The Success of Business Forecasting: Comparisons across Industries, Countries and Time

Tobias F. Rötheli[1]

[1] Department of Economics, University of Erfurt, Germany

Correspondence: Tobias F. Rötheli, Nordhäuser Strasse 63, PF 900 221, D-99105 Erfurt, Germany.

Abstract

This study assesses the accuracy of forecasts by industry branches. Such an investigation provides a perspective on the relative benefits of forecasting in different industries. Accuracy of forecasting is assessed by econometrically investigating expectations data on firms' production drawn from surveys covering manufacturing. Such data is available for only few countries and few historical periods. We study U.S. data covering the 1980s and German data over the period from 1991 to 2018. We first present rankings of industries according to forecast accuracy for both countries. Then the historical gap between the two countries' data set is put to use to assess the stability and the dynamics in the relevance of forecasting in different branches of industry. We identify several industries that – across time and place – are among the most (e.g., electric machinery) and least accurate forecasters (e.g., the food industry). By contrast in some industries forecasting performance appear to undergo noticeable changes over time: the reported evidence suggests that forecasting has lost some of its potential in the printing and textile industries while gaining over time in the nonelectric machinery and in the metals industry. The findings can help management to make decisions regarding the allocation of resources to forecasting.

1. Introduction

Business forecasting has a long history and a wide range of applications (Rötheli, 2007, Friedman, 2013). A broad literature covers available techniques and applications (Newbury, 1953, Armstrong, 2001, Jain, 2004, Elliott, Graham & Timmermann, 2016, Fildes et al. 2019). Concerning the actual use of forecasting techniques we have the benefit of findings from asking businesses what types of forecasting they actually practice (Newbury, 1953, Dalrymple, 1987, Peterson, 1993). What is less well researched is the question which branches of business most successfully engage in forecasting. From a normative viewpoint – considering what resources should be allocated to forecasting – this is an important issue. In particular, the firm's output is the key variable to forecast since this forecast determines the appropriate allocation of inputs into the production process. It is important to reflect on the fact that forecasting yields different returns in different branches of industry. More specifically, the benefits of business forecasting depend on (i) the extent to which the product mix in an industry varies over the business cycle (expansion versus contraction), (ii) the extent to which the optimal size of the operation depends on the state of the economy, and (iii) on the accuracy of the forecasts available (Rötheli, 2018).

Accordingly, business forecasting should be most prevalent in industries where firms substantially change the output mix in different stages of the business cycle, where the firm's optimal size of operation depends on the amplitude of the cycle and where, dollar for dollar, forecasting is most successful at sharpening foresight. As it turns out, it is very difficult to directly assess which branches of industry – according to these dimensions – benefit more or less from forecasting. Hence, the question at hand is reframed in this study: Here, we assess which industries are more or less successful at forecasting. Proceeding on the assumption of rational resource allocation, more accurate forecasting signals where forecasting has the strongest impact.[1,2] This then is the

---

[1] If resource allocation is affected by boundedly rational choices then differences in forecast accuracy are also affected by historical processes determining the status and competence of forecasters in different branches of industry.

central tenet of the approach presented here: forecast accuracy can be interpreted as an indicator of the relevance and importance of business forecasting in a branch of industry. This form of interpreting predictive accuracy is new in the forecast literature. It opens the door for best practice assessments of the sort that industries in countries not covered in this study may evaluate their forecasting practices. If, e.g., they find their industry high up in our rankings and acknowledge their own scant use of resources for forecasting, then improving this management tool seems advisable.

2. Methodology and Data

The question of forecasting accuracy is tackled here by studying survey data on firms' output forecasts. It has to be emphasized that such surveys are very rare. In fact for the U.S. there is no such survey currently run. The only such data for the U.S. comes from a survey of the Conference Board which ran a survey on CEOs' expectations for which we have data for the period from 1978 to 1988. This is a quarterly survey and asks respondents about expectations concerning business conditions in their own industry six months ahead and current conditions compared with six months ago. Accordingly, the data allow a match of forecasts and realizations. For one other major economy, Germany, a survey of industry firms is undertaken since the early 1990s. The German data come from the IFO institute and cover the period starting in 1991 up to 2018. The monthly IFO survey asks firms about expected output over the coming three months. The question concerning realizations asks for the change in production relative to the previous month. Accordingly, a match of horizons in the German case means that we convert the monthly data into quarterly observations and with respect to realizations we use the average change over a period of three months.

