Testing Smart Crowds for the Economy

Carina Antonia Hallin, Julian Johannes Umbhau Jensen, Collective Intelligence Unit, Copenhagen Business School, Denmark
Oded Koren, Nir Perel, Shenkar School of Engineering, Design and the Arts, Israel
Sigbjørn Tveteraas, Department of Industrial Economics, Risk and Planning, University of Stavanger, Norway

1. INTRODUCTION

Conventional forecasts of macroeconomic indicators by financial institutions are typically less dynamic than the monthly or quarterly forecasts of performance variables in firms. The latest financial crisis underlined the need for monitoring and forecasting household credits, debts and savings and unemployment as these indicators are essential for avoiding an economic recession (Estrella and Mishkin, 1998). Hence, there is a justification to develop new and innovative prediction methods for reducing uncertainties about fluctuations in credits, debts and savings of households, as gauging these variables more precisely enables to provide a more accurate estimation of GDP (Zabai, 2017).

In the last decade, new prediction methods such as 'the wisdom of crowds' have evolved with the potential to advance accuracy in predictions and decision-making (Davis-Stober et al., 2014; Surowiecki, 2004). In this study we examine which crowds are smart in terms of predicting key economic indicators. We set up a prediction contest about the economy 12 months ahead in the Copenhagen Region of Denmark to test which crowds are smartest, sampling citizens, frontline employees across banks and financial experts from academic and financial institutions. As the first comparative study on the prediction capabilities of these crowds, we investigate the group-wise forecast accuracy and personal characteristics that define the respondents with the best predictions.

2. THE WISDOM OF CROWDS AND PREDICTIONS OF THE ECONOMY

It is long known that predictions by crowds can outperform individual experts (Galton, 1907), which is popularized as "the wisdom of crowds" (Surowiecki, 2004). Crowds are thought to excel because they aggregate a large number of independent beliefs; and the average beliefs tend to rule out extreme values (Clemen, 1989). A diverse crowd with relevant knowledge of the problem will be able to offer more diverse perspectives and to fill in more of the missing pieces, resulting in better solutions (Hong and Page, 2001). Moreover, aggregating crowd members' predictions can mitigate biases associated with their individual beliefs (Davis-Stober et al., 2014). Researchers have repeatedly documented gross inaccuracies in the estimates and predictions of individual experts with some few exemptions that show that individuals can be superforecasters (Mellers et al., 2015). Yet, in many prediction tasks groups have reaffirmed the accuracy of crowds (Larrick and Soll, 2006; Wolfers and Zitzewitz, 2004). For the wisdom of the crowds to be more accurate, it depends on the number of participants and the diversity of the expertise of each individual participant. The more participants involved and the more diverse the participants are, the lower the margin of error (Hong and Page, 2004).

So what does the wisdom of the crowds have to do with economic forecasting? According to Hayek (1945), a centrally planned economy will never be able to match the efficiency of the open market because what is known by a single agent is only a small fraction of the sum of knowledge held by all members of society. In other words, the central planning agency can be thought of as an individual forecast and the "open market" as a multiple-source forecast. The University of Michigan Consumer Sentiment Index and the consensus forecasts are examples of 'wisdom of crowds' forecasts that reflect the knowledge available and have been used in the prediction of US recessions (Estrella and Mishkin, 1998). However, to our awareness studies have not yet identified if there are specific crowds accumulating knowledge that are smarter than others in predicting short-term fluctuations in economic indicators, and if so, what are the specific characteristics that make them better at predicting these fluctuations.

3. METHOD

The predictions were based on survey-based data collection from three different samples of crowds: a representative citizen sample of the Copenhagen Region in Denmark, frontline employees in banks from the same region, and financial experts from the academic and private sector. The representative respondent sample was a panel-based survey by the data analytics firm YouGov that consisted of 1239 citizens living in the region. The citizens gathered points for answering the surveys which were translatable into winning products. Two sub-samples, savers and borrowers, were subset from this sample, based on their placement on the scales of Personal Savings Orientation (PSO)(Dholakia et al., 2016) and Attitude towards Credit (AtC)(Pinto et al., 2004). The 230 respondents scoring highest on the AtC scale were subset as borrowers, while the 218 respondents with highest scores on the PSO scale were subset as savers. The sample of frontline employees from banks located in the Copenhagen Region, collected through the cooperation of the Finance Union for employees in Danish banks, consisted of 78 respondents. YouGov facilitated the data gathering and respondents collected monetary rewards for the best predictions. The sample of financial experts from academia and the private sector consisted of 17 respondents, data gathering was facilitated by YouGov, and monetary rewards were given for the best predictions. The inclusion criteria for the financial expert sample were based on a list of the most proficient economists in Denmark as reported by the Danish finance newspaper Børsen (Bjørnskov, 2017).

