Kiel

Working Papers

Kiel Institute
for the World Economy

Identification of a Core-Periphery Structure Among Participants of a Business Climate Survey

by Ulrich Stolzenburg and Thomas Lux

No. 1659 | December 2010

Web: www.ifw-kiel.de

Kiel Working Paper 1659 | December 2010

Identification of a Core-Periphery Structure Among Participants of a Business Climate Survey

Ulrich Stolzenburg and Thomas Lux

Abstract:

Processes of social opinion formation might be dominated by a set of highly influential agents acting as 'opinion leaders'. Here we explore whether such a perspective could shed light on the dynamics of a well known economic sentiment index. To this end, we hypothesize that the respondents of the survey under investigation form a core-periphery network, and we identify those agents that define the core (in a discrete setting) or the proximity of each agent to the core (in a continuous setting). As it turns out, there is significant correlation between the so identified cores of different survey questions. Both the discrete and the continuous cores allow an almost perfect replication of the original series with a reduced data set of core members or weighted entries according to core proximity. Using a monthly time series on industrial production in Germany, we also compared experts' predictions with the real economic development. The core members identified in the discrete setting showed significantly better prediction capabilities than those agents assigned to the periphery of the network.

Keywords: Business climate, network topology, economic expectations

JEL classification: C42, C83, D85

Ulrich Stolzenburg
Department of Economics
University of Kiel
Olshausenstraße 40
24118 Kiel, Germany
E-mail: stolzenburg@economics.uni-kiel.de

Thomas Lux
Kiel Institute for the World Economy
24100 Kiel, Germany
Phone: +49 431-8814 278
E-mail: thomas.lux@ifw-kiel.de
E-mail: lux@bwl.uni-kiel.de

Identification of a Core-Periphery Structure Among Participants of a Business Climate Survey

An Investigation Based on the ZEW Survey Data[1]

Ulrich Stolzenburg
Department of Economics *Olshausen Str. 40, 24118 Kiel*
University of Kiel, Germany *E-Mail: stolzenburg@economics.uni-kiel.de*

Thomas Lux
Department of Economics *Olshausen Str. 40, 24118 Kiel*
University of Kiel, Germany *E-Mail: lux@bwl.uni-kiel.de*
and
Kiel Institute for the World Economy *Hindenburgufer 66, 24105 Kiel*

December 17, 2010

Abstract: Processes of social opinion formation might be dominated by a set of highly influential agents acting as 'opinion leaders'. Here we explore whether such a perspective could shed light on the dynamics of a well known economic sentiment index. To this end, we hypothesize that the respondents of the survey under investigation form a core-periphery network, and we identify those agents that define the core (in a discrete setting) or the proximity of each agent to the core (in a continuous setting). As it turns out, there is significant correlation between the so identified cores of different survey questions. Both the discrete and the continuous cores allow an almost perfect replication of the original series with a reduced data set of core members or weighted entries according to core proximity. Using a monthly time series on industrial production in Germany, we also compared experts' predictions with the real economic development. The core members identified in the discrete setting showed significantly better prediction capabilities than those agents assigned to the periphery of the network.

Keywords: business climate, network topology, economic expectations

JEL-codes: C42, C83, D85

[1] Financial support by the Volkswagen Foundation through their grant for the project 'Financial Markets as Complex Networks' is gratefully acknowledged. Thomas Lux also wishes to thank the Japanese Society for the Promotion of Science (JSPS) for financial support and the Graduate School for Social Sciences at ICU Tokyo for its hospitality at the time when the research reported in this paper was conducted.

1 Introduction

Expectations play a major role in determining economic outcomes. Data on economic expectations are usually collected in the form of surveys of so-called consumer or producer sentiment. Due to the complexity of formulating a forecast concerning such abstract objects like GDP (gross domestic product) or inflation rates, such surveys are typically constructed in a qualitative way: Agents are given a limited set of possible answers such as 'good', 'bad' or 'equal' for their assessment of the future development of some economic quantity.

Despite the widespread attention devoted to the often monthly publication of the results of such surveys, economic literature on these data somewhat surprisingly is not that voluminous. One important reason for this neglect is that modelling of expectations has for a long time been dominated by the paradigm of so-called 'rational expectations' (Miller, 1994). This modelling device assumes economic agents formulate expectations in harmony with the underlying economic model. As a consequence, a model builder would try hard in order to close his model deriving model-consistent expectations. In this way, no role was left for expectations as a driving factor *sui generis* which explains a widespread lack of interest in data on expectations. The empirical analysis of expectation data that emerged from the rational expectation school was basically self-referential: Its aim was to test the rational expectations hypothesis. While the typical outcome of such analyses was that rational expectations could be overwhelmingly rejected for survey data (Nardo, 2003) this did little damage to the insistence on rational expectations as a key building block of macroeconomic models.

The 'rational expectation' revolution of the 70s and 80s has, however, meanwhile triggered a counter-revolution that emphasizes once abandoned alternative concepts such as 'animal spirits' (Akerlof and Shiller, 2009). In this view (that has also characterized the then dominating Keynesian economics up to the 70s), expectations of economic agents are not the mere, rationally computed reflection of objective probabilities of future events. Rather, economic expectations are shaped by the socio-economic climate. An optimistic or pessimistic disposition among the population might emerge endogenously and be propagated via direct or indirect channels of communication. Under this perspective, it is not so much external events impacting on the economic system but a collective process of opinion formation that governs the evolution of expectations on future economic conditions. Waves of optimism or pessimism might then leave their imprint on real economic activity as deeds (to consume or not consume, to invest or not invest) would follow the psychological state of consumers and businessmen. This change of perspective also triggered a new interest in behavioral theories of expectation and opinion formation in economics (see, for example, Branch, 2004; Carroll, 2003; Hohnisch et al., 2005; Lux, 2009; Menkhoff, Rebitzky and Schroeder, 2009; Westerhoff, 2010). Within this emerging literature , we ask a question that has to our knowledge never been explored for economic expectation data before: Can we identify a subset of agents that act as opinion leaders?

Processes of opinion formation are very complex. Obviously, tendencies of herd behavior are present when some opinion leaders are able to influence other people or even broad masses of the population. These opinion leaders can be politicians, journalists of important newspapers, TV programs, other media or private persons that like to share their views with their colleagues and friends. Opinion leadership and some degree of herd behavior are present not only in the formation of political views, opinions about rock stars, sports, fashion, lifestyle trends and the need of consumer electronics (cf. Casti, 2010). It might also be prevalent in the expectation formation process with respect to economic categories like asset prices and the business climate expectations. Data of one of the leading business climate surveys in Germany will be scrutinized under this perspective.

The data used in our investigation is the complete set of individual responses to the so-called ZEW Financial Market Test. Approximately 330 financial experts are asked each month to predict the future development of a number of economic indicators. Our aim is to apply a technique designed to identify a core-periphery network structure among the participating experts. By identifying a core of experts among the participants, we might be able to isolate a group of respondents that acts as opinion leaders in this monthly survey.

