Identifying Experts in Virtual Forecasting Communities

Florian Teschner
Karlsruhe, Germany., teschner@gmail.com

Christof Weinhardt
Karlsruhe, Germany., weinhardt@kit.edu

Identifying Experts in Virtual Forecasting Communities

Florian Teschner
Karlsruhe Institute of Technology (KIT)
Teschner@kit.edu

Christof Weinhardt
Karlsruhe Institute of Technology (KIT)
Weinhardt@kit.edu

ABSTRACT

Macroeconomic forecasts are used extensively in industry and government even though the historical accuracy and reliability is questionable. Over the last couple of years prediction markets as a community forecasting method have gained interest in the scientific world and in industry. An arising question is how to detect valuable user input and identify experts in such online communities. Detecting such input would possibly enable us to improve the information aggregation mechanism and the forecast performance of such systems. We design a prediction market for economic derivatives that aggregates macroeconomic information. Using market-based measures we find that user input can be evaluated ad-hoc. Further analysis shows that aggregated measures outperform established methods -such as reputation- in identifying forecasting experts. Moreover, using data from a two year field-experiment we find that expertise is stable for longer time horizons.

Keywords

Online Communities, Expert Identification, Prediction Markets, Macroeconomic Forecasting

INTRODUCTION

A wide and important range of policy decisions are made on the informational basis of economic forecasts such as GDP growth. It is a well-established fact that traditional economic forecast models lack the necessary accuracy (Osterloh 2008; McNees 1992; Schuh 2001). Simplified, the current approaches mix expert knowledge with historic extrapolation. They are thus inadequate to capture rapid economic changes, as exemplified in the 2008 recession. Internet communities offer the advantage of instant information exchange. But how can online communities be designed to facilitate information aggregation of macroeconomic variables? Running such a community confronts us with the questions of how to identify valuable user input. Especially in the domain of forecasting, ad-hoc evaluation of participant estimates might help increasing the overall forecast performance. Moreover, it seems fruitful to identify experts in such communities for qualitative surveys or interviews. A common approach is use self-rated expertise or reputation-based approaches. However a recent study shows that reputation based expertise is not a good predictor for future forecast performance (Armstrong 2008).

Over the last couple of years prediction markets as a game-like forecasting method have gained interest in the scientific world and in industry. They facilitate and support decision making through aggregating expectations about events (Hahn and Tetlock, 2006). The roots of their predictive power are twofold; the market provides the incentives for traders to truthfully disclose their information and an algorithm to weight opinions (Arrow et al., 2008). We thus setup a prediction market for economic variables called Economic Indicator Exchange (EIX). The EIX play money prediction market is specifically designed to continuously forecast economic indicators such as GDP, inflation, IFO index, investments, export and unemployment figures in Germany.

Evaluating market-based expert identification measures we find that user input can be evaluated ad-hoc. Furthermore, we show that the aggregated measures perform better than other methods (such as reputation) identifying forecasting experts. Finally by analyzing individual forecast performance over time we find that these measures reliably predict future forecast performance.

The remainder of this paper is structured as follows: The second section gives a brief review of previous markets for economic variables. Furthermore expert identification in online communities is discussed. Section thee summarizes the research questions. The forth section presents the IS-artifact and details the field experiment setting. The subsequent section details our measures used to rate user input and identify experts. Section six evaluates these measures from different perspectives. Finally section seven concludes this paper.

RELATED WORK

Prediction Markets as Online Communities

A common approach to economic forecasting is to identify experts based on reputation who can make a prediction. These experts use statistical models combined with heuristics, which are based on an expert's experience and intuition. However reputation-based experts are prone to biases and political influence and generally do not perform better than novices in forecasting future events (Armstrong 2008). Furthermore macroeconomic forecasts suffer from the optimism bias (Batchelor, 2007) and imitation behavior (Osterloh, 2008). Group decision making is a technique often applied to deal with these limitations. Internet communities offer the advantage of instant information exchange and group decision that is not possible in a real-life. An arising question is how to build and maintain internet communities to forecast macroeconomic variables. Furthermore how can well-informed people be motivated and incentivized to participate in information sharing and collaboration? A certain type of online communities, so called prediction markets have emerged as a forecasting tool for wide range of applications.

Prediction markets facilitate and support decision making through aggregating expectations about events (Hahn and Tetlock, 2006). In most cases they allow anonymous participation, which may increase the likelihood of nonconformist to participate and reveal information. The roots of their predictive power are twofold; the market provides the incentives for traders to truthfully disclose their information and an algorithm to weight opinions (Arrow et al., 2008).