The different samples went through the same prediction contest, predicting 1, 3, 6 and 12 months ahead of the survey-month, November 2017. Prediction variables were as follows; per capita mortgage debt, per capita bank debt, per capita credit debt, per capita bank deposit and the number of full-time unemployed in the Copenhagen Region. Credit debt refers to quick loans with no safety, no direct product relation and a runtime of maximum 3 months, as defined by the Danish competition and consumer authority. Prior to the predictions, respondents were given the actual economic figures for the variables from the 4 previous months, 2-5 months earlier. The forecast accuracy on the five financial variables was calculated using the actual economic data, once available, provided by Center for Regional and Tourism studies (CRT), an economic institute working closely with Denmark Statistics (DST). Due to the lag in economic data, the predictions for 12 months ahead could not be part of the conclusions of this preliminary paper.

To measure prediction accuracy we used the Mean Absolute Percentage Error (MAPE). MAPE is a standard measure of forecast accuracy, which, for example, is used in forecasts of demography (Swanson, 2000), economics (Hyndman and Koehler, 2006) and machine learning (Espinoza et al, 2005). MAPE shows how much the collective prediction of the crowd deviated from the actual outcome in absolute percentage terms for the different economic indicators. The crowd attempts to predict a future outcome for each variable with the collective prediction given by the average of the individual members' predictions over time: *crowd*=the different samples, *i*=1…n are the individuals of the crowd, *t*=3 prediction times, *k*=the five different variables.

$$mape_k^{crowd} = \frac{100\%}{n} \sum_{t-1}^{n} \left| \frac{prediction_{tk}^{crowd} - actual_{tk}}{actual_{tk}} \right|$$

4. RESULTS

The sample of citizens consisted of 52.5% females and an average age of 47 years. The frontline sample constituted of 78 respondents perfectly split between females and males and an average age of 46.5 years. The expert sample consisted of 17 respondents with 17.6% female. The average age of this group was 50 years.

Table 1. Demographics – Age and Gender

Sample group	N	Female	Male	Average age
Citizens	1239	52.5%	47.5%	47.1 (17.6)
Citizens (Borrowers)	230	51.7%	48.3%	39.1 (15.1)
Citizens (Savers)	218	61.9%	38.1%	45.6 (17.5)
Frontline	78	50%	50%	46.5 (11.4)
Experts	17	17.6%	82.4%	50.1 (10.6)
Total	1334	693	641	45.7

The MAPE results in Table 2 shows a clear ranking of the prediction capabilities of the groups in all variables. Experts provide the most accurate predictions, followed relatively close by the frontline. Then comes the subsample of savers, followed by the full citizen sample and finally, and with the least performing predictions on all accounts by the subsample of borrowers. The main takeaway from these results is that wisdom of crowds does actually improve when the crowds' participants have relevant or - even better - expert knowledge about the subject matter. Even if the experts outperform the other groups, it is interesting to note that the frontline is not far behind on forecasting accuracy.

Table 2. MAPE Analysis

Sample	N	Mortgage	Bank debt	Credit	Bank deposit	Unemployment	Average
Citizens	1239	6.80%	8.61%	57.12%	6.55%	11.87%	18.19%
Borrowers	230	19.23%	16.31%	149.72%	15.96%	29.11%	46.07%
Savers	218	4.37%	6.59%	36.31%	5.54%	6.58%	11.88%
Frontline	78	1.90%	4.72%	5.59%	2.71%	3.46%	3.67%
Experts	17	0.97%	4.30%	2.80%	1.97%	3.42%	2.69%

5. CONCLUSION

The results of the prediction contest show a clear tendency of what groups provide the best prediction accuracy of the variables. Based on the results, we can conclude that the more knowledgeable the crowd is with finance and economics, the better the prediction performance is. Furthermore, we can conclude that the respondents scoring high on PSO and AtC have a distinct difference in predictive performance. The mindset of the respondents seems to have an effect on predictive performance.

