Institute for Employment Research

The Research Institute of the Federal Employment Agency

IAB-Discussion Paper 25/2012

Articles on labour market issues

The formation of experts‘ expectations on labour markets

Do they run with the pack?

Norbert Schanne

ISSN 2195-2663

The formation of experts' expectations on labour markets: Do they run with the pack?

Norbert Schanne (IAB)

Contents

Abstract

Expectations regarding the economic development might be correlated due to various reasons: because individuals use the same public information and similar evaluation methods, and because of social learning or herding amongst peers. We analyse to what extent expectations are driven by herd behaviour, and if it contributes to make expectations more realistic.
In a novel survey the CEOs of the local departments of the German Federal Employment Agency report their expectations on unemployment in the short run. In this data we can discriminate between close and less-close peers to overcome the reflection problem and to quantitatively assess answers regarding the initial questions.
We find strong evidence for herding in expectation formation. The size of effect is robust across various specifications and remains even when controlling for forecasts from experts external to the survey. The social multiplier approximately doubles the effect of information (signals) included in the model. Compared to counterfactual expectations without herding constructed from the estimates, herding seems to improve the accuracy of the expectations.

Zusammenfassung

Erwartungen hinsichtlich der wirtschaftlichen Entwicklung können aufgrund verschiedener Ursachen korreliert sein: weil Individuen die selbe Information und ähnliche Auswertungsmethoden verwenden, oder weil sie andere beobachten, deren Erwartungen adaptieren, ihnen folgen. Wir analysieren, wieweit Erwartungen durch Herdenverhalten getrieben ist und ob dies dazu beiträgt, Erwartungen realistischer werden zu lassen.

In einer neuen Erhebung berichten die Geschäftsführungen der lokalen Arbeitsagenturen der Bundesagentur für Arbeit ihre Erwartungen bezüglich der kurzfristigen Arbeitslosigkeitsentwicklung. Dieser Datensatz erlaubt es uns, zwischen nahen (wichtigeren) und entfernten Bezugspersonen unter den Geschäftsführungen zu unterscheiden, um das Reflektionsproblem in Peer-Effekt-Studien zu lösen und so quantitative Antworten auf die Ausgangsfragen zu finden.

Wir finden starke Evidenz für Herdenverhalten in der Erwartungsbildung. Die Größenordnung des Effektes ist über verschidene Spezifikationen robust und besteht selbst dann, wenn wir für die Erwartungen von (externen) Prognoseexperten kontrollieren. Der Effekt von Signalen, von lokaler Information, wird durch die soziale Interaktion nahezu verdoppelt. Im Verhältnis zu kontrafaktischen Erwartungen, bei denen die Wirkung der sozialen Interaktion herausgerechnet wird, scheint die Genauigkeit der Erwartungen durch das Herdenverhalten zuzunehmen.

JEL classification: C 31; D 83; E 24; J64

Keywords: Economic expectations; Expectation formation; Herding; Information cascades; Labour market forecasts; Peer effects; Social learning; Spatial dependence

Acknowledgements: I would like to thank Silke Delfs and Reinhilde Zimmer for providing access to the survey used in this study. I gratefully acknowledge research assistance by Bastian Stockinger and Pirmin Schanne. Thanks even to Katrin Schmidt, Christian Hutter and numerous colleagues at IAB for comments on earlier versions on this paper, as well as to participants at seminars at University Regensburg, IAB Nuremberg, the 32nd International Symposium on Forecasting at Boston, the 66th European Meeting of the Econometric Society at Malaga and the 31st CIRET conference at Vienna. Remaining errors are to my fault.

1 Introduction

Expectations regarding forthcoming events and the future development of the economy are essential for the plans and decisions of economic and political actors. There's an ongoing debate on the formation and the (frequently missing) rationality of individuals' expectations, focusing on two questions: Which information (private or public, to-date or outdated etc.) do people evaluate in which way when formulating their expectations? Are individuals' expectations really looking forward, or do individuals when announcing their expectation pursue some other objective than predicting future outcomes (e.g. building up reputation)? In the present study on unemployment expectations we will address in particular the first of the two questions: our focus is to empirically ascertain the impact of social learning in expectation formation – that is, of observing and mimicking the expectations of other individuals – and to separate it from the effect of learning from public information.

Professional forecasters in the US[1], a small, intensely interactive group, seem to have fairly realistic expectations, using available information efficiently and adapting quickly to new trends (depending on the sample and the investigation design, see e.g. Chroushore, 1998, 2010 for an overview addressing to a large extent inflation expectations). Expectations of non-forecasters are, however, found to be less realistic. Private households seem to learn on average slowly about economic trends, with only a small part of the population receiving new information (e.g. from reading newspapers) and a large though declining fraction of the population relying on outdated information (Carroll, 2003; Curtin, 2003). Nevertheless, they even tend to extrapolate past trends into the future, causing overly pessimistic expectations at the end and too optimistic early in a recession (Tortorice, 2012).