The most basic trading mechanism for prediction markets is based on a continuous double auction for one stock which represents the outcome of an event. The stock will pay 1 if an event has the predicted outcome and else the stock will be worthless. Market participants form expectations about the outcome of an event. Comparable to financial markets, they buy if they find that prices underestimate the event in question and they sell a stock if they find that prices overestimate the probability of an event. Thus communication in such a system is limited to the market language; bids and offers.

Markets for Economic Outcomes

Financial markets for macroeconomic variables have been used since the 80s. The Coffee, Sugar and Cocoa Exchange established a futures market on the consumer price index allowing traders to hedge on inflation. The market, however, was closed due to low interest (Mbemap 2004). In 1993 Robert Shiller argued for the creation 'Macro Markets' which would allow a more effective risk allocation (Shiller 1993). In an attempt to set up a market to predict economic variables in 2002 Goldman Sachs and Deutsche Bank created the so called 'Economic Derivatives' market. It tries to predict macroeconomic outcomes such as ISM Manufacturing, change in Non-Farm Payrolls, Initial Jobless Claims and consumer price index (Gadanecz et al., 2007). The traded contracts are securities with payoffs based on macroeconomic data releases. The instruments are traded as a series (between 10-20) of binary options. For example a single data release of the retail sales in April 2005 was traded as 18 stocks. In order to maximize liquidity the market operators use a series of occasional Dutch auctions just before the data releases instead of the more common continuous trading on most financial markets. Thus the market provides hedging opportunities against event risks and a short horizon market forecast of certain economic variables. By analyzing the forecast efficiency Gurkaynak and Wolfers (2006) find that market generated forecasts are very similar but more accurate than survey based forecasts[1].

In an attempt to forecast inflation changes in Germany, Berlemann and Nelson (2005) set up a series of markets. The markets feature continuous trading of binary contracts. In a similar field experiment Berlemann et al. (2005) use a similar system in order to aggregate information about inflation expectations in Bulgaria. All in all, the reported forecasts results in both experiments are mixed but promising.

Identifying Experts in Online Communities

Various approaches have been employed to identify experts in online communities. However, researchers have detected the need to employ systematic approaches in finding local knowledge (Davis and Wagner, 2003). The most common approach of presenting expertise is a list of people ranked by the number of inputs they submit. These lists may reflect whether a person knows about a topic, but it is difficult to distinguish that person's relative expertise levels or even judge about a single contribution. Using a large help-seeking community, Zhang et al. (2007) use social network analysis methods to form expertise networks. They are able to automatically evaluate relative expertise. However, the system does not provide any real-time evaluation of user input.

[1] One must note that the Bloomberg survey forecasts are published on Fridays before the data release, whereas the auction was run -and the forecast was generated- on the data release day.

Linked to the idea of identifying expertise is the notion of lead-user detection. Lead users *"are users whose present strong needs will become general in a marketplace"* (von Hippel 1986; Urban and von Hippel 1988). Following this notion, Spann et al. (2006) used a virtual stock market for box office revenues to identify lead-users. They point out that in virtual stock market communities there are two mechanisms at work selecting users. First, a self-selection effect occurs as users who are not interested drop out of the community. Secondly, a performance effect takes place, as users performing below average run out of cash. In their setting, they show that a portfolio based ranking correlates well with a survey based lead-user detection analysis. However, it remains unclear if a portfolio based approach is the best measure to capture expertise and if these expertise is persistent over time.

RESEARCH QUESTIONS

As online communities become more important for people to seek und share information, one key question is how to identify valuable user input. Especially in the domain of forecasting, ad-hoc evaluation of participant estimates might help increasing the overall forecast performance. As most reputation settings aggregate individual expertise over a range of actions, it seems useful to identify experts on a more detailed level. Given that participants can be successfully identified as above average forecasters, this subsequently leads to question if individual forecast performance stays stable over a long time period. Hence, if the identification measures can be used to predict future forecast performance.

Whereas in macroeconomic forecasting domain the outcome (e.g. growth in the last quarter) is known after a relative short time span, there are some cases such as technology foresight and product innovation in which the outcome might unknown. However, especially in those cases we would like to measure the valuable input and identify lead-users. As a consequence another question is, if we can the find expert identification measures which do not rely on knowing the outcome.