Before turning to the specifics of the data, consider how the performance of a forecaster is typically assessed. A widely used procedure to evaluate forecasting performance is to first compute the forecast error (i.e., the difference between the value actually realized from the forecast value) and then to assess a measure like the root mean square error (Armstrong & Collopy, 1992). This procedure is useful when comparing (ex post) different forecasting models or heuristics against each other. However, when the task is to evaluate forecast performance of different forecasters in different circumstances (e.g., different industries), the forecast accuracy (or the lack thereof) has to be set relative to the variability of the indicator to be forecast. For this purpose, the correlation of the forecast with the actual realization is a relevant measure. Specifically, we estimate a regression of the form

$$Q_t = \alpha + \beta Q_{t,t-1}^F + \varepsilon_t \qquad . \tag{1}$$

Here, Q_t stands for the actual value of the variable (e.g., ex-post output) and $Q_{t,t-1}^F$ stands for the forecast (i.e., the ex-ante value of output). This specification is also typical in studies that evaluate different models of expectations (Pesaran & Weale, 2006). The relevant statistic to assess for an evaluation then is the standard error of the estimate, or alternatively, the R^2 statistic of the estimate.[3] The approach set out in equation (1) is appropriate for the type of data analyzed here. Both the U.S. and the German data ask respondents for a categorical choice. In the U.S. case the Conference Board allowed firms to choose from five possible assessments: "substantially better", "moderately better", "the same", "moderately worse", and "substantially worse". Answers in the respective categories receive a weight of 100, 75, 50, 25, and 0, respectively. Averaging over respondents in an industry yields the data presented in the published survey and analyzed here.[4] In the German survey firms answer regarding the course of a variable by choosing one of the three categories "will decrease", "will remain unchanged", or "will increase". The method developed by Carlson & Parkin (1975) is applied to this data set in order to compute percentage changes form the qualitative data.

---

[2] A rare comparative study of firms' forecast accuracy, although not regarding output forecasting, has been published by Cassar (2014). The focus on start-up firms expectations of that study does not lend itself for an comparison with the findings of the present study. Tanaka et al. (2019) offer an assessment of firms accuracy in making predictions of the macroeconomic variables. Again their findings cannot be compared to our insights.

[3] The R^2 of such an estimate is the relevant statistic because it measures the fraction of the variation in output that is accounted for by the forecast of output. A measure of one indicates a perfect score while a variable of zero indicates that the forecast is completely uninformative.

[4] The data come from the publication Business Executives' Expectations (1978-1988).

Returning to the argument for measuring forecast accuracy by way of estimating equation (1): given the transformations that go into computing the (forecast and actual) data in both the U.S. and Germany, we need to make allowance for the fact that these transformations can induce biases into these series.[5] Hence, the estimate of a (non-zero) α-coefficient and a (non-unitary) β-coefficient helps to correct for such distortions in the indicators constructed from the survey data.[6] Summing up, the accuracy of industry forecasts is assessed by running individual regressions where the realization of output in an industry is regressed on the forecast from the previous period and a constant term. The R^2 for each branch of industry is then the measure of forecast accuracy.[7]

3. Results

We start with the results for the U.S. The findings of the statistical analysis are organized in Table 1. The ordinary least squares method yields unbiased estimates of coefficients. For the sake of completeness standard errors of the estimated coefficients are also included in the table.[8] Table 1 presents the regression results for 10 branches of industry where data are available over the period from 1978 to 1988. For the U.S. data we make the adjustment of pulling together (by an averaging of the data) the branches of primary and fabricated metals into one category. This makes possible a comparison of the U.S. to the German data. When looking at the R^2 values across industries we find remarkable differences: the printing industry is the case with the best and the food industry the case with the worst performance in forecasting.

Consider next the regression findings for Germany. In order to make the findings comparable with those from the U.S. we seek for common terms denoting industries.[9] For the German industry named "Herstellung von Kraftwagen und Kraftwagenteilen" we apply the term "transportation equipment". Further, the German industry "Herstellung von elektrischen Ausrüstungen" is termed "electric machinery" while the term "nonelectric machinery" is used for the branch "Maschinenbau". Table 2 thus shows the results for the same set of industries covered in Table 1 for the U.S.