From Borgatti and Everett (1999) we learn that there are different intuitive interpretations for a core-periphery network structure: First, there may exist two groups of experts, core members and periphery members. Each person definitely belongs to one of these groups and can be assigned either a 1 or a 0 indicating core or periphery membership. Later this will be referred to as the discrete model. Second, agents might be characterized by a continuous measure of influence or proximity to the core. Instead of the binary dichotomy of the discrete model, in this framework, each individual might be assigned a real number that identifies its core proximity, which Borgatti and Everett denote as 'coreness'. This model will be referred to as the continuous model. Both concepts will be applied to the ZEW data set, and we will also explore how much overlap there is between the cores identified for different categories of questions.

The rest of the paper is organized as follows: In part 2, the origin and structure of the underlying dataset is briefly explained. Part 3 contains a description of the discrete model followed by its results, part 4 covers the continuous model and its results. In part 5, the results of the network analysis are used as weighting schemes to construct a new business climate indicator based on the identified core membership of the respondents. Part 6 uses data on real economic activity to assess the predictive performance of individual experts and compares the average predictive success of identified core and periphery members. Part 7 concludes.

2 Dataset

2.1 The ZEW Financial Market Test

Each month 300 to 350 financial experts take part in the survey called ZEW Financial Market Test conducted by the Center for European Economic Research (German acronym: ZEW) at the University of Mannheim. Participating experts are asked to assess the current economic situation and to predict the development of key economic indicators for Germany and a few other important economies like France, Italy, Japan, USA and the United Kingdom. The questionnaire consists of a number of assessments for the development of business climate, interest rates, inflation, stock market indices, commodity prices, profitability in various sectors and others. The most prominent ZEW indicator of future business climate is the six-month expectation for the overall economic situation in Germany. Survey participants are given three answer categories for their assessment plus the additional "I don't know"-option. In six months, the economic situation may improve (+), stay the same (0) or deteriorate (-). Depending on the type of economic indicator, the three given categories can also be good (+), normal (0) and bad (-) instead, e.g. for the question on respondents' view of the current state of the economy.

The published ZEW Indicator reflects the difference between the percentage of optimists and the percentage of pessimists. So if all experts were optimistic, a maximum score of 100 would be the result. If all experts were pessimistic, the indicator would show a score of minus 100. Therefore, the domain of the ZEW Indicator extends from -100 to +100 with an average score of approximately 30 points between 12/1991 and 03/2008. The development of the ZEW Indicator can be observed in Figure 6 (blue line).

Our data cover the period December 1991 - March 2008, so that we have a record of 196 monthly observations. 196 months, multiplied by more than 300 participants each month sum up to more than 60,000 filled-in questionnaires. None of the experts participated in all of the 196 months, so we have incomplete time series. A total number of 1310 different individuals participated at least once.

In our analysis we focused on some selected survey questions that should be most relevant for the group of German financial experts participating in the survey. Table 1 lists these issues together with the original question numbers of the ZEW survey on the left side.

2.2 Dataset Reduction

Since we are interested in network structures, we need frequently participating experts in order to be able to detect similarities and specifics in their answers. So we filtered

1	Business climate, current situation, Germany
7	Business climate, 6 month expectation, Germany
13	Inflation, Germany
19	Short run interest rates, Germany
25	Long run interest rates, Germany
31	DAX
32	DOW JONES
37	Exchange rate Euro - US$

Table 1: Selected Survey Questions

out those experts who participated less than (an arbitrary threshold of) 60 times. After eliminating all less frequently participating individuals, we are left with 372 remaining experts, whose participation patterns are depicted as dots in Figure 1. The vertical axis depicts all remaining experts of the reduced dataset, while on the horizontal axis the 196 months are shown. Now, the remaining 372 experts with a total of 48,915 questionnaires have an average number of 131 participation months.

From Figure 1 we are able to roughly distinguish three groups of experts: Those who participated approximately up to the year 2000 in the upper part of the figure, those in the lower part who participated mainly from 2000 up to 2008, and those who participated during the whole time series from 1991 to 2008. Groups one and two have almost no temporal overlap regarding survey participation.

As a closeness or association measure we defined the number of identical predictions, divided by the absolute number of overlapping survey participations. This measure will always be bounded between 0 and 1. An additional condition of at least 20 overlapping participations was applied to make sure all association measures that enter our subsequent analyses have a sufficient basis. The rate of identical assessments is interpreted as the connection weight between two persons in the network.

For our core-periphery models we assume that members of the core should have a close connection (high correlation of given answers), that the connection of core-members with periphery-members should be lower (medium correlation of answers) and that periphery-members should have a weak connection among each other (relatively low correlation in survey answers). Remember that this association matrix can be generated for each of the 62 different survey questions, so for each question our analysis may result in a unique network structure.

Figure 2 shows the associations between 372 rather frequently participating individuals. As usual, our association matrix is symmetric, since our association measure has no direction. The right side shows the color code for the strength of the connection weights, which are defined as the shares of identical assessments. Note that each agent has a

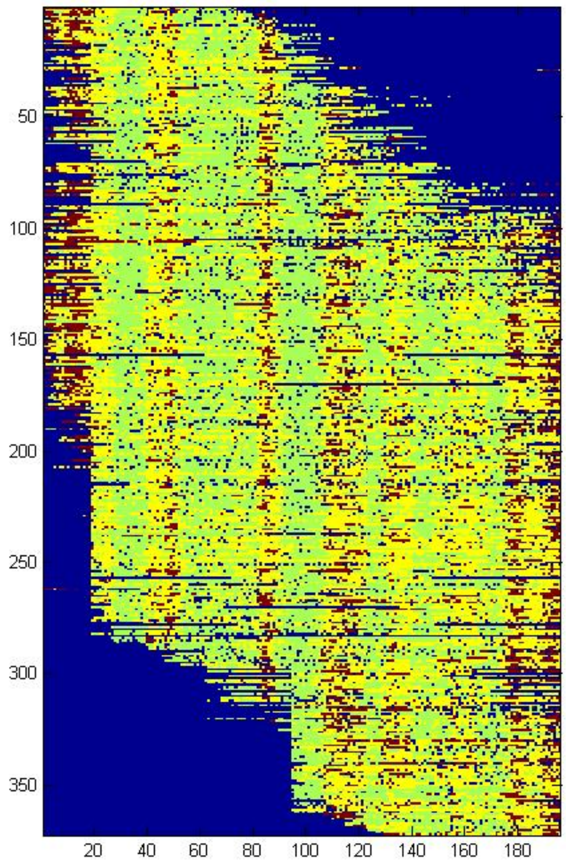


Figure 1: Responses patterns of those experts with more than 60 active participations in the survey, a total of 372 experts, sorted by participation time (horizontal axis: survey months). Blue areas show months with missing responses of the respective expert. Using survey question 7 on business climate expectations we added more information to this Figure: Green color indicates predictions of an improvement of the economic situation (+), yellow color indicates predictions of stable business conditions (0) and red color identifies experts with a pessimistic prediction (-).

correlation of exactly 1 with his or her own survey answers, so we observe dark red dots on the main diagonal.