With these preliminary findings in mind, this paper will elaborate further on the characteristics that define the different groups and their prediction performance. Characteristics cover a wide range of possible factors, from respondent demographics to personality traits, purchase behavior and mental health. We expect that this will provide insights into how to construct the right crowd for predictions purposes of fluctuations in the economy. A preliminary conclusion we can draw from the empirical findings is that both a broad and deep insight in the respective topic for prediction matter to become a smart crowd.

ACKNOWLEDGEMENTS

This research received funding from a block grant in the research program "Nordic Finance and the Good Society" by Center for Corporate Governance, Copenhagen Business School. We would like to thank Center Director, Dr. Lars Ohnemus, for encouraging us to apply for a grant in the program to test smart crowds in the financial sector. We would also like to thank the Danish Finance Union for employees in Danish banks for their collaboration in the administration of surveys to frontline employees. Finally, but not least, we would like to thank Research Assistant Frederik K. Larsen for his contributions during the project period.

REFERENCES

Bjørnskov, C. 2017. Debate: Here are Denmarks most proficient national economists. *Børsen.dk*. Found on the 08/02/2019 at https://borsen.dk/nyheder/opinion/artikel/1/345855/debat_her_er_danmarks_dygtigste_nationaloekonomer.html

Clemen, R.T. 1989. Combining Forecasts: A Review And Annotated Bibliography. *International Journal of Forecasting*. 5:559-583.

Davis-Stober, C.P., Budescu, D.V., Dana, J., and Broomell, S.B. 2014. When Is A Crowd Wise?. *Decision*. 1(2), 79-101.

Dholakia, U., Tam, L., Yoon, S. and Wong, N. 2016. The Ant and the Grasshopper: Understanding Personal Saving Orientation of Consumers. *Journal of Consumer Research*. Volume 43, Issue 1, 1 June 2016, Pages 134–155. https://doi.org/10.1093/jcr/ucw004

Espinoza, M., Suykens, J.A.K., De Moor, B. 2005. Load Forecasting Using Fixed-Size Least Squares Support Vector Machines. In *Cabestany J., Prieto A., Sandoval F. (eds) Computational Intelligence and Bioinspired Systems. IWANN 2005. Lecture Notes in Computer Science.* vol 3512.

Estrella, A., and Mishkin, F.S. 1998. Predicting U.S. Recessions: Financial Variables as Leading Indicators. *The Review of Economics and Statistics*. Vol. 80, No. 1 (Feb., 1998), 45-61.

Galton, F. 1907. Vox Populi. *Nature*. 75:450-451.

Hayek, F.A. 1945. The use of knowledge in society. *The American economic review*. 35(4): 519-530.

Hong L. and Page S. 2001. Problem solving by heterogeneous agents. *Journal of Economic Theory*. 97(1):123-163.

Hong L. and Page S. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*. 101(46):16385-16389.

Hyndman, R.J. and Koehler, A.B. 2006. Another Look At Measures Of Forecast Accuracy, *International Journal Of Forecasting*. Volume 22, Issue 4, 2006, Pages 679-688. Https://Doi.Org/10.1016/J.Ijforecast.2006.03.001.

Larrick, R.P., and Soll, J.B. 2006. Intuitions about combining opinions: Misappreciation of the averaging principle. *Management Science*. 52:111-127.

Mellers, B., Stone, E., Murray, T., Minster, A., Rohrbaugh, N., Bishop and M., Tetlock, P. 2015. Identifying And Cultivating Superforecasters As A Method Of Improving Probabilistic Predictions. *Perspectives On Psychological Science*. 10(3), 267–281. Https://Doi.Org/10.1177/1745691615577794

Pinto, M.B., Mansfield, P.M. and Parente, D.H. 2004. Relationship Of Credit Attitude And Debt To Self-Esteem And Locus Of Control In College-Age Consumers. *Psychological Reports*. 2004, 94, 1405-1418.

Swanson, D.A., Tayman, J. and Barr, C.F. 2000. A note on the measurement of accuracy for subnational demographic estimates. *Demography*. 37: 193. DOI: Https://doi.org/10.2307/2648121

Suroweicki, J. 2004. The Wisdom Of Crowds. *Doubleday, New York, Ny.*Wolfers, J., and Zitzewitz, E. 2004. Prediction markets. *Journal of Economic Perspectives*. 18(2):107-126.

Zabai, A. 2017. Household Debt: Recent Developments And Challenges. In *BIS Quarterly Review*. December 2017.