A VIRTUAL FORECASTING COMMUNITY FOR ECONOMIC VARIABLES

In October 2009 a play money prediction market was launched specifically designed to forecast economic indicators such as GDP, inflation, investments, export and unemployment figures in Germany. The goal is to forecast the indicators over longer time periods in advance and continuously aggregate economic information. The market called Economic Indicator Exchange (EIX)[2] was launched in cooperation with the leading German economic newspaper 'Handelsblatt'. The cooperation aims at reaching a wide and well informed audience interested in financial markets and economic development. We thus expect no problems understanding the indicators and the concept of trading. The market is publicly available over the Internet and readers where invited to join. The registration is free and requires besides a valid email address just minimal personal information.

Market & Contract Design

The market design features a continuous double auction without designated market maker. Participants are allowed to submit marketable limit orders with 0.01 increments through the web-based interface. After registration participants are endowed

$$p = 100 + \alpha + \left(\frac{I_{t0} - I_{t-1}}{I_{t-1}}\right) \qquad with\,\alpha = 10 \qquad (1)$$

with 1,000 stocks of each contract and 100,000 play money units. We propose to represent continuous outcomes with one stock and define a linear payout function. Contracts for each economic indicator are paid out according to equation 1.

A contract is worth: 100 +/- α times the percentage change for an indicator in play money (e.g. a change of 2.1 % results in a price of 121). We set α to 10. Therefore the representable outcome ranges from -10% to infinity. To represent the whole outcome range from -100% to infinity α could be set to one. Previous work indicates that market participants find it difficult to estimate minor changes in the underlying (Stathel et al. 2009). Hence we propose to scale the minor changes to a certain level. Looking at historical data there were no events where German GDP dropped 10% per quarter. The rationale for setting α to 10 was the deliberation that participants find it more intuitive to enter integers in order to express reasonable accuracy. Additionally German statistical data releases rarely come with more than one decimal.

Table 1 summarizes the economic variables tradable on the market. Due to the payout function and the selection of the corresponding units; all stock prices are expected to roughly range between 50 and 150. Therefore participants could similarly gain by investing in specific indicators. The indicators are a mix of leading -forecasting the economy- (e.g. Investments) and lagging -describing the state of the economy-(e.g. Unemployment numbers) economic indicators. To facilitate longer forecast horizons every indicator is represented by three independent stocks each representing the next three

---

[2] www.eix-market.de

data releases (t_1, t_2, t_3). As a consequence the initial forecast periods vary between one month for monthly released indicators and up to 3 quarters for quarterly released variables. One day before the release date the trading in the concerned stock is stopped. Finally the stocks are liquidated according to the payout function defined in equation 1. As soon as the trading in one stock stops a new stock of the same indicator (e.g. t_4) is introduced into the market. This means that participants received 1,000 new stocks of the respective indicator. All in all participants are able to continuously trade 18 stocks at all times.

The web portal features more information such as available account information for individual traders which includes the number of shares held in each contract, the balance of the cash account, the total value of their deposit, a list of outstanding buy and sell orders, as well as a list of trades. The portal also provides more information on the prizes traders can win; the operational principle of the prediction market including a video tutorial and frequently asked questions, as well as up-to-date news stream related to the German economic development.

Indicator	Unit	Data Release Cycle	Payout Number	Payout Function
Exports	%-Changes$_{t-1}$	monthly	25	$100 + \alpha \times \left(\frac{I_{t0} - I_{t-1}}{I_{t-1}}\right)$
GDP	%-Changes$_{t-1}$	quarterly	8	$100 + \alpha \times \left(\frac{I_{t0} - I_{t-1}}{I_{t-1}}\right)$
IFO Index	ABS-Changes$_{t-1}$	monthly	16	ABS(IFO Index)
Inflation	%-Changes$_{t-12}$	monthly	25	$100 + \alpha \times \left(\frac{I_{t0} - I_{t-12}}{I_{t-12}}\right)$
Investments	%-Changes$_{t-1}$	quarterly	8	$100 + \alpha \times \left(\frac{I_{t0} - I_{t-1}}{I_{t-1}}\right)$
Unemployment	Million (ABS)	monthly	25	$100 + \left(\frac{ABS(Number)}{100.000}\right)$

Table 1. Economic variables

Trading Interface

The trading interface is displayed in figure 1. Participants have convenient access to the order book with 10 accumulated levels of visible depth (I1), the price development (I2), the account information (I3) and market information (I4) such as the last trading day. As additional information the Handelsblatt provides access to an up-to-date economic news-stream (I5) and finally the indicator's last year's performance (I6) is displayed. Participants are able to customize their trading interface individually. By clicking the small arrows the six information panels open and close. In the default setting, only the trading mask and the six headlines are visible. After each submitted order the chosen interface is saved per user. On user return the system opens the previously used interface elements on default. Moreover, a short description of the market comprising the respective payoff function is shown as part of the trading screen. The design reflects recent findings that designers have to avoid overloading participants cognitively (Blohm et al. 2011).