In Figure 1 we clearly see the two groups of experts without a temporal overlap in their survey participations. The blue areas in the lower left and in the upper right corner of Figure 2 indicate that our association measure is not available for these segments of our reduced dataset. That is why we focused our network analysis on the middle part of all frequently participating experts. Thus, after reducing the dataset further we are left with 186 experts participating more or less regularly over the whole time span

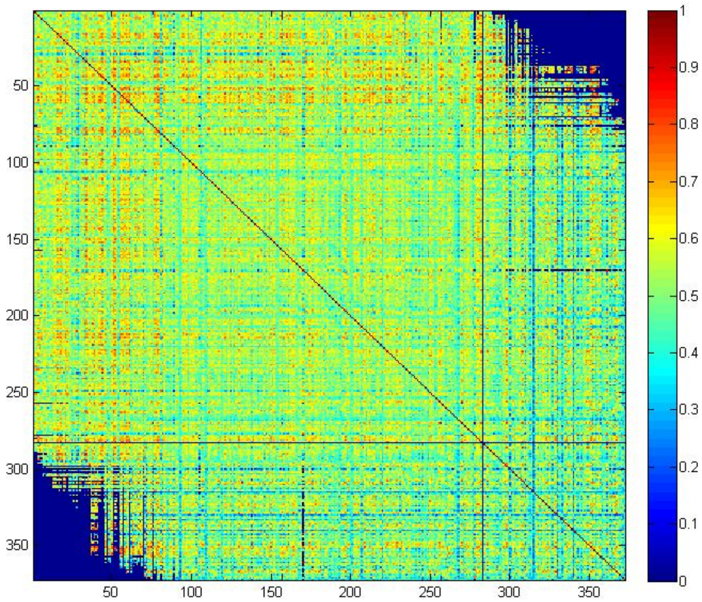


Figure 2: Connection between experts in reduced dataset

between 1991 and 2008. Since the primary goal of this paper is an application of a network identification method, such a dataset reduction does not reduce the validity of our study, but excludes cases that due to their infrequent or non-overlapping activity would be unsuitable for our analysis.

3 Discrete Model

3.1 Optimization Problem

We follow the methodology proposed by Borgetti and Everett (1999) to identify the core-periphery structure of our data set. In the discrete model each of the 186 experts is assigned either a value of 1 or 0. Core members are identified by a 1, periphery members by a 0. Two core members should have a relatively strong connection (high association measure), other combinations (core-periphery, periphery-periphery) a relatively weak association. Now, by multiplying the binary string of core memberships with its own transpose, we determine the so-called pattern matrix. It will carry ones only at core-core positions and zeros elsewhere. We search for the binary string of core membership that has a structure as similar as possible to the matrix of identical assessment shares (data matrix).

Let c_i be the core membership of the expert at position i of the binary string. Then the pattern matrix $\boldsymbol{P}$ is determined by multiplying the whole vector $\boldsymbol{c}$ by $\boldsymbol{c}'$. Thus, elements of the pattern matrix δ_{ij} are given by

$$\delta_{ij} = \begin{cases} 1 & \text{if } c_i = 1 \quad \text{and } c_j = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Now our goal is to choose the vector c such that the resulting pattern matrix $\boldsymbol{P}$ (with elements δ_{ij}) has a similar structure as the data matrix $\boldsymbol{A}$ (with elements a_{ij}). Structural similarity is measured by the so-called 'matrix correlation'. For this purpose, we define 'matrix correlation' as the correlation between the two vectors, that are constructed by simply stacking all matrix columns. Now the optimization problem is to choose a vector $\boldsymbol{c}$ such that a pattern matrix is constructed, that has a maximum matrix correlation with the data matrix $\boldsymbol{A}$.

$$\max_{c} Corr(vec(\boldsymbol{A}), vec(\boldsymbol{P})). \tag{2}$$

Note that core membership does not necessarily give evidence whatsoever of a better quality in the survey answers. Since the criterion for becoming a member of the core relies on a similar patterns of answers of a group of experts, a person that follows the majority opinion in his survey answers is more likely to be detected as a core member than a person with a more idiosyncratic assessment of the economic outlook. In this way, we may identify some sort of core or opinion leaders of agents rather than persons with outstanding economic expertise.

3.2 Optimization Method

A modified genetic algorithm has been used to find the optimal binary string. As objective function of this algorithm we use the matrix correlation between the data matrix and the pattern matrix.

Our genetic algorithm is initialized by randomly generating an initial population of N genes (binary strings of 186 digits). Most of the time we used a relatively small population of $N = 20$ genes. Each of the binary strings has a certain level of 'fitness', which is the value of the objective function given in eq. 2. Now we evolve the population using the standard genetic algorithm operators of 'reproduction', 'mutation', 'crossover', and 'election'.[2] In order to accelerate the computations, we added another operation that we christened 'genetic engineering' in the context of this algorithm.

During the first step of 'reproduction', the old population is evolved by allowing each of the binary strings to enter the new population with a probability that is proportional to

[2] The exact design of the genetic operations follows the scheme presented in `http://www.bwl.uni-kiel.de/vwlinstitute/gwif/files/handouts/abmef_alt/VGeneticLearningin\%20EconomicModls120706.pdf`. Pertinent GAUSS code can be found at `http://www.bwl.uni-kiel.de/vwlinstitute/gwif/downloads_handoutsv.php?lang=de`.

its fitness value: The higher the genes' fitness, the higher the probability of entering the next generation. With a rather low probability, some of the binary digits of a string are flipped, so that genes are evolved by a 'mutation' step into a new population. Another step in the process is the so-called 'crossover operator'. Two binary strings are cracked at a random point, and the first part of string A is combined with the second part of string B and vice versa to create the offspring of the pertinent 'parents'. This way, we are left with two new candidates for the new population. In the economics learning literature, often the standard set of genetic operators has been expanded by including a so-called 'election' operator (cf. Arifovic, 1996; Lux and Schornstein, 2005). We also follow this practice. The 'election' operator implies backtesting of these candidates prior to admission to the new population. In particular, they are only admitted to the population if their fitness is at least as high as that of one of their parents. Otherwise, the parents are copied rather than their offspring. In this way we avoid the population being invaded by inferior candidates.

Our additional operation called 'genetic engineering' accelerates the convergence of the algorithm. We select the binary string with highest fitness value of the current population so far. Then we try to improve its fitness value by eliminating one core member and by replacing it with another. The core member to eliminate should be one with a relatively low connection to its fellow core members, while the one replacing it should be a periphery member with a relatively high association to all other core members. To this end we compute the connections of all current core members with all other core members and determine the mean of these values. We also pick each periphery member and compute the connection with all core members and again determine the mean. Then we can identify the core member with the lowest average connection to all other core members and flip its core membership value from 1 to 0. Also, we locate the periphery member with the highest average share of associations to all core members hitherto and flip its core membership value from 0 to 1. Afterwards, we put the resulting string back into the evolving population. By using this additional operation we can manipulate our population target-oriented and do not have to rely solely on blind exploration of the search space by the random evolutionary steps of 'mutation' and 'crossover'.