Surprisingly, most studies on unemployment expectations treat individuals' responses as independent from each other although they, when asked about their sentiments, report similar expectations. Of course, it might be argued that – at least for professional forecasters – individuals expect a similar development because they utilize the same or at least related information; as well, they employ similar methods and models (see Zarnowitz/Lambros, 1987; Keane/Runkle, 1990; Hey, 1994). Thus, strongly diverging expectations generated using the same data and methods would indeed come at a surprise; diverging expectations would result from diverging (private) information. However, correlation of expectations might even be due to social interaction if people either communicate their private information to other persons or conclude on the information received by other individuals from observing the announcements, actions and decisions amongst their peers. In the extremum of 'herding' (Banerjee, 1992) or an 'informational cascade' (Bikhchanandi/Hirshleifer/Welch, 1992) social learning may cause total disregard of the respective personal information.[2] As a consequence, expectations can follow blindly a wrong direction or, on the opposite, converge fast to the correct direction/value. Where the informational cascade

[1] To our knowledge, forecasting studies using European data focused on GDP, inflation and stock-market expectations; corresponding studies analysing unemployment expectations in European countries are to date missing.

[2] Both models, as well as a number of follow-up studies, have a similar structure: Individuals receive private signals; they observe the behaviour of other persons who had to decide earlier, and conclude from their observations on the social aggregate over previous signals; and the individuals make their decisions according to a combination of their private signal and the socially aggregated signal. In both models two

ends up and when it is broken, is a question of the initial direction and of (public) information that becomes available throughout the cascade. Various studies, e.g. Bikhchanandi/Hirshleifer/Welch (1998), Smith/Sørensen (2000), Çelen/Kariv (2004), Chamley (2004), Manski (2004), Acemoglu/Ozdaglar (2011) and Acemoglu et al. (2011), provide deeper (theoretical, simulation and experimental) insight under which conditions – unboundedness of private believes, continuous rather than discrete signals, repeated decisions in a stationary environment, etc. – social learning warrants convergence to the correct expectation/decision. However, it might be interesting even to quantitatively assess the effect of social learning, and to test if herding contributes to more realistic expectations – despite the conceptual differences between herding theory (modeling a sequence of decisions in continuous time in a static world) and empirical herding analysis (with data observed in discrete time in a dynamic world) as it has been emphasized by Welch (2000).

Some papers – addressing expectations on unemployment and other macroeconomic figures such as GDP growth, inflation or stock market indices – use the serially lagged consensus, that is the expectations' average, as herding variable (e.g. Bewley/Fiebig, 2002; Rangvid/Schmeling/Schrimpf, 2012). This might be a misspecified measure since herding should arise amongst fairly contemporaneous observations and not with those announced a month or quarter ago. Even less studies test for herding in contemporaneous expectations. Pons-Novell (2003) provide evidence of herding amongst some groups of forecasters in the Livingston Survey. Rülke/Tillmann (2011), in contrast, reject herding of the FOMC members with regard to unemployment sentiments. Both studies rely on the contemporaneous consensus. As we will argue, this measure may be also problematic: empirical identification of social learning, herding and imitation among peers is crucial, but not warranted in a linear-in-means peer model.

Manski (1993) shows that, in a model entailing on the right hand an expectation about the peers' observations of the dependent variable and expectations about the distribution of explanatory variables amongst the peers, the parameters of this model are frequently not identified when estimating conditional expectations. If a variable's expectation is linear (that is, estimable by the equally weighted arithmetic mean) the expectation of the dependent variable on the left hand will be reflected by the expectation of the dependent variable (across the peers) on the right hand, and likewise will be the expectations of the explanatory variables. The reflection problem may be overcome by nonlinear exclusion restrictions

(or three) subsequent individuals acting in the same fashion are sufficient to initiate a non-optimal social outcome (be it sitting in the worse restaurant, or having the wrong believe with regard to future development of unemployment), and only the first individual of these two is required to have the wrong information. The private information of all following individuals becomes in general irrelevant once the cascade has started and the herd began to move. Bikhchanandi/Hirshleifer/Welch (1998) and Chamley (2004) demonstrate that cascades occur with a high probability.

The models of Banerjee (1992), Bikhchanandi/Hirshleifer/Welch (1992), Smith/Sørensen (2000) and Acemoglu et al. (2011) describe games in which each individual has only to make a singular decision. Learning from mistakes, or gaining reputation as a person with good information, is not possible in this framework whereas both might be possible if decisions have to be made (expectations have to be formed) repeatedly (as in Manski, 2004). However, the effects of learning and, in particular, of reputation will be only minor if the correct outcome varies across time and if the private signals regarding this outcome are redistributed every period: then, a person knows about the relative reliability of the (observable) public information compared to the private signal and the socially aggregated information; but not the reliability of her signal's current realization or the reliability of the signal she presumes a single peer to have received.

e.g. due to a nonlinearity in the expectation or by imposing a network structure amongst the peers (for theoretical discussions and an overview across various applications see Soetevent, 2006 and Bramoullé/Djebbari/Fortin, 2009).