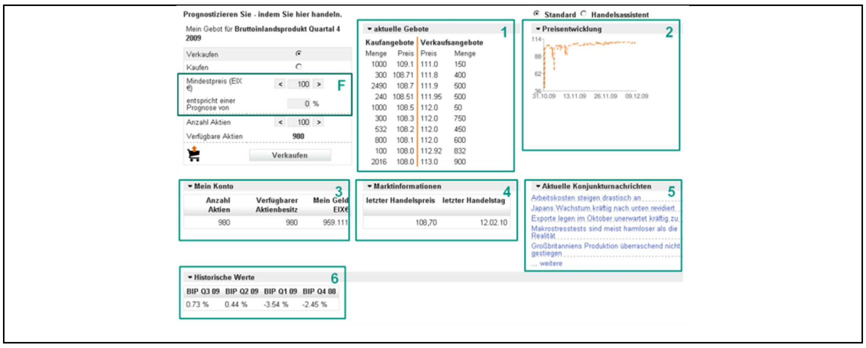

Figure 1. Trading screen with open information panels (1-6)

Incentives

As mentioned the market is a free to join play money market. Previous work has shown that play-money markets perform equally well as real-money markets at predicting future events (Rosenbloom and Notz, 2006). Note also that due to legal restrictions on gambling the EIX prediction market has to rely on play money. In order to motivate participants intrinsically we provided two interface features; traders could follow their performance on a leader board and they could form groups with others to spur competition with friends. To increase participants' motivation and to provide incentives to contribute information we hand out prizes worth 36,000 Euro. Incentives are divided in two parts (a) monthly prizes and (b) yearly prizes. The 8 yearly prizes (total value 10,000 Euro) are handed out according to the portfolio ranking at the end of the market. The monthly prizes are shuffled among participants who fulfill two requirements for the respected month: (i) they increase their portfolio value and (ii) they actively participate by submitting at least five orders. Both incentives are clearly communicated through the interface. For the yearly prizes the leader board indicates the current status of all participants. The monthly winning status is displayed individually just after each login.

Software Architecture

In addition to the key design elements of the EIX prediction market described, one also has to design the web-based trading software as well as the facilities handling information about the traders' accounts, the order matching and quote updates from a technical point of view.

The EIX prediction market software is an advancement of two previously run (Stathel et al. 2009). The system is implemented in Grails. It features a modularized architecture in order to keep it easy to maintain and expendable by services and functionality. Due to the previously unknown number of users the software platform has to be scalable. The system can be described from three perspectives; IT-infrastructure, application logic and the core order management. The IT-infrastructure is provided by the Forschungszentrum Informatik, Karlsruhe (FZI), it consists of three physical servers; a Squid reverse proxy -caching the static pages, a designated PostgreSQL server for the database and a tomcat application server - running the application logic. The application logic has been set up following the model-view-controller concept. Therefore it is separated in three layers; one handling the external communication e.g. the website presentation, one for the internal database querying and finally one running the core order processing. As the core element the order management processes all incoming orders. The EIX market employs the commonly used trading mechanism; the continuous double auction (CDA). In a CDA known e.g. from the Deutsche Börse system Xetra, traders submit buy and sell orders which are executed immediately if they are executable against orders on the other side of the order book (Madhavan 1992). If orders are not immediately executable, orders are queued in an order book and remain there until they are matched with a counter-offer, or are actively deleted by either the market operator or the submitting participant. Orders are executed according to price/time priority, i.e. buy orders with a higher limit and vice versa sell orders with a lower limit take priority. In case several orders were placed with the same limit price, the orders which were submitted earlier are executed first. One of the main advantages of using a CDA is the fact that markets with a CDA pose no financial risk for market operators as they are a zero-sum game. Moreover, the CDA allows for continuous information incorporation into prices and consequently traders are capable of quickly reacting to events. This setting helps to fulfill the virtual community requirements identified by Leimeister et al. (2004) of security, up-to-dateness and quality of the provided content.