3.3 Results

Figure 3 shows the same correlation matrix as in Figure 1 with two differences. First, it shows only the central square of Figure 1, because we restrict our network analysis to those 186 experts (numbered 95 to 280 in Figures 1 and 2) with strong temporal overlap in survey participations. Second, a rearrangement of experts with respect to core membership has been conducted to visualize the network structure. The yellowish square in the upper left corner indicates that the corresponding individuals have a fairly higher association among each other than with the rest of the individuals. A core-periphery network structure appears clearly visible in this rearranged association matrix.

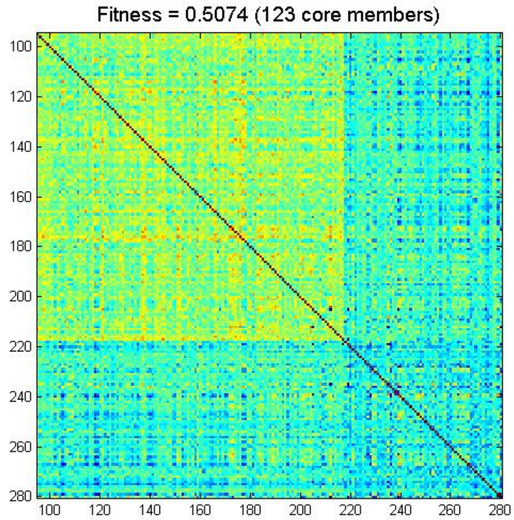
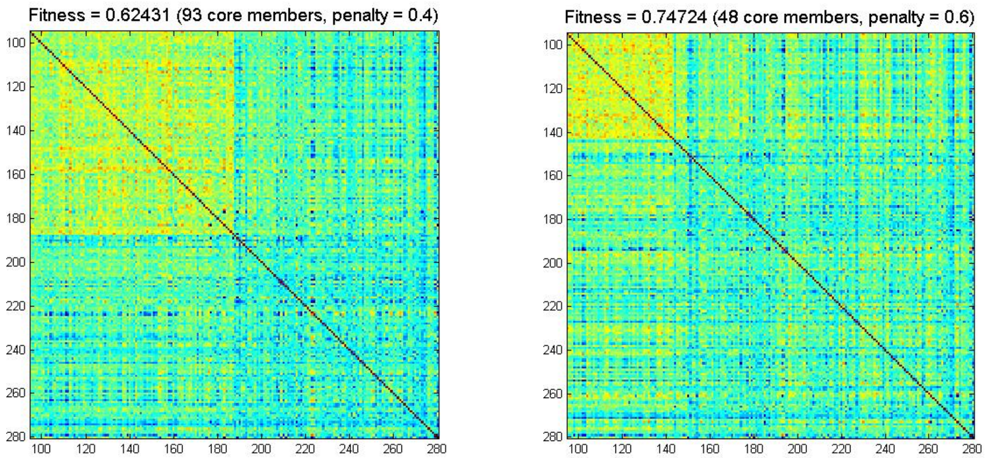


Figure 3: Binary core-periphery structure for responses concerning the expected economic situation in Germany in 6 months (Survey question 7)

A potentially interesting extension of the discrete model consists in using a penalty function in order to limit the size of the core. Note that in our previous analysis the number of experts entering the core was completely open and determined in a data-driven way. One might be interested in a core of only 50 out of 186 agents, while our optimization in Figure 3 resulted in a core of 123 out of 186 agents. The number of agents entering the core can be influenced if we add a penalty factor for each additional agent that enters the core. The genetic algorithms' objective function is then simply expanded by this factor.[3] A positive (negative) penalty factor will result in a core with fewer (more) core members. Figures 4(a) and 4(b) show results of such an experiment with a moderate and a relatively high penalty factor. Instead of 123 members, the core consists of only 93 members (a) and 48 members (b). Since we lack clearly defined criteria for a target size of the core, we leave a more systematic investigation of the effect of penalty functions to future research.

[3] Technically, this can be done by simply multiplying the correlation coefficient of the fitness function by the term $\left(\frac{N}{N_{core}}\right)^{\rho}$. N is the absolute number of agents (in our case 186), N_{core} is the number of core members in a certain core-periphery subdivision and ρ is the chosen penalty factor (or penalty exponent).



(a) (b)

Figure 4: Binary core-periphery structure for responses concerning the expected economic situation in Germany in 6 months (Survey question 7) with a moderate (a) and a high penalty factor (b) for additional core members.

3.4 Correlations between Survey Questions

We conducted the above analysis (without a penalty factor) for different survey questions as listed in Table 1. For each of them we get as a result one binary vector of core membership values. If the detected core members really constitute something like the group of main opinion leaders of survey participants, we may expect the algorithm to find approximately the same core members, no matter what survey question were investigated. So a comparison of optimized coreness vectors for different questions has been conducted. Instead of counting concurrent core and periphery members, we computed simple correlation coefficients between both vectors, because later this procedure can also be applied to the continuous model. Coefficients are listed in Table 2.

First, we observe that almost all values are positive, which is supportive of the idea that we were able to find some sort of true network structure among the agents. It is, however, not clear a priori whether these numbers are in any sense significant or not. Note that if the core consists of exactly fifty percent of the population on average, independent random distributions of core membership across survey questions would be characterized by equally positive and negative correlations. However, our cores are data-driven and often include more than half of the underlying population. In this case, the expected correlation between core membership for different questions might, in fact, be positive even if the pertinent core members were independently drawn in both cases.

To check whether significant correlations prevail between the optimized core and pe-

Qu.	1	7	13	19	25	31	37
1	1	**0.303****	**0.219****	0.093	**0.231****	0.058	**0.193***
7		1	**0.313****	**0.175***	**0.329****	**0.168***	0.113
13			1	**0.156***	**0.174***	0.050	-0.131
19				1	0.063	0.110	**0.278****
25					1	**0.301****	**0.155***
31						1	**0.208***
37							1

(*: Close to critical values of bootstrap distribution; roughly significant at 5%)
(**: Far beyond critical values; highly significant)

Table 2: Correlation coefficients of optimized core membership vectors between different survey questions.

riphery memberships of different survey questions, we apply a bootstrap test. For each pair of survey questions we reshuffle the ones and zeros in both vectors and determine the correlation of both vectors. This procedure was repeated 1000 times, so that we obtained a distribution of bootstrap correlations. The true correlation can be compared with this bootstrap distribution. By running this bootstrap exercise, we are able to assess the significance of correlation coefficients. In most cases, we can conclude that the correlation of coreness vectors among different survey questions is significantly positive, even if the magnitude of this association seems not excessively large.

4 Continuous Model

4.1 Optimization Problem

In the continuous model, individuals are characterized by different levels of 'coreness'. Each expert belongs to the group of opinion leaders to a greater or lesser extent. Each of the 186 experts is assigned a real number to identify his degree of core proximity or 'coreness'. Apparently, two experts with high values are expected to have a strong connection (high correlation in survey answers), while experts with low values should have a rather weak connection to each other.