The survey employed in this paper – collected amongst the CEOs of the local departments of the German Federal Employment Agency (FEA) – allows to assess to each respondent her location. By assuming more intense communication amongst geographical neighbours and amongst CEOs forced to meet frequently, we are able to discriminate between close and less-close peers. Thus, we can impose a network structure amongst the survey participants which so far has not been possible in the literature on (macro-)economic expectations. The cross-sectionally dependent model is estimable with spatial econometric techniques since we observe a complete network and thus have hardly a problem of omitted peers, missing network nodes or 'edge effects' (omitted spatial neighbours). The dependence structure is similar to a spatial Durbin model (e.g. LeSage/Pace, 2009), i.e. a model including a spatially lagged dependent variable to account for the 'endogenous effect' amongst peers, spatially lagged exogenous variables for the 'contextual effect' as well as a spatially correlated error term for the 'correlated effect'.

We model unemployment-growth expectations to be affected by previous local unemployment and vacancies as the observable market fundamentals according to a matching function, private (for us and for other CEOs unobservable) signals about job destruction, job creation and plant closures, and additionally by endogenous and contextual peer effects. We find strong evidence for the interdependence of the expectations among peers which we interpret as social learning. Our estimate for the social multiplier – a measure for the endogenous peer effect – mounts approximately to two, robustly across various specifications; that is, social interaction roughly doubles the effect which we assess directly to observable characteristics. The size of this effect persists even if we include public professional unemployment forecasts as an alternative source of social information. Furthermore, though expectations in our survey are overly pessimistic, we find that social learning brings them closer to the realized development. Herding seems indeed to make expectations more rational.

However, our paper is limited with regard to some aspects. First, we abstract from strategic herding due to reputation effects (described e.g. by Scharfstein/Stein, 1990; Lamont, 2002; Ottaviani/Sørensen, 2006); admittedly, we wouldn't be able to distinguish reputation effects from social learning with the available information. The labour-market experts in the FEA survey gain reputation from effectively reducing unemployment, not from having expectations close to those of their principals; hence, reputation bias is likely more severe amongst professional forecasters.[3] Second, we do not directly test if expectations are rational although we touch this issue in Section 7. Supposedly, the costs of underestimating unemployment (e.g. in terms of lacking the resources necessary to get unemployed quickly into training or work) may exceed the costs of overestimating (e.g. in terms of ALMP participation below the capacity frontier). Thus, rational CEOs would have an asymmetric loss

[3] Note that Keane/Runkle (1990) argue that survey responses amongst professional forecasters are more realistic. However, various studies finding reputation bias stronger amongst professional forecaster than amongst non-professional may provide support for our perspective.

function whereas we employ mean-square loss as it is custom. Third, we leave to further research whether the reported expectations correspond to subsequent action.

The paper is organized as follows. Section 2 describes the survey and the record data considered throughout the investigation. Section 3 discusses the basic model, Manski's Reflection Problem and a number of estimation issues arising in spatially autoregressive panel models. Section 4 entails the analysis of the expectation formation process in general. Sections 5, 6 and 7 focus on particular aspects of expectation formation: learning from public information, the detection of an informational cascade through its fragility and the question if expectations become more realistic by herding. Central results are recapitulated in the conclusion.

2 Data

2.1 The FEA Management Survey

The following section describes the Management Survey of the FEA which has been started in November 2008.[4] In addition to this survey we use record data generated inside the FEA in labour-market administration processes and provided by the FEA statistics. The responsibility for the survey is at the FEA's labour-market monitoring service. The survey is collected at a monthly frequency amongst the CEOs (Vorsitzende der Geschäftsführung, VG) of the local offices of the FEA. Probably, these are more expert than the participants in the Michigan Survey of Consumers but have less forecasting expertise than the respondents in the Survey of Professional Forecasters or the Livingston Survey (as the three surveys frequently employed in the literature on unemployment expectations). Reporting date is around the record day of FEA statistics; the realized numbers of unemployed persons, participants in labour-market programmes and employees in the current month are unknown at the date of response, the corresponding numbers of the previous months have already been published (except the number of employees according to the FEA register which is first released with a three-months delay). The set of questions and the possible items corresponding to the questions vary; a short summary is provided in Table 9 in the appendix. However, the question which is of central interest in our paper has been observed continuously since the beginning:
How do you expect unemployment in your district to develop within the next three months (besides the usual seasonality)?
The answer allows five items [Decline strongly, Decline, Stay constant, Increase, Increase strongly] which are associated numerically to the values $\{-2, -1, \ldots, 2\}$. Although the variable is ordinal and only defined on a 5-groups Likert scale, we treat it like an interval variable throughout the analysis; this allows us not only to use quantiles of this variable but even means and standard deviations. However, we try to be careful about interpreting numbers.