A Two Stage Experiment

The EIX-market-game was setup as a one year field-experiment. As we received positive feedback and promising forecast results, we decided to continue the experiment for a second year. We started the second market period on October 1st 2010. As the first market closed on October 31st 2010, we had a smooth transition. Every market participant who registered for the first version was automatically transferred to the second round. No new registration was required and the website layout, web-address and institutional setting remained the same. In order to continuously improve our platform, we added some minor features and slight changes to the market design. E.g. the price for IFO index stocks is directly related to the underlying (P = IFO index (points)). The intuition was to make it easier for participants to translate a prediction into a limit-price. Due to a lower number of sponsors, the amount of prize money was reduced. We handed out three prizes worth 1,030 Euro per month – 12,360 Euro overall.

METHOD

A common approach to measure participation is to take the number of user inputs as a proxy for engagement. However that does not quantify the input's quality. In systems without reputation systems, one might rely on a self-rated assessment. Previous work showed that self-rated expertise is not a good indicator for forecasting ability. In the following section we present quantifiable measures to rate participation input in online markets.

Quantifying Single Participation Input

The most natural rating of a user input in markets is to calculate the resulting profit. However, this has several drawbacks. First off all, the profit can only be calculated ex-post if the participant closes the position. Secondly, as various market segments have different underlying uncertainty, one should consider risk adjusted profits (e.g. Sharpe ratio).

An immediate approximation of an order's information content is the price impact. The price impact approximates the permanent impact of a trade under the assumption that information impacts are permanent and realized at the x-minute mark. Following a trade, liquidity suppliers adjust their beliefs about the fundamental value of an asset depending on the information content of a trade (cf. Zhang et al. 2011). Let $Ask_{i;t}$ be the ask price for a $stock_i$ at time t and $Bid_{i;t}$ the respective bid price. $Mid_{i;t}$ denotes the mid quote then the simple price impact of a trade is calculated as follows:

$$PI_{i,t} = D_{i,t} * \left(\frac{Mid_{i,t+x} - Mid_{i,t}}{Mid_{i,t}}\right) \qquad (2)$$

$D_{i;t}$ denotes the trade direction, -1 for a sell and +1 for a buy order. The price impact provides an indication of the information content of a trade. As prediction markets incorporate information slower than financial markets we set x to 180 minutes[3].

In most prediction markets we can observe the outcome, i.e. the fundamental value of each stock. Therefore, we can ex-post measure the information content of each order. If an order moved the price in the right direction with respect to the final outcome of the stock, it is informed; whereas an order moving the price in opposite direction to the final outcome price, it is uninformed.

Thus we present the following score to capture this process:

$$Score_{i,o} = \begin{cases} 1, price_{o,i} \le fv_i \wedge o_{type} = BUY \\ 1, price_{o,i} \ge fv_i \wedge o_{type} = SELL \\ 0, price_{o,i} > fv_i \wedge o_{type} = BUY \\ 0, price_{o,i} < fv_i \wedge o_{type} = SELL \end{cases} \qquad (3)$$

The price of an $order_o$ for the $stock_i$ is represented as $price_{o,i}$. The fundamental final outcome value of a stock is represented by fv_i. In other words the score rates an order as profitable or not. Moreover we can extend the $Score_{i,p}$ by multiplying it with the order size. The order-size can be interpreted as the confidence a trader places in his bet. In the following section we will refer to this as Score(Q)[4].

Quantifying Overall Participation

Finally, individual inputs need to be combined and participants need to be ranked according to a simple metric. The most common approach is to display a ranking based upon the accumulated individual profits. However such a ranking is judging traders by both their effort and their skill. As described it seems useful to find (forecasting) experts and separate these from those investing a lot of time. Thus, we propose to rank participants according to the previously defined measures. Furthermore these measures can be calculated and displayed for the market's sub-categories. In our market setting, one can imagine a trader to be good at predicting exports (due to his company insights) but below average in predicting the unemployment numbers.

RESULTS

The following section first presents some descriptive market statistics such as participant activity. Evaluating our previously described measures we find that the Price Impact can be used to identify relevant user input in real-time. Moreover we find

---

[3] We tried various timespans ranging from 3 to 24 hours, all leading to the qualitatively same results.