Again, by multiplying the vector of coreness values $\boldsymbol{c}$ with its own transpose, we get to the pattern matrix $\boldsymbol{P} = \boldsymbol{c} \cdot \boldsymbol{c}'$, exactly as in the discrete model. There is now a continuum of coreness values across respondents. Figure 5(a) shows an example of a pattern matrix, resulting from the opimization process. By sorting all individual experts for their coreness values, we are left with a continuum of expected connection weights. Experts with high values in the upper-left corner have by far a higher expected connection with each other than two experts with low values. We search for the vector of optimal values that serves best to resemble the true network structure. As objective function we use the matrix correlation between the data matrix and the pattern matrix; and the pattern matrix should have a structure as close as possible to the matrix of identical assessment shares (data matrix). Our optimization problem is (again):

$$\max_{c} Corr\big(vec(\boldsymbol{A}), vec(\boldsymbol{P})\big) \tag{3}$$

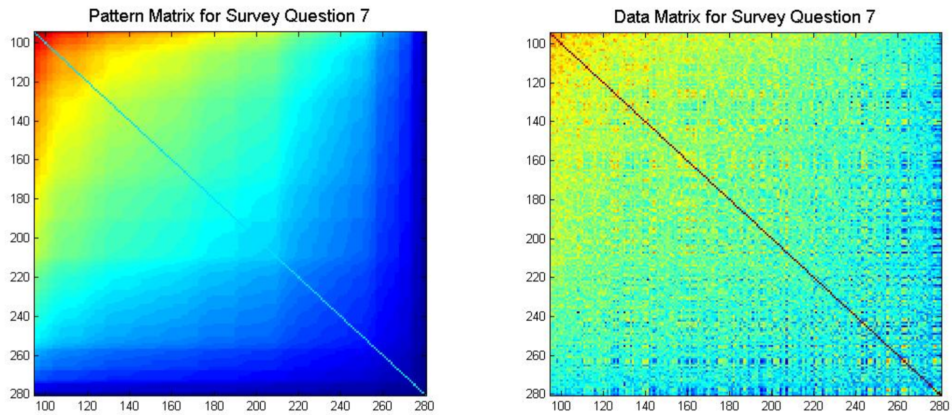
Note that all values in vector c may assume real rather than binary values now, so the optimization problem is far more complex than in the discrete case. Accordingly, computation for this optimization problem is far more time-consuming. Note again that a high coreness value does not necessarily give evidence whatsoever of a better quality in the survey answers. Since the criterion for becoming a member of the core relies on a smilar answering pattern of a group of experts, a person that follows the majority opinion in his survey answers is more likely to be detected as a core member than a person with more independent expectations.

4.2 Optimization Method

We used the algorithm of Nelder-Mead to solve our multi-dimensional maximization problem (cf. Press et al., 2007, chap. 10). In our setting with 186 experts it is able to find the optimal coreness vector in a workable time span.

Since the objective function is a scale-free correlation coefficient it is only the relative magnitude of coreness values that matters. Different starting vectors will, therefore, lead to different sets of optimal values, albeit with the same structure of relative coreness values (provided there is a unique solution to our optimization problem). In repeated runs, we indeed found results that were very close in relative coreness. To make optimal vectors of different survey questions comparable, we normalize all values by division by the mean. Moreover, we restricted the vector of coreness values to positive real numbers, because otherwise an optimal vector with many negative values (and therefore lower mean) would result in a nomalized vector with very high variance, which renders normalized coreness values incomparable again. Hence comparability suggests a restriction of the model to positive values.

We also ran our continuous model in a different setting without the restriction to positive values. This allows for two groups instead of only one group, members of the main crowd and antagonistic agents, that share a strong connection in each group, but have a weak connection or negative correlation to members of the other group. A pattern matrix in that case would be characterized by positive values also in the lower right corner, because negative coreness values of the second group are multiplied with other negative values and result in positive values again. In almost all of these experiments less than 10 experts out of 186 were assigned slightly negative values by our algorithm. In other words, the data rather reject the idea of two antagonistic groups of experts with high associations among themselves and low associations with members of the other group. Moreover, comparability with our discrete models results is another advantage of the model restricted to positive values. So we decide not to allow for negative coreness values in our main analysis. It indeed appears implausible in our present setting that there should exist two well-defined groups with persistently antagonistic views on the future development of economic variables among the respondents.



(a) (b)

Figure 5: Pattern matrix(a) and data matrix(b), sorted for coreness values.
Survey question 7: Economic situation in 6 months in Germany.

4.3 Results for the Continuous Model

Analogously to the discrete case, Figure 5(b) shows the same association matrix as in Figure 2 with two differences. First, it again shows only the central square, because we restricted our network analysis to only 186 experts, numbered 95 to 280 in Figures 1 and 2, with strong temporal overlap in survey participations. Second, a rearrangement of experts with respect to coreness values has been conducted to visualize the network structure. Yellowish and light green areas in the upper left corner indicate that corresponding individuals have a fairly higher association among each other than with the individuals below. A core-periphery network structure seems again clearly detectable in this rearranged association matrix by visual inspection. Moreover, the similarity with the corresponding pattern matrix in Figure 5(a) is also observable.

Similar results can be obtained for different survey questions. For reasons of space and redundancy the respective figures are not displayed in this paper.

4.4 Correlation between Survey Questions

We conducted again the above analysis for the different survey questions as listed in Table 1. If the detected values really indicate some kind of robust core of financial experts, we would expect the algorithm to find approximately the same relative coreness values, no matter what question was asked. Correlation coefficients of coreness vectors for different questions are listed in Table 3.

Qu.	1	7	13	19	25	31	37
1	1	**0.351****	**0.268****	**0.256****	**0.280****	0.138	0.042
7		1	**0.464****	**0.273***	**0.509****	**0.310***	0.122
13			1	**0.386***	**0.315***	0.130	-0.046
19				1	**0.299***	0.122	**0.211****
25					1	**0.357****	0.103
31						1	**0.230***
37							1

(*: Close to critical values of bootstrap distribution; roughly significant at 5%)
(**: Far beyond critical values; highly significant)

Table 3: Correlation coefficients of optimized coreness vectors between different survey questions.

Again, we observe that almost all values are positive, which is supportive of the idea that we were able to find some sort of true network structure among the agents. Note also how close the magnitudes of the correlations are to those reported for the discrete model in Table 2 throughout, although the later are computed from a completely different set of binary (0-1) coreness vectors. To check whether optimized core and periphery memberships are stable between different survey questions, we apply again the same bootstrap test. Again the values of both vectors are shuffled and the correlation coefficient of both reshuffled vectors is determined. This procedure has been repeated 1000 times, so that we obtained a distribution of bootstrap correlations. The true correlation was compared with the bootstrap distribution. By running this bootstrap exercise, we are able to assess the significance of correlation coefficients. In most cases, we can again conclude that the correlation of coreness vectors among different survey questions is significantly positive. The complete pattern of results is in very close correspondence to those reported in Table 2.