Again, there are major differences in forecast performance with the nonelectric machinery industry the case with the best and, just as in the U.S., the food industry the case with the worst performance in forecasting.[10] The next section puts the similarities and differences across the two sets of estimates into perspective.

---

[5] This is particularly well documented for the Carlson-Parkin (1975) method where typically a non-unitary β-coefficient is estimated because the threshold level (a critical input in the method) is arbitrarily taken to be one.

[6] Note that we are not interested in testing hypotheses concerning these coefficient values. Accordingly, possible autocorrelation of residuals is of no relevance here. Furthermore, least-square estimates are efficient independent of whether residuals are normally distributed or not.

[7] The problems described here are typically not present in studies of forecast accuracy that use quantitative forecasts, e.g., for inflation or GDP (Bank of England, 2015, Günay, 2018)

[8] Due to autocorrelation in the residuals the estimated standard deviations are not efficiently estimated. However, this is not critical since we are not interested in testing hypotheses concerning parameters.

[9] Note that the demarcation of industry branches are not necessarily the same over the two countries and data sets covered.

[10] The differences in the fit of the regressions between across the U.S. and German should not be a matter of concern. They are due to the conceptual differences between the two surveys in particular the specifics of how the data gathered are transformed into published series. In particular, the overall higher R^2 values for the German industries should not be interpreted as showing better (national) forecasting abilities. The generally lower R^2 values in the estimates with the U.S. data must be seen on the background of the rather arbitrary weighting of the survey responses.

Table 1. Estimates with U.S. Data

| Industries: | Variables: | | |
	Constant	Expectation	R^2
Food	29.908 (8.351)	0.363 (0.054)	0.102
Textiles	14.522 (9.207)	0.610 (0.182)	0.185
Apparel	24.396 (9.693)	0. 469 (0.186)	0.150
Paper	20.967 (11.121)	0.574 (0.191)	0.162
Printing	15.081 (8.159)	0.655 (0.144)	0.334
Chemicals	20.179 (11.076)	0.588 (0.188)	0.192
Metals	21.187 (11.557)	0.525 (0.199)	0.104
Nonelectric Machinery	20.935 (9.935)	0.541 (0.176)	0.128
Electric Machinery	15.186 (9.596)	0.637 (0.165)	0.211
Transportation Equipment	18.915 (12.083)	0.570 (0.216)	0.158

Numbers in parentheses are Huber-White-Hinkley heteroskedasticity consistent standard errors.

Table 2. Estimates with German Data

| Industries: | Variables: | | |
	Constant	Expectation	R^2
Food	-0.079 (0.014)	0.449 (0.059)	0.366
Textiles	-0.162 (0.020)	0.741 (0.071)	0.487
Apparel	-0.080 (0.019)	0. 564 (0.060)	0.502
Paper	-0.114 (0.021)	0.829 (0.091)	0.554
Printing	-0.111 (0.016)	0.717 (0.065)	0.521
Chemicals	-0.163 (0.024)	0.988 (0.072)	0.608
Metals	-0.091 (0.019)	1.000 (0.051)	0.751
Nonelectric Machinery	-0.053 (0.011)	0.885 (0.030)	0.870
Electric Machinery	-0.066 (0.015)	1.058 (0.058)	0.759
Transportation Equipment	-0.009 (0.023)	0.727 (0.065)	0.663

Numbers in parentheses are Huber-White-Hinkley heteroskedasticity consistent standard errors.

4. Continuity and Change

In this section we compare the findings reported in Tables 1 and 2. By doing so we seek to integrate the findings from different places and different times. Instead of judging the differences in our two samples as an unfortunate fact, we seek to most fully draw interesting inferences from the available data. Even with the conceptual differences in the two surveys covered, the use of the regression approach and the focus on national rankings (avoiding comparing R^2 measures across countries) can yield remarkable insights.

As a start, consider the rankings of industries according to their forecasting accuracy for both countries. Table 3 presents these rankings for the U.S. and the German data.[11] For the interpretation keep in mind that the data studied does not only come from different countries, but also come from different times. I.e., the 1980s for the U.S. and the period since the 1990 for Germany. When considering this point several interesting insights can be gained from a comparison of the two rankings. For such an exercise Table 3 also reports the changes in the rankings of industries: in the rightmost column we report how many places an industry has gained or lost when comparing the newer German data with the older U.S. data. In particular, these gains and losses relate to the industry reported in the middle column with the German data. Firstly, the electric machinery industry is close to the top in both rankings. Secondly, the food industry is at the bottom of the rankings for both countries. Two more branches of industry (paper and apparel) differ just by one rank. Hence, seen over space and time these industries do not differ markedly when it comes to the possibilities and necessities of forecasting. These industries document the continuity in forecasting in these particular industries: even amid structural changes the role of business forecasting in these industries shows stability over time.