[4] More precisely, before multiplying the order order-size with the Score, we adapt the Score, by exchanging the 0s by -1s. (e.g. A not profitable order leads to a negative Score(Q)).

that aggregated profits and Score(Q) can be used to identify expert participants. Predicting future forecast performance we find that the measured expertise is stable over a long time period.

Participant Activity

The following data includes the time span from 30th October 2009 till 31st of October 2011. In total 1,235 (1,006 in the first round) participants registered at the EIX market, of those 809 (680) submitted at least one order. Upon registration we asked participants to self-assess their knowledge of the German economy (further coded as *Self-Assessment*).

Altogether participants submitted 79,334 (45,808) orders resulting in 34,028 (22,574) executed transactions. Figure 2 shows the market activity over time. In the respected time frame 107 (47) stocks were paid out. In order to keep participants active and informed we sent out a weekly newsletter summarizing up-to-date economic news.

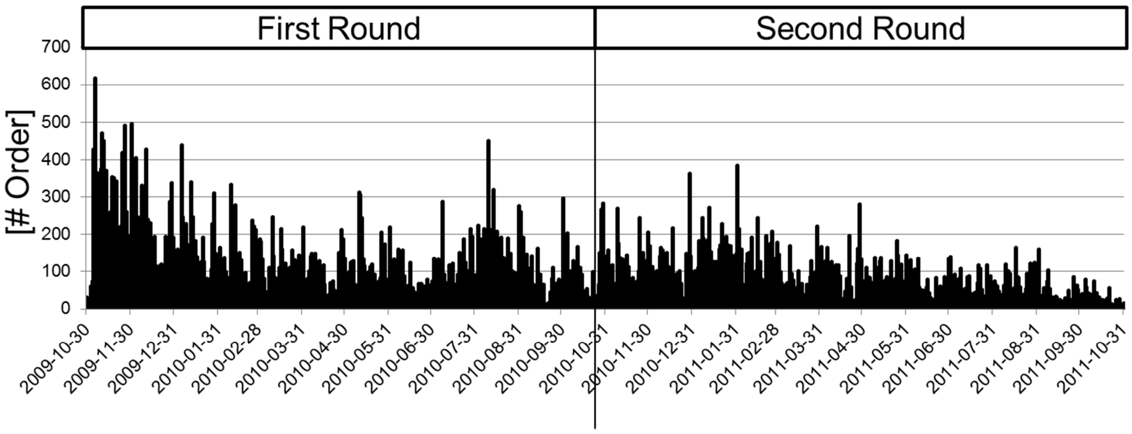

Figure 2.Activity over the game period

Previous work showed that the market-generated forecasts performed well in comparison to the 'Bloomberg'- survey forecasts, the industry standard (Teschner et al. 2011).

Identifying Informative Participation

In order to evaluate measures to identify relevant user input in real time, we start by correlating the measures on a trade by trade basis. Table 2 depicts the results.

	Score	Score (Q)	Profit	Price Impact
Self-Assessment	-0.08[a]	-0.1[a]	-0.04[a]	-0.002
Score	1	0.32[a]	0.17[a]	0.002
Score (Q)	-	1	0.46[a]	0.01[c]
Profit	-	-	1	0.02[a]

Table 2. Correlation analysis[5]

We see that Score, Score(Q) and Profit are –as expected- well correlated. Furthermore it seems that the Self-Assessment does not provide any information about the input relevance (negatively correlated).

One should note that Score, Score(Q) and Profit are available ex-post and thus do not enable us to separate input dependent on relevance. However, it seems reasonable to assume that profit is the most accurate measure to express input relevance. The two measures which are available in real-time are the Self-Assessment and Price Impact. Hence in order to evaluate which one contains more information regarding the input relevance (ad-hoc) we ran an OLS-regression predicting Profit.

[5] The superscript 'a' denotes significance at the 0.1%, 'b' at the 1% level and 'c' at the 5% level

Table 3 shows the results. We see that the self-assessment has a negative, small coefficient, indicating that participants are overconfident about their forecast abilities and self-rated user expertise is not related to forecast accuracy. This is in line with previous work (e.g. Armstrong 2008; Riedl at al. 2010).

	Profit	
	Estimate	*t*-stat.
Intercept	10,933[a]	3.35
Self. Assessment	-3,687[a]	-7.44
Price Impact	21,053[a]	4.05

Table 3. Predicting positive user contributions[4]

However we also see that the price impact is a better predictor for a successful, positive contribution.

