology*, 46A, 779-811.

Brennscheidt, G., 1993. *Predictive Behavior - An Experimental Study.* Lecture Notes in Economics and Mathematical Systems Vol. 403.

Burns, A. F., Mitchell, W. C. (1946). *Measuring business cycles.* New York: NBER.

Chapman, L.J., Chapman, J.P., 1969. Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, 74, 271-280.

Clemen, R.T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5, 559-583.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2001. Coincident and leading indicators for the euro area. *The Economic Journal*, 111, C62-C85.

Fukuda, S., Onodera, T., 2001. A new composite index of coincident economic indicators in Japan: how can we improve forecast performances? *International Journal of Forecasting*, 17, 483-498.

Garner, A.C., 1982. Experimental Evidence on the Rationality of Intuitive Forecasters, In: V.L. Smith (Eds.), *Research in Experimental Economics*, Greenwich. Conn: JAI Press, pp. 113-128.

Granger, C.W.J. (1989). Combining forecasts - twenty years later. *Journal of Forecasting*, 8, 167-173.

Goodwin, P., Fildes, R., 1999. Judgmental Forecasts of Time Series Affected by Special Events: Does Providing a Statistical Forecast Improve Accuracy? *Journal of Behavioral Decision Making*, 12, 37-53.

Hey, J.D., 1994. Expectations Formation: Rational or Adaptive or...? *Journal of Economic Behavior and Organization*, 25, 329-349.

Lim, J.S., O'Connor, M., 1996. Judgmental forecasting with time series and causal information. *International Journal of Forecasting*, 12, 139-153.

Megna, R., Xu, Q., 2003. Forecasting the New York State economy: The coincident and leading indicators approach. *International Journal of Forecasting*, 19, 701-713.

Markidakis, S., Winkler, R.L., 1983. Averages of forecasts: some emprical results. *Management Science*, 29, 987-996.

Rua, A., Nunes, L.C., 2005. Coincident and leading indicators for the euro area: A frequency band approach. *International Journal of Forecasting*, 21, 503-523.

Sanders, N.R., 1997. The impact of task properties feedback on time series judgmental forecasting tasks. *Omega: International Journal of Management Science*, 25, 135-144.

Sackman, H. (1974). *Delphi Assessment, Expert Opinion, Forecasting, and Group Process.* Santa Monica: Rand Corporation, Rand report R-1283-PR.