4 Results of the current and recent waves are available in the FEAs intranet. For further information and data access, contact the corresponding author.

Respondents typically answer in consultation with their top-level staff; thus, the answers can be considered as an institutional expectation rather than an individual. A major advantage is that the answers are not anonymously. We know the agency district (that is the location) of each respondent. As a consequence, we can easily assign to each respondent her geographical neighbours, her Regional Division[5] (RD) and thus those regions reporting to the same principal, as well as those respondents sharing the same benchmarking group[6]. Responses are available for all 176 agency districts for every month since November 2008 (till April 2012, the last month considered here); as participation in the survey is not voluntary, it is not affected by non-response or panel mortality.

Throughout the entire observation period, 676 times a CEO expected unemployment to increase strongly within the next tree months, 2056 times to increase (moderately), 2249 times to remain at the same level, 2371 times to decline, and only 40 times to decline strongly. These (the 40 decline-strongly responses) are observed in only 12 districts, of which in three districts the CEOs expected unemployment to decline strongly five times and in one district ten times. The asymmetry in the answers may be partly due to the German job miracle: unemployment remained, against any economic intuition, rather low during the credit-crunch crisis – and thus had not too much potential to decline during the recovery.

Table 1: Unemployment sentiments – descriptive statistics

Statistic	Overall	I/2009	II/2009	I/2010	II/2010	I/2011	II/2011
Median	0.000	1.000	1.000	0.000	-1.000	-1.000	0.000
Mean	0.129	1.335	1.014	0.127	-0.609	-0.775	-0.390
Std. Dev.	0.985	0.509	0.577	0.824	0.529	0.447	0.607
Within Std. Dev.	0.981	0.326	0.406	0.719	0.372	0.329	0.434
Cross Std. Dev.	0.543	0.504	0.562	0.620	0.526	0.440	0.579

Descriptive figures allow a first illumination of the data. Common statistics are reported in Table 1; numbers in column 2 refer to the entire observation period whereas columns 3 to 8 show the corresponding statistics calculated for the respective half-year. According to the median over the entire observation period and the semi-annual medians the CEOs seem to expect a more or less cyclical development of unemployment, such that increases balance with decreases and that, over a longer period, unemployment remains rather constant. However, they seem to be slightly pessimistic; the total mean and the period-specific averages are always higher than the corresponding median values. The dispersion of the sentiments is dominated by the variation over time. The within standard deviation[7] has the same size as the total standard deviation; the cross-sectional standard deviation

[5] Regional Divisions form the intermediate organisational level between the local and the national. The ten RDs have approximately the size of the major German federal states and collect between 8 (Berlin/Brandenburg) and 33 (Northrhine-Westfalia) agency districts.

[6] Each agency district belongs to a benchmarking group implemented according to comparable economic conditions (the procedure is described by Rüb/Werner, 2008 and Blien/Hirschenauer/Thi Hong Van, 2010); these groups do not coincide with the Regional Divisons. CEOs (and their top-level staff) have to participate in periodical meetings of their benchmarking group.

[7] Within Std. Dev. $= \sqrt{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it}-\bar{y}_i)^2}$; Cross Std. Dev. $= \sqrt{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it}-\bar{y}_t)^2}$.

(computed using the deviations from the time-specific mean) has significantly smaller size. However, a large amount of the temporal variation in the sentiments seems to be due to the variation in the first half-year in 2010. The standard deviation in other half-years mounts to a size similar to that of the cross-sectional standard deviation.