Identifying Informed Participants

Next we aggregate all individual user input on a user level. Most common in markets is to rank participants according to their cumulated portfolio. This provides us with a benchmark scenario. Using an OLS-regression we try to test which aggregated measure contains the most information about the final individual rank (Table 4, left side). Again we find that the self-assessment contains no information about a user's overall performance. Turning to the informative variables we see that the aggregated Score(Q) and aggregated Profit predict the rank very well. For example the estimates show that the higher the profit the lower (better) the overall rank. As one might argue that the rank is not directly indicating that user is an expert we split participants in three groups; experts (Top 100), average, and low performer (Bottom 100). Dropping the average group, we have to two groups left. Running a Logit-regression (Table 4, right side) we see that a higher aggregated Profit/Score(Q) increase the chances that a participant belongs to the expert group. Hence this confirms our previous findings.

	Rank (OLS)		Experts (Logit)	
	Estimate	*t*-stat.	Estimate	χ^2
Intercept	499.9[a]	10.4	-2.4[a]	13.5
Self. Assessment	-16.8	-1.3	0.15	1.14
#Orders	0.02	0.9	0.0017[c]	6.2
Sum. Profit	-0.00007[a]	-5.57	0.000004[a]	17.1
Avg. Score	-142.1[b]	-2.3	3.28[a]	12.2
Sum. Score(Q)	-0.0006[a]	-3.47	0.00003[a]	17.0
Sum. Price Impact	293.3	1.06	-6.2	0.2

Table 4. Predicting individual overall performance[4]

Is Forecast Performance Stable over Time

The question arises if the quality of user input is stable over time. Moreover, we would like to know if previously measured expert knowledge can be used to predict future forecast performance. Following our previous methodology, we use measures from the first round (year) to predict the users' performance in the second round (year). We find that a higher Score (Q) in the first round is correlated with a lower (better) rank in the second year. Surprisingly the ranking as well as the profits in the first round have no significant predictive power.

	Rank (Round 2) (OLS)	
	Estimate	*t*-stat.
Intercept	185.8[a]	5.1
Rank (Round 1)	-0.01	-0.2
Sum.Profit (Round 1)	-0.0	-1.1
Avg. Score (Round 1)	-87[c]	-1.8
Sum.Score(Q) (Round 1)	-0.0001[b]	-2.3
Sum.Price Impact (Round 1)	-5.4	-0.1

Table 5. Predicting future forecast performance[4]

Hence we conclude that our employed measures can be used to identify experts, and this expertise is stable over a two year period.

CONCLUSIONS

Internet communities offer the advantage of instant information exchange and group decision that is not possible in a real-life. We designed an online community facilitating information aggregation of macroeconomic variables. Following the notion of lead-user detection, we test various measures to identify forecast experts.

Evaluating market-based expert identification measures we find that user input can be evaluated ad-hoc. This is important as this might be used to improve the aggregated community forecasts. In a second step we show that the aggregated measures can be used to identify forecasting experts. Benchmarking with standard methods such as reputation or a portfolio ranking, we show that our measures outperform other methods. Finally, by analyzing individual forecast performance over time we find that these measures are stable over time and reliable predict future forecast performance.