5 Reconstructing the ZEW Index with Core Agents

5.1 Construction

The popular ZEW Indicator of Economic Sentiment is based solely on survey question 7, which asks for the economic situation in 6 months and allows improvement (+), an unchanged situation (0) and deterioration (-) as answer categories. Defining experts who expect an improvement of the economic situation as optimists, those who expect a deterioration of the economic situation as pessimists and those expect the current economic situation to remain stable during the next 6 months as neutral agents, one computes the index as a diffusion index, i.e. fraction of optimistic minus fraction of pessimistic agents. Results on the predictive capacity of the index can be found in Breitung and Jagodzinski (2001) and Hüfner and Schröder (2002). Nolte and Pohlmeier (2007) compare the performance of predictions based on the ZEW index with forecasts of linear times series models and random walk forecasts and found no evidence of superiority of the aggreagte survery forecasts.

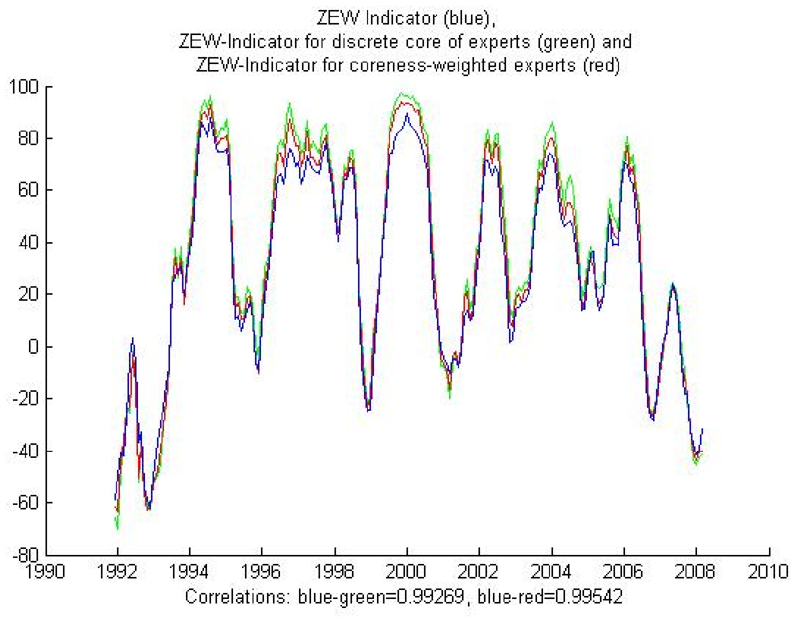


Figure 6: ZEW Indicator and weighted indicators from network analysis.

5.2 Weighted Indicator from Reduced/Transformed Dataset

A share of optimists and pessimists can be computed for a reduced dataset as well, so we are able to compute another hypothetical business climate indicator with a reduced set of core agents. Moreover, we can also use results from our continuous model to create a weighting scheme for the ZEW Indicator. By using only the core members from the discrete models' optimization result, time series of optimists and pessimists have been generated and another indicator was constructed (green line). By using coreness values from the continuous model as weights, we are able to generate a third indicator, which is depicted as red line in Figure 6.

For the indicator weighted by the discrete optimal core, only about one third of all available questionnaires have been included in the generation of this curve, but it still has a correlation with the original ZEW indicator of 0.993. The coreness weighted indicator has a correlation of 0.995 with the original ZEW Indicator and uses about one half of all available questionnaires. So we see that the detected core of experts carries almost exactly the same information as the complete dataset. In a sense this shows that there is quite some redundancy in the complete record of responses, although the ZEW has a smaller number of respondents than some other popular surveys. This redundancy by itself suggests some kind of correlation among the individual participants and our above results strongly support some kind of hierarchical core-periphery structure.

We conducted a bootstrap exercise to investigate whether the core is really different from the periphery in its proximity to the aggregate forecasts of the complete set of respondents. The design of an informative test for this question is a relatively delicate task. One obstacle, of course, is the large number of missing data for individual agents that makes it impossible to track the performance of any core or periphery member over all periods. Another complication is that our data-driven cores are mostly larger than the periphery. As a consequence, the core will trivially carry a higher weight in the aggregation than the smaller periphery. Hence, the average of the core opinions should automatically be closer to the overall aggregate than the average of the periphery. To neutralize the influence of the number of core members, we proceed in the following way: Instead of using the complete core from our optimization, we picked a random sample of 40 individuals from the group of core members (with replacement) and computed a new indicator with this random combination of experts. Similarly, we picked a random combination of 40 individuals from the group of periphery members and constructed another indicator. Correlation coefficients of both constructed indicators with the original ZEW Indicator were computed to investigate whether they show a similar pattern as the ZEW Indicator.

Figure 7 shows histograms of these correlations from 1000 bootstrap repetitions. We found that core members are significantly better than periphery members in replicating the ZEW Indicators' temporal pattern. We performed a t-test on the equality of means of both correlation samples.[4] As expected, the null hypothesis of equal mean values could be strongly rejected. The significantly higher correlation of indicators constructed from randomly sampled core members confirms that the identified core of

[4] We used a one-sided mean difference t-test for independent samples with unequal variances. The null hypothesis is H0: equality of means, which is tested against the alternative H1: Correlations of core members' indicators (red bars in figure 7) have a higher mean value (one-sided test). H0 was strongly rejected with a t-statistic beyond 60.

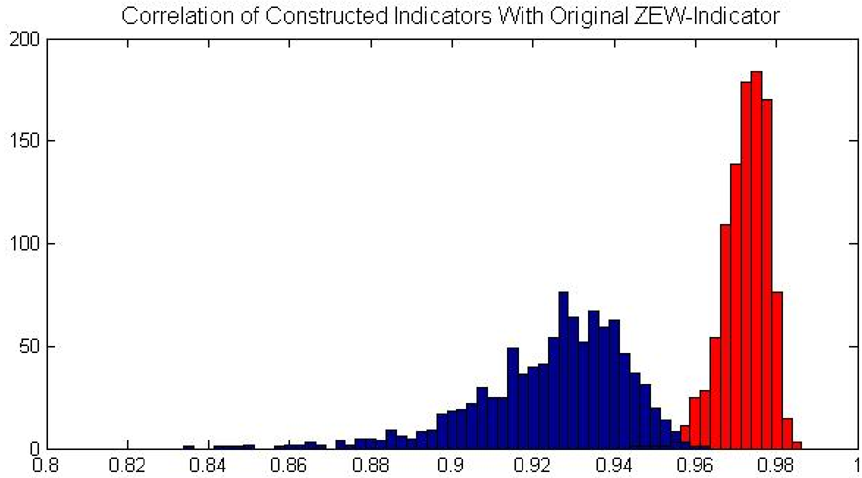


Figure 7: Bootstrap distributions as histograms of the correlation between the original ZEW Indicator and constructed indicators based on random samples of core (red) and periphery experts (blue).

experts is closest to the information of the Original ZEW Indicator, while the opinions of periphery members contain more unsystematc noise.[5]

6 Forecast Value of Core and Periphery

We are also interested in the predictive power of economic sentiment indicators. Since the ZEW Indicator reflects expectations about Germany's overall economic situation, for comparison we need a measure of real economic activity. This measure should have the same (monthly) frequency as the ZEW Indicator in order to assess its predictive power properly. Growth rates of gross national product are published quarterly, so instead of it we use total industrial production as a proxy for economic activity. For this purpose we use seasonally adjusted time series available in the OECD online database.[6]

Now since an assessment of the overall economic situation (business climate) does not refer to an absolute level of production but to economic dynamics, i.e. the change of production, we need to transform our total industrial production index into some sort of economic growth measure. Let $Prod_t$ denote the index of total industrial production. Then $\Delta Prod_t$ is a one-month difference of our production index from $t-1$ to t. Since all industrial production index values have an order of magnitude around 100, simple

[5] This result is also stable if we take random draws from core and periphery subsets that have been forced to be of equal size. For example, in Figure 4(a) experts are divided into 93 core and 93 periphery members. Our bootstrap exercise for this case showed the same pattern as in Figure 7 and the t-statistic of the mean difference test had a value of about 60.