Table 3. Ranking of industry branches according to forecasting accuracy

U.S.	Germany	Change in Ranking
Printing	Nonelectric Machinery	+7
Electric Machinery	Electric Machinery	0
Chemicals	Metals	+6
Textiles	Transportation Equipment	+2
Paper	Chemicals	-2
Transportation Equipment	Paper	-1
Apparel	Printing	-6
Nonelectric Machinery	Apparel	-1
Metals	Textiles	-5
Food	Food	0

By contrast, there are industries where the forecast performance has changed noticeably over time. Consider two industries that rank markedly lower in the German as opposed to the U.S. ranking. Printing makes the case that structural changes affect production processes and the need for forecasting. Industry studies indicate that vast changes in the access to fast computers and printers have fundamentally changed the printing industry (Romano, 2004). The newly available highly flexible production processes have made output forecasting almost irrelevant. Similarly, the textile industry ranks much lower in the more up-to-date German data. In this line of industry cheap and highly flexible international suppliers have rendered forecasting less relevant (Sethi & Sethi, 1990, Saleeshya & Raghuram, 2012).

Further, there are also two industries – nonelectric machinery and metals – that in terms of forecasting performance are markedly higher up in the German ranking compared to the U.S. ranking. It is beyond the scope of this study to detail the industry trends that can account for these developments. Yet, we can report evidence that such trends are operating when we look at changes over the course of the three decades of data for Germany. For this purpose we estimate the same regressions shown before but now for two subsamples of equal length (1991 to 2004 and 2005 to 2018). Doing so we find that the R^2 increases in these two industries over time: for both the metals industry (the R^2 goes from 0.699 to 0.799) as well as the nonelectric machinery industry (the R^2 goes from 0.809 to 0.881) this measure of forecasting accuracy shows an increase.

Hence, a trend toward higher forecasting accuracy in these two industries is apparent not only when comparing the (earlier) U.S. and the (later) German measures but also over time in the German data. Consider also the

---

[11] Note that the difference in forecasting accuracy across industries is the result of several forces. The reason for the food industry to be at the bottom of both rankings could particularly be the result of the difficulty of prediction, i.e., the high volatility of demand and competitors' decisions.

decline in the accuracy measures reported before for printing and textiles in the rankings between the U.S. and Germany. For printing in Germany the accuracy measure declines over the two sub-periods from 0.545 to 0.527. For the textile industry the corresponding findings are even more telling: here the R^2 declines from 0.535 to 0.424. Hence, in all industries whose rankings have changed by more than two places the econometric evidence indicates persistent changes over time.

5. Summary and Conclusions

This paper makes survey data on forecasting useful for the assessment of the relative opportunities and potentials of forecasting in different industries. The study covers industry branches in the U.S. and in Germany and identifies industries that are consistently among the most/least accurate forecasters. The reported results allow an assessment of the question of who stands to benefit from business forecasting. An industry consistently high up in the ranking of forecasting accuracy concerning production is the electric machinery industry. An industry at the bottom of both rankings is the food industry. This sort of finding can help management in their decisions regarding the allocation of resources to forecasting. Consider, e.g., the findings reported by Watson (1996) regarding the underdeveloped use of forecasting in the Scottish electronics industry. With a view to our findings it appears that decision makers in this industry should make an effort to actualize the potential of forecasting.

When looking at the branches of industry that show markedly different rankings in the U.S. and the German data we find that these differences are linked to historical trends in industries. In all industries with a noticeable difference in rankings between the two national data sets we also find that the direction of change is discernable over time. When dividing the newer and longer German data set into subsamples it is found that these changes also show over subperiods. Hence, judging from the data studied here it appears that in many industries forecasting shows continuity and that in the cases where change occurs it tends to happen rather gradually. The approach presented here lends itself to be applied to data from different countries and periods. Generalizing the findings along these lines promises to yield further insights into the continuity and trends in business forecasting.