REFERENCES

1. Armstrong, J. S. (2008). Combining Forecasts. PRINCIPLES OF FORECASTING: A HAND BOOK FOR RESEARCHERS AND PRACTITIONERS, J. Scott Armstrong, ed., Norwell, MA: Kluwer Academic Publishers, 2001.
2. Arrow, K. J., R. Forsythe, M. Gorham, R. Hahn, R. Hanson, J. O. Ledyard, S. Levmore, R. Litan, P. Milgrom, F. D. Nelson, G. R. Neumann, M. Ottaviani, T. C. Schelling, R. J. Shiller, V. L. Smith, E. Snowberg, C. R. Sunstein, P. C. Tetlock, P. E. Tetlock, H. R. Varian, J. Wolfers, and E. Zitzewitz (2008, May). Economics: The promise of prediction markets. Science 320(5878), 877–878.
3. Batchelor, R. (2007). Bias in macroeconomic forecasts. International Journal of Forecasting 23(2), 189 – 203.
4. Berlemann, M., K. Dimitrova, and N. Nenovsky (2005). Assessing market expectations on exchange rates and inflation: A pilot forecasting system for Bulgaria. SSRN eLibrary.
5. Berlemann, M. and F. Nelson (2005). Forecasting inflation via experimental stock markets some results from pilot markets. (Ifo Working Papers No. 10).
6. Blohm, I., Riedl, C., Leimeister, J. M. and Krcmar, H. (2011). "Idea Evaluation Mechanisms for Collective Intelligence in Open Innovation Communities: Do Traders outperform Raters?" 'Proceedings of the International Conference on Information Systems (ICIS)', Shanghai, China, Paper 19
7. Davis, A. and Wagner, J. (2003). "Who knows? On the importance of identifying “experts” when researching local ecological knowledge," *Human ecology* (31:3), pp. 463--489.
8. Gadanecz, B., R. Moessner, and C. Upper (2007). Economic derivatives. BIS Quarterly Review.
9. Gurkaynak, R. and J. Wolfers (2006, January). Macroeconomic derivatives: An initial analysis of market-based macro forecasts, uncertainty, and risk. NBER Working Papers 11929, National Bureau of Economic Research, Inc.
10. Hahn, R. W. and P. C. Tetlock (2006). Introduction to information markets. In R. W. Hahn and P. C. Tetlock (Eds.), Information Markets: A New Way of Making Decisions, pp. 1–12. Washington D.C.: AEI Press. incollection George Introduction to Information Markets.
11. Madhavan, A. (1992, June). Trading mechanisms in securities markets. Journal of Finance 47(2), 607–41.

12. Mbemap, M. (2004). Economic derivatives and the art and science of investment. Journal ofDerivatives Accounting 1, v–viii.
13. McNees, S. K. (1992). How large are economic forecast errors? New England Economic Review (Jul), 25–42.
14. Osterloh, S. (2008). Accuracy and properties of german business cycle forecasts. Applied Economics Quarterly (formerly: Konjunkturpolitik) 54(1), 27–57.
15. Riedl, C., Blohm, I., Leimeister, J. and Krcmar, H. (2010). "Rating scales for collective intelligence in innovation communities: Why quick and easy decision making does not get it right" 'International Conference on Information Systems (ICIS), Saint Louis'.
16. Rosenbloom, E. S. and W. Notz (2006, feb). Statistical Tests of Real-Money versus Play-Money Prediction Markets. Electronic Markets 16(1), 63–69.
17. Schuh, S. (2001). An evaluation of recent macroeconomic forecast errors. New England Economic Review, 35–56.
18. Shiller, R. J. (1993). Macro markets: creating institutions for managing society's largest economic risks / Robert J. Shiller. Clarendon, Oxford.
19. Spann, M., Ernst, H., Skiera, B. and Soll, J. H. (2006). "Identification of Lead Users for Consumer Products via Virtual Stock Markets," *Journal of Product Innovation Management* (26(3)), pp. 322-335.
20. Stathel, S., S. Luckner, F. Teschner, C. Weinhardt, A. Reeson, and S. Whitten (2009, May). AKX - An Exchange for Predicting Water Dam Levels in Australia. In Proceedings of the 4th International Symposium on Information Technologies in Environmental Engineering, Thessaloniki, Greece, pp. 78–90.
21. Teschner, F., Mazarakis, A., Riordan, R. and Weinhardt, C. (2011). "Participation, Feedback & Incentives in a Competitive Forecasting Community". 'Proceedings of the International Conference on Information Systems (ICIS)', Shanghai, China, Paper 16.
22. Urban, G. and Von Hippel, E. (1988). "Lead user analyses for the development of new industrial products," *Management science*, pp. 569--582.
23. Von Hippel, E. (1986) "Lead users: a source of novel product concepts," *Management science*, pp. 791--805.
24. Wolfers, J. and E. Zitzewitz (2004, Spring). Prediction markets. Journal of Economic Perspectives 18(2), 107–126.
25. Wolfers, J. and E. Zitzewitz (2006). Interpreting prediction market prices as probabilities. Working Paper Series 2006-11, Federal Reserve Bank of San Francisco.
26. Zhang, J., Ackerman, M. and Adamic, L. (2007). "Expertise networks in online communities: structure and algorithms" In Proceedings of the 16th international conference on World Wide Web, ACM, pp. 221--230.
27. Zhang, S. S., Wagener, M., Storkenmaier, A. and Weinhardt, C. (2011). "The Quality of Electronic Markets" 'Proceedings of the Forty-Fourth Annual Hawaii International Conference on System Sciences', Koloa, Hawaii, pp. 1-10.