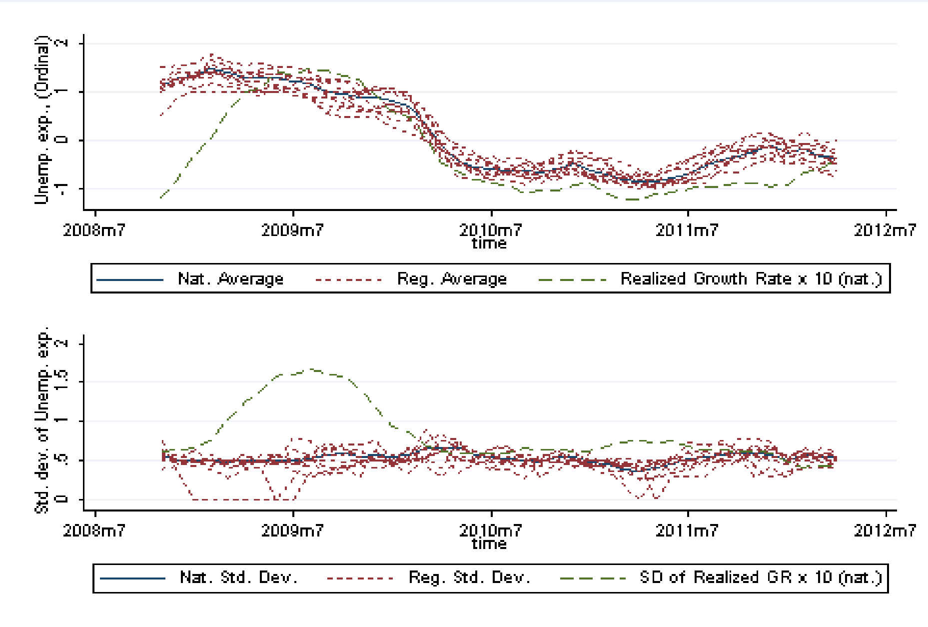

Figure 1: Development of unemployment sentiments

The early year 2010 is even more outstanding when we inspect the series of monthly national and regional mean expectations in Figure 1. Sentiments of local CEOs are averaged at the RD level to compute regional means. The figures indicate little variation in expectations across regions. They seem to develop in a similar fashion, more or less parallel to the national average; only in a single RD the expectations move towards 'declining unemployment' earlier than the average as the lowest dashed line (in the period from December 2009 to March 2010) in the upper diagram of Figure 1 indicates. Time-specific standard deviations with regard to the national and each regional data set (the lower diagram) are in general close to or below 0.5 (in one RD, all CEOs gave the same response in early 2009); only in Spring 2010 it exceeded 0.5. Hence, the distribution of responses seems to exhibit a shift of the first moment but stable second moments.

Figure 2 contrasts the expected and the realized development of unemployment, both monthly averaged over the local districts in each regional division. We divide the data in two samples because of the expectations' shift in early 2010. Throughout the entire period before January 2010 all regional averages across expectations take on values above 0.5; that is, the majority of CEOs in every regional division expected unemployment to increase. Averages higher than 1.5 – corresponding with a regional majority of CEOs expecting a strong increase – are however rare. This concentration of sentiments' averages does not correspond to the realized development of unemployment at the forecast horizon: we observe both declining unemployment (with a regional average growth rate of roughly -2.5 percent within three months, or -10 percent annually) as well as strong increases (up to 10 percent within a quarter, corresponding with an annual growth rate of roughly 40

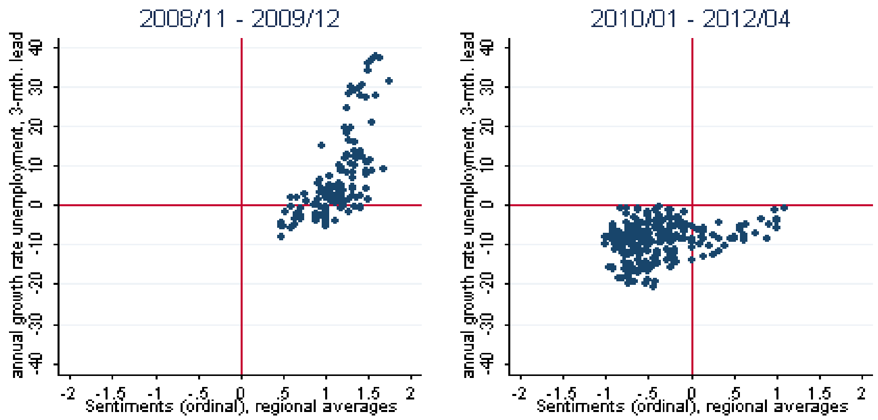

Figure 2: Sentiments and realized unemployment growth

percent).

In the period from January 2010 onward both expectations and realized unemployment growth shifted downward; all regional averages in the right sentiment-realization diagram are located in the second and third quadrant. That is, unemployment declined everywhere within the next quarter (with an estimated annual growth rate up to -20 percent). Nevertheless, the sentiments seem again too pessimistic: In a number of regional divisions, the majority of CEOs still expected unemployment to increase (in particular in early 2010), and the mass of the distribution is slightly right from the value corresponding to a moderate decline (in the direction of the 'no change' item).

2.2 Prospective public information

Public information plays a key role particularly in models for expectations of non-professional forecasters. In the model of Bikhchanandi/Hirshleifer/Welch (1992), external public information may cause the collapse of an informational cascade; thus, the effect associated with herding likely exhibits a response. In the model of Carroll (2003), only a fraction of private households receive news about professional forecasts and adapt their expectations accordingly; social learning might interact with learning from public information. We discuss both in detail later, hence it seems adequate to clarify our notion of public information before.