Acknowledgements

References

Armstrong, S. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners.* New York (N.Y.), Kluwer Academic Publishers. https://doi.org/10.1007/978-0-306-47630-3

Armstrong, J. S., & Collopy, F. (1992). Error Measures For Generalizing About Forecasting Methods: Empirical Comparisons. *International Journal of Forecasting, 8*(1), 69-80. https://doi.org/10.1016/0169-2070(92)90008-W

Bank of England (2015). *Evaluating Forecast Performance, London.* Independent Evaluation Office.

Business Executives' Expectations (1977-1988): *A Quarterly Report from the Conference Board Based on a Survey of Chief Executives.* New York (N.Y.), The Conference Board.

Carlson, J. A., & Parkin, M. J. (1975). Inflation Expectations. *Economica, 42*(166), 123-138. https://doi.org/10.2307/2553588

Cassar, G. (2014). Industry and Startup Experience on Entrepreneur Forecast Performance in New Firms. *Journal of Business Venturing, 29*(1), 137-151. https://doi.org/10.1016/j.jbusvent.2012.10.002

Jain, C. L. (2004). Business Forecasting Practices in 2003. *Journal of Business Forecasting Methods & Systems, 23*(3), 2-6.

Croushore, D. D. (1993). Introducing: The Survey of Professional Forecasters, Business Review. *Federal Reserve Bank of Philadelphia, 3*(6), 3-15.

Dalrymple, D. J. (1987). Sales Forecasting Practices: Results From a United States Survey. *International Journal of Forecasting, 3*(3-4), 379-391. https://doi.org/10.1016/0169-2070(87)90031-8

Elliott, G., & Timmermann, A. (2016). Forecasting in Economics and Finance. *Annual Review of Economics, 8,* 81-110. https://doi.org/10.1146/annurev-economics-080315-015346

Fildes, R., Ma, S., & Kolassa, S. (2019). *Retail Forecasting: Research and Practice.* Munich Personal RePEc Working Paper No. 89356. https://doi.org/10.1016/j.ijforecast.2019.06.004

Friedman, W. A. (2013). *Fortune Tellers: The Story of America's First Economic Forecasters.* Princeton,

Princeton University Press. https://doi.org/10.1515/9781400849864

Günay, M. (2018). Forecasting Industrial Production and Inflation in Turkey with Factor Models. *Central Bank Review, 18*(4), 149-161. https://doi.org/10.1016/j.cbrev.2018.11.003

Newbury, F. D. (1953). *Business Forecasting: A Survey of Business Practices and Methods.* New York, Controllership Foundation.

Pesaran, M. H., & Weale, M. (2006). Survey Expectations, Chapter 14 of Handbook of Economic Forecasting, 1, Eds. Graham E., Granger, C. W. J. & A. Timmermann, 715-776. https://doi.org/10.1016/S1574-0706(05)01014-1

Peterson, R. T. (1993). Forecasting Practices in Retail Industry. *Journal of Business Forecasting Methods & Systems, 12*(1) 11-14.

Romano, F. (2004). *An Investigation Into Printing Industry Trends*. Rochester (NY), A Research Monograph of the Printing Industry Center at Rochester Institute of Technology.

Rötheli, T. F. (2007). Business Forecasting and the Development of Business Cycle Theory. *History of Political Economy, 39*(3), 481-510. https://doi.org/10.1215/00182702-2007-019

Rötheli, T. F. (2018). Should Business Rely on Business Cycle Forecasting? *Central European Journal of Operations Research, 26*(1), 121-133. https://doi.org/10.1007/s10100-017-0477-8

Saleeshya, P. G., & Raghuram, P. (2012). Lean Manufacturing Practices in Textile Industries: A Case Study. *International Journal of Collaborative Enterprise, 3*(1), 18-37. https://doi.org/10.1504/IJCENT.2012.052367

Sethi, A. K., & Sethi, S. P. (1990). Flexibility in Manufacturing: A Survey. *International Journal of Flexible Manufacturing Systems, 2,* 289-328. https://doi.org/10.1007/BF00186471

Tanaka, M., Bloom, N., David J. M., & Koga, M. (2019). Firm Performance and Macro Forecast Accuracy. *Forthcoming Journal of Monetary Economics.* https://doi.org/10.1016/j.jmoneco.2019.02.008

Watson, M. C. (1996). Forecasting in the Scottish Electronics Industry. *International Journal of Forecasting, 12*(3), 361-371. https://doi.org/10.1016/0169-2070(96)00670-X