[6] The index of total industrial production can be downloaded from http://stats.oecd.org.

differences can be considered as an approximation for percentage changes. The relevant survey question predicts the economic situation in six months. Now we have to decide what kind of measure $\hat{Y}_{t,t+6}$ should be used to best assess the quality of our prediction, that is, to compare our sentiment index with. The most plausible value for comparison appears to be the six-month difference of total industrial production. Thus we use for the evaluation of the agents' forecasts the sign of

$$\hat{Y}_{t,t+6} = Prod_{t+6} - Prod_t = \sum_{s=1}^{6} \Delta Prod_{t+s} \tag{4}$$

Figure 8 shows the pattern of our real economic measure defined by equation 4. We added a vertical line at 0% to split all values of $\hat{Y}_{t,t+6}$ into the categories improvement (+) and deterioration (-), that survey participants were able to choose from. A third category for the stable, unchanged situation (0) with more or less arbitrary threshold values could also be included, but this would require us to come up with a definition of a neutral interval. To avoid this complication, we mainly restrict our attention to the division only into positive ($\hat{Y}_{t,t+6} > 0$) and negative values ($\hat{Y}_{t,t+6} \leq 0$). Now we are able to compare predictions and real economic outcome of individual survey participants.

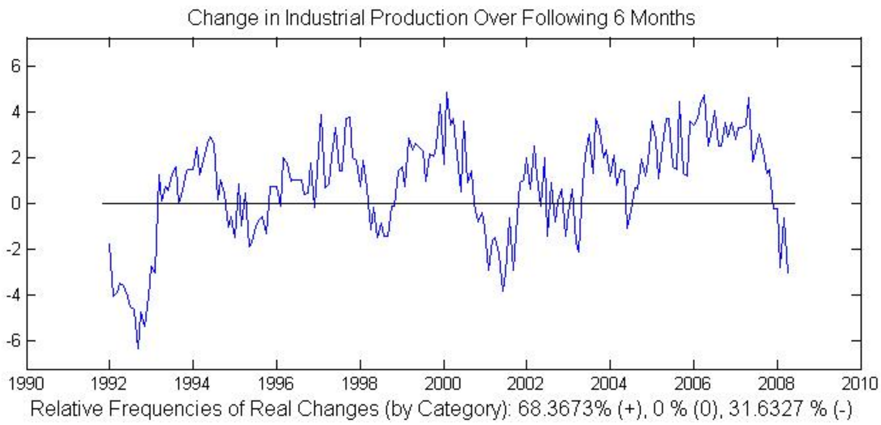
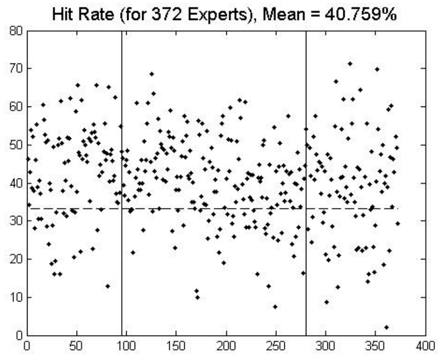
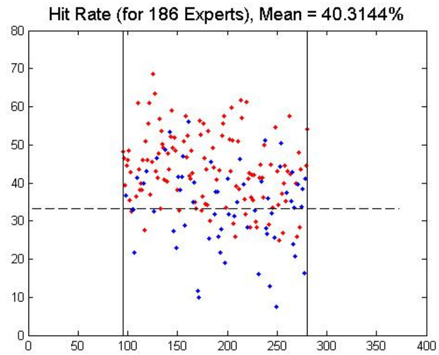


Figure 8: Six-month difference of total industrial production in Germany ($\hat{Y}_{t,t+6}$) as a measure of real economic progress during the following six months, 1992-2008.

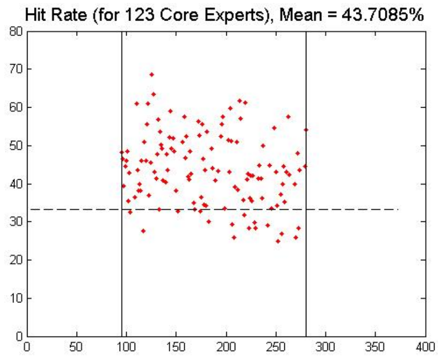
Let the number of successful predictions of improvements(+) or declines(-) divided by that experts total number of survey participations be defined as the personal prediction hit rate. Figure 9(a) shows individual hit rates of prediction for all experts in the reduced dataset, sorted for the time of their survey participations. Two vertical lines identify the middle part of experts that participated over the whole survey timespan. A dashed horizontal line at a hit rate of 33.3% is also included for the sake of orientation, because if agents were simply rolling dice to predict the future development in three categories, that hit rate should be reached on average. Note that in our dataset, experts were choosing the option of an unchanged economic situation (0) in about 43.5% of all cases. All these stability predictions are doomed to be wrong in a setting with only two alternatives for the real economic development. Therefore, an *average* prediction hit rate of about 56.5% is the maximum to be achieved if an agent were correct in all her non-zero predictions.



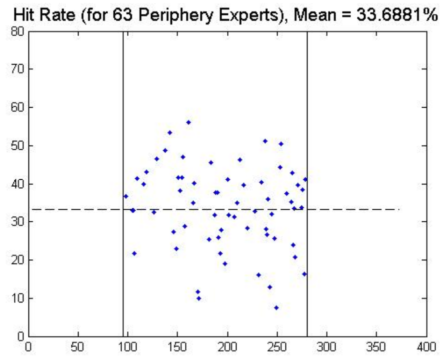
(a)



(b)



(c)



(d)

Figure 9: Prediction hit rates of individual agents, mean value of hit rates in title. Groups: All 372 agents of reduced data set (a), only 186 experts with continuous participation (b), 123 core members only (c) and 63 periphery members only (d).