Though the local labour market data reported by the FEA statistics which is available to each CEO as well as outside the Federal Employment Agency is in general public data, we consider it here as private information with regard to trend reversals: cyclical shifts may be hard to detect besides the seasonal and irregular fluctuation. Another source of public information are the business-cycle forecasts published by the major German economic research institutes.[8] Their forecasts on GDP growth and unemployment in the years 2009 till 2011 are listed by month of release in Table 2.

Table 2: Availability of public information on trend reversals

Released in	Publishing institute	GDP growth rate forecasts for				unemployment forecasts (in '000 persons) for			
		2009	2010	2011	2012	2009	2010	2011	2012
Oct 08	IMK	0.2	–	–	–	3263	–	–	–
Dec 08	ifo	-2.2	-0.2	–	–	3471	3971	–	–
Dec 08	IfW	-2.7	0.3	–	–	3665	3949	–	–
Dec 08	IMK	-1.8	–	–	–	3882	–	–	–
Jan 09	RWI	-4.3	0.5	–	–	3727	4633	–	–
Mar 09	IfW	-3.7	-0.1	–	–	3642	4251	–	–
Apr 09	IMK	-6.0	-0.5	–	–	3718	4688	–	–
Jun 09	ifo	-6.3	-0.3	–	–	3581	4337	–	–
Jun 09	IfW	-6.0	0.4	–	–	3576	4365	–	–
Jun 09	RWI	-5.0	1.2	–	–	3481	4133		–
Jul 09	IMK	-6.5	-0.4	–	–	3575	4448	–	–
Sep 09	IfW	-4.9	1.0	–	–	3444	3881	–	–
Oct 09	IMK	-5.0	1.2	–	–	3470	4075	–	–
Dec 09	ifo	-4.9	1.7	1.2	–	3426	3607	3617	–
Dec 09	IfW	–	1.2	2.0	–	–	3827	3935	–
Dec 09	IMK	-4.9	2.0	–	–	3424	3600	–	–
Jan 10	RWI	–	1.4	1.6	–	–	3475	3565	–
Mar 10	IfW	–	1.2	1.8	–	–	3443	3275	–
Apr 10	IMK	–	1.5	1.4	–	–	3382	3313	–
Jun 10	ifo	–	2.1	1.5	–	–	3233	3043	–
Jun 10	IfW	–	2.1	1.2	–	–	3199	2952	–
Jun 10	IMK	–	2.0	1.5	–	–	3226	3048	–
Jun 10	RWI	–	3.4	2.2	–	–	3250	3055	–
Sep 10	IfW	–	3.4	1.7	–	–	3235	2958	–
Oct 10	IMK	–	3.5	1.9	–	–	3236	2933	–
Dec 10	ifo	–	3.7	2.4	–	–	3242	2943	–
Dec 10	IfW	–	3.7	2.3	1,3	–	3252	2984	2778
Dec 10	IMK	–	3.7	2.5	–	–	3240	2963	–
Jan 11	RWI	–	–	2.9	2,4	–	–	2875	2467
Mar 11	IfW	–	3.6	2.8	1,6	–	3244	2992	2803
Apr 11	IMK	–	–	2.7	1,7	–	–	2944	2758
Jun 11	ifo	–	–	3.3	2,3	–	–	2944	2683
Jun 11	IfW	–	–	3.6	1,6	–	–	2970	2687
Jun 11	IMK	–	–	4.0	2,3	–	–	2949	2740
Jun 11	RWI	–	–	2.9	2,3	–	–	2965	2650
Sep 11	IfW	–	–	2.8	0,8	–	–	2984	2868
Sep 11	RWI	–	–	2.9	1.0	–	–	2965	2805
Oct 11	IMK	–	–	3.2	0.7	–	–	2977	2865
Dec 11	ifo	–	–	3.0	0.4	–	–	2975	2800
Dec 11	IfW	–	–	2.9	0.5	–	–	2976	2863
Dec 11	IMK	–	–	3.0	-0.1	–	–	2976	2900
Mar 12	IfW	–	–	–	0.7	–	–	–	2775
Mar 12	IMK	–	–	–	0.3	–	–	–	2876
Jun 12	ifo	–	–	–	0.7	–	–	–	2866
Jun 12	IfW	–	–	–	0.9	–	–	–	2866

Forecasts on unemployment (and GDP growth) published by the major German economic research institutes. ifo = ifo institute Munich; IfW = Institute for the World Economy Kiel; IMK = Institute for Macroeconomic Policy; RWI = RWI Essen. The ZEW centre for european economic research Mannheim (ZEW) published only GDP growth rate forecasts. In January 2012, the Halle Institute for Economic Research (IWH) didn't provide their forecasts retrospectively for the period before 2011; thus, their forecasts are not listed.

It might be easier to identify a trend reversal in these publicly available forecasts, as we will demonstrate with Table 2; for comparison, note that realized unemployment numbers

8 Those forecast which attract the greatest deal of attention are, supposedly, published by the Ifo Institute for Economic Research (Ifo, Munich), the Institute for the World Economy (IfW, Kiel), the ZEW Centre for European Economic Research (ZEW, Mannheim), the Macroeconomic Policy Institute (IMK, Düsseldorf), the Halle Institute for Economic Research (IWH, Halle) and the Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI, Essen); these institutes are (or have been) moreover involved in the Joint Economic Diagnosis (Gemeinschaftsdiagnose). With regard to labour markets, forecasts of the Institute for Employment Research (IAB, Nuremberg) receive high attention as well.

had been on average 3.258 million in 2008, 3.415 million in 2009, 3.238 million in 2010 and 2.972 million in 2011. All institutes expected unemployment to rise in 2009; throughout most of the year 2009, by five to ten percent within the remaining year and by another ten to 25 percent in the following. In late 2009, the expected rise of unemployment became smaller, and in January 2010 the first institute expected unemployment to rise by less than five percent within the same and the subsequent year. In the next forecast published in the first quarter of 2010, a research institute already predicted a decline in unemployment.

3 Empirical design

3.1 The basic model of labour market expectations

In the following paragraphs, we will construct a fairly simple model which we hold to be valid in describing the CEOs expectation formation process. The core of the model consists of three variables related to two components: First, we suppose that the CEOs expect the short-run unemployment dynamics to follow a matching process: jobs are created by matching unemployed persons and vacant jobs. Thus, unemployment and vacancies (included in annual differences to eliminate seasonality and non-stationarity) are two central observable indicators describing the market fundamentals.[9] Second, CEOs typically will receive information on planned job destruction (firm closures and job separation) which is not collected by the FEA statistics; this, as well as information on vacancies not reported to the statistics, form private signals on labour market dynamics. We denote the sentiments with y and the vector entailing the market fundamentals (unemployment and vacancies in logs) with x.

Furthermore, we assume that the CEOs take also information from recent periods into consideration; this may result in belief persistence or slow/delayed adaption to new information. The entire information available up to period $t-1$ is already incorporated in the sentiments reported in the previous period; thus, we include the first lag of the dependent variable (denoted with $L.y$) as additional regressor on the right hand (following Carroll, 2003). The autoregressive component may also reflect that CEOs adapt to prediction errors over time – which in fact would correspond to a Moving-Average (MA) process.

It is unlikely that the CEOs consider only the reported market fundamentals referring to the district they are responsible for. Regional labour markets are interdependent, thus it would be irrational not to look at unemployment and vacancies in other regions. The development in other regions, or the aggregate development, may help to assess whether ups and downs in unemployment are national phenomena affecting all regions or a local phenomenon leveling across the regions; it has been shown that consideration of spatial co-developments provides prospective information (e.g. Schanne/Wapler/Weyh, 2010;

[9] Current unemployment and vacancies are two of the numbers published monthly by the FEA at various regional levels; their development is typically discussed in the first chapter of the monthly German Labour Market Report (together with employment which is published with a delay of one/two months). Validity of vacancies as a leading indicator regarding unemployment at a three-months forecast horizon has been shown recently by Schanne (2010).

Mayor/Patuelli, 2012). CEOs aggregate the information observable across other regions, that is across those districts under responsibility of their peers, to a conditional expectation which we denote with $E(x|p)$; here, p represents the characteristics defining the peer group or the social network. In addition to this, we expect a CEO and her top level staff to communicate with other CEOs/their top level staff (i.e. with the persons considered as their peers) in their daily business. Supposedly labour-market expectations are subject matter of communication in the time between two waves of the survey. Then, a (rational) CEO would draw information from the (preliminary) sentiments announced by her peers; the aggregate over the peers' sentiments is $E(y|p)$.

Thus, we need to consider four sources of information in our model of expectation formation: the own history of sentiments (which accounts for learning from own mistakes as well as past information), statistical (verifiable) information directly related to an observation, an aggregate of contextual statistical (verifiable) information related to the peers, and social learning. A regression model with these four sources of information can be written as follows:

$$y = \alpha + x' \beta + E(x|p)' \gamma + E(y|p) \eta + L.y \phi + e \tag{1}$$

When estimating, that is when taking expectations conditional on x and p, and under the assumption that the unobservable signals (the disturbances) are somewhat related amongst peers with $E(e|x,p) = g(p;\delta)$, eq. (1) becomes:

$$\begin{aligned} E(y|x,p) = \alpha + E(x'\beta|x,p) + E(E(x|p)'\gamma|x,p) + E(E(y|p)\,\eta|x,p) \\ + E(L.y|x,p)\,\phi + g(p;\delta) \end{aligned} \tag{2}$$