The titles of Figures 9(a) to 9(d) include the average hit rate of predictions for a certain subgroup of survey participants. Figures 9(a) and 9(b) show that average hit rates for the whole reduced dataset of 372 experts and the middle part of 186 experts have almost the same value. Figures 9(c) and 9(d) show that average hit rates for core members of the discrete model (see Figure 3) are significantly higher than those of periphery members, with 43.7% for core members and 33.7% for periphery members. It is remarkable that all experts with a very poor prediction record (hit rate below 25%) have been identified by our algorithm as periphery members. Again, a one-sided test for mean difference of both sets of hit rates shows that the difference of means is highly significant with a t-statistic of 6.7. Relating these average hit rates to the maximum average hit rate of about 56.5% in our setting, the difference between core and periphery members is even more impressive.[7]

This exercise of comparing average hit rates of core and periphery members can also be conducted with smaller core sizes. Using the smaller core from Figure 4(a), identified with a discrete optimization including a penalty factor for additional core members, 93 core members have an average hit rate of 45.2%, compared with 35.4% for 93 periphery members. Using the even smaller core from Figure 4(b), average hit rates are 50.1% for 48 core members and 36.9% for 138 periphery members.[8] However, when using another discrete core with only 33 members (not shown in the paper), this outstanding performance of core members is not maintained, because here periphery members have a slightly higher prediction hit rate on average.

It should also be mentioned that significantly better prediction success rates of core members could not be found in a setting with a third alternative of an unchanged economic situation (0) of $\hat{Y}_{t,t+6}$. That is, when we introduce a certain range about zero that we define as a situation without recognizable change of economic conditions, and compute our prediction hit rates on the basis of three categories for the real development instead of only two, core members are not found to be significantly better in predicting the future development than periphery members. Of course, with the introduction of a more or less arbitrary neutral range some of the differences in Fig. 4 between core and periphery members will also be 'neutralized', i.e. small positive and small negative changes will be assigned to the same class. As the category of neutral assessment is a well-known problematic feature of qualtitative surveys it is unclear how much weight we should place on the significance or lack of significance of differences between core and periphery under a binary or ternary classification of real economic changes. In any case, it appears remarkable that in the binary classification, the core members show consistently superior predictive performance over a broad range for the size of the core (with and without penalty functions used in the core detection algorithm).

7 Conclusion

Inspired by recent interest in modeling the expectations of economic agents, we have analyzed an index of economic sentiment from a new perspective. Rather than following the cumbersome standard assumption of individual's reporting their rationally expected forecasts based on their idiosyncratic information sets, we adopt a network perspective. Both with a discrete and with a continuous approach, the respondents can be classified into a core and periphery group (or a continuous structure with different

[7] Core members are no more optimistic on average, so even if positive changes of $\hat{Y}_{t,t+6}$ are more numerous in Figure 8, this also can not be the reason for their superior prediction record.

[8] It is needless to mention that mean differences in all cases so far are highly significant.

degrees of ‘coreness’). Even more supportive of such heterogeneity, we find significant correlation between the coreness patterns for different survey questions. The so identified core might consist of opinion leaders (consistent with behavioral theories) or might consist of those with a similar (not necessarily superior) information set (consistent with standard neoclassical theory). In fact, while the results from the last section seem to suggest core members are better informed than agents from the periphery on a first view, that need not necessarily be so. Even if core members had significantly better forecasts of future economic activity, we could also imagine a causal effect that is devoid of informational differences. If sentiment is an important driving force for real activity (following the ’animal spirits’ hypothesis), the better predictive performance could be a mere consequence of core members being the opinion leaders in the social process (or being close to the ‘true’ opinion leaders that need not necessarily belong to the group of the ZEW survey particpants) that is responsible for the ups and downs of sentiment. Their better predictive performance would then be a reflection of their hierarchical status and would have nothing to do with information about future economic circumstances. Hence, if the opinion dynamics has some direct causal influence on real activity, ‘coreness’ per se should account for higher correlation of predictions with future outcomes. Our present analysis cannot distinguish between this direct causation from animal spirits to economic activity and the interpretation of causality as evidence for ‘rational expectations’ of future events. In any case, the core-periphery dichotomy speaks in favor of redundancy of the information contained in the complete set of questionaires. This result is very much in line with the finding of an ‘effective’ number of agents that is smaller than the nominal number in the estimation results of a structural model of opinion formation (Lux, 2009).

References

[1] Akerlof, G., Schiller, R. (2009), *Animal Spririts: How Human Psychology Drives the Economy, and Why it Matters for Global Capitalism*, Princeton University Press.

[2] Arifovic, J. (1996), The Behavior of the Exchange Rate in the Genetic Algorithm and Experimental Economies, *Journal of Political Economy 104*, 510–541.

[3] Borgatti, S. and Everett, M. (1999), Models of Core/Periphery Structures, *Social Networks 21*, 375–395.

[4] Branch, W. (2004), The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations, *The Economic Journal 114*, 592–621.

[5] Breitung, J., and Jagodzinski, D. (2001), Prognoseeigenschaften alternativer Indikatoren für die Konjunkturentwicklung in Deutschland, *Konjunkturpolitik 47*, 292–314.

[6] Carroll, C. (2003), Macroeconomic Expectations of Households and Professional Forecasters, *The Quarterly Journal of Economics*, Feb., 269–298.

[7] Casti, J.L. (2010), *Mood Matters*, Springer.

[8] Everett, M. and Borgatti, S. (1999), The centrality of groups and classes, *The Journal of Mathematical Sociology 23(3)*, 181–201.

[9] Everett, M.G. and Borgatti, S.P. (2000), Peripheries of cohesive subsets, *Social Networks 21(4)*, 397–407.

[10] Hohnisch, M., Pittnauer, S., Solomon, S. and Stauffer, D. (2005), Socioeconomic interactions and swings in business confidence indicators, *Physica A 345*, 646 - 656.

[11] Hüfner, F., and Schröder, M. (2002), Prognosegehalt von ifo-Geschäftserwartungen und ZEW-Konjunkturerwartungen: Ein ökonometrischer Vergleich, *Jahrbücher für Nationalökonomie und Statistik 222(3)*, 316-336.

[12] Lux, T. and Schornstein, S., (2005), Genetic learning as an explanation of stylized facts of foreign exchange markets, *Journal of Mathematical Economics, Elsevier, 41(1-2)*, Feb., 169-196.

[13] Lux, T. (2009), Rational Forecasts or Social Opinion Dynamics: Identification of Interaction Effects in a Business Climate Index, *Journal of Economic Behavior and Organization 72*, 638 - 655.

[14] Menkhoff, L., Rebitzky, R. and Schroeder, M. (2009), Heterogeneity in Exchange Rate Expectations: Evidence on the Chartist-Fundamentalist Approach, *Journal of Economic Behavior and Organization 70(1/2)*, 241-252.

[15] Miller, P.J. (1994), *The rational expectations revolution: Readings from the front line*, The MIT Press.

[16] Nardo, M. (2003), The Quantification of Qualitative Survey Data : A Critical Assessment, *Journal of Economic Surveys 17*, 645-668.

[17] Nolte, I. and W. Pohlmeier (2007), Using forecasts or forecasters to forecast, *International Journal of Forecasting 23*, 15 - 28.

[18] Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (2007), *Numerical recipes 3rd edition: The art of scientific computing*, Cambridge University Press, New York.

[19] Westerhoff, F. (2010), An agent-based macroeconomic model with interacting firms, socio-economic opinion formation and optimistic/pessimistic sales expectations, *New Journal of Physics 12*, 075035.