With linear social expectations on y (in the endogenous peer-related effects), the *Reflection Problem* (see Manski, 1993) becomes obvious when applying the Law of Iterated Expectations to eq. (2). When sequentially conditioning first on p and then on x, it is possible to receive reduced-form parameters in the scalars $\frac{\phi}{1-\eta}$, $\frac{\alpha}{1-\eta}$ and the vector $\frac{1}{1-\eta}(\beta\eta + \gamma)$ but, with exception of β not the structural parameters themselves from the following equation:

$$E(y|x,p) = \tfrac{\alpha}{1-\eta} + x'\beta + E(x'|p)\tfrac{\gamma+\beta\eta}{1-\eta} + E(L.y|x,p)\,\tfrac{\phi}{1-\eta} + \tfrac{g(p;\delta)}{1-\eta} \tag{3}$$

However, a nonlinearity in the model might be sufficient to overcome the identification problem in eq. (2) (see Manski, 1993, Sec. 3; Brock/Durlauf, 2001). Identification in models for an ordered categorial variable will be investigated in Subsection 3.3.

When sampling issues are negligible (because the data covers the entire population or the sampling accounts for this particular structure), spatially autoregressive models are an alternative to a nonlinear functional form(see Manski, 1993, Sec. 2.6): Since distance (as well as neighbourhood relations or communication intensity) to a third unit varies across observations, a weighted average with weights accounting for these distance relations shows variation itself. The intuition behind is that one needs to discriminate between peers and non-peers (or close and distant peers), that is to observe variation, in order to identify the effect of peers (an argument raised in a similar form by De Giorgi/Pellizari/Redaelli, 2010). Likewise, identification of spatially autoregressive effects in a cross section or in a panel saturated with time-specific effects is impossible if a spatial system entails only neighbouring regions; without time-specific effects, the temporal variation identifies the endogenous peer effect (see, e.g., Kelejian/Prucha, 2002; Baltagi, 2006; Baltagi/Liu, 2009). Identification issues in spatial autoregressive panels are discussed in Subsection 3.2.

In our data we are able to investigate alternative peer-relation structures. First, we know the geographical location of each respondent. In daily business communication of (top level) staff members in different agencies is typically most intense with those in close agency districts, that is with those providing job-seekers and vacancies within commuting distance. Thus, using geographical distance (or contiguity) of regions as a measure for communication structures (and for peer relations) in expectation formation seems plausible. Political structures in which groups of CEOs meet periodically provide an alternative network structure which may serve as a robustness check. Here, utilization of benchmarking groups as social network seems advantageous over using Regional Divisions since it is not possible to separate social learning from reputation effects when analysing the latter.

3.2 Estimation of spatially autoregressive panels

In terms of spatial econometrics, eq. (1) is a dynamic panel regression with a spatially lagged dependent variable (endogenous effect), spatially lagged exogenous variables (contextual effect) and, likely, spatially autocorrelated disturbances (correlated effect). With different sets of weights for the endogenous, the contextual and the correlated effect, the empirical model becomes:

$$y_{i,t} = \phi y_{i,t-1} + \eta \sum_{j=1}^{n} w_{(1)ij} y_{j,t} + \alpha + x_{i,t}\beta + \sum_{j=1}^{n} w_{(2)ij} x_{j,t}\gamma + e_{i,t} \tag{4}$$

with disturbances allowing for possible time-invariant unobserved heterogeneity, serial and network autocorrelation:

$$e_{i,t} = \mu_i + u_{i,t} \tag{5a}$$

$$u_{i,t} = \delta \sum_{j=1}^{n} w_{(3)ij} u_{j,t} + \nu_{i,t} \tag{5b}$$

$$\nu_{i,t} = \rho \nu_{i,t-1} + \varepsilon_{i,t} \qquad . \tag{5c}$$

Eqs. (4) and (5) are in matrix notation with ι_T a $T \times 1$ vector of ones, I_T the T-dimensional unit matrix (likewise for dimensions n and nT), L_T the lag operator matrix (a $T \times T$ matrix with the first sub-diagonal containing ones and all other elements set to zero), $W_{(1)}, W_{(2)}, W_{(3)}$ as $n \times n$ matrices containing the spatial weights, Y, e and ε as Tn-dimensional column vectors arranged such that $Y = (Y_1', \ldots, Y_T')'$ with $Y_t = (y_{1,t}, \ldots, y_{n,t})'$, μ a n-dimensional vector of region-specific fixed effects and X a $(Tn) \times k$ matrix:

$$Y = \phi(L_T \otimes I_n)Y + \eta(I_T \otimes W_{(1)})Y + \iota_{Tn}\alpha + X\beta + (I_T \otimes W_{(2)})X\gamma + e \quad \text{(6a)}$$

$$e = (\iota_T \otimes I_n)\mu + \left(I_T \otimes (I_n - \delta W_{(3)})^{-1}\right)\left((I_T - \rho L_T)^{-1} \otimes I_n\right)\varepsilon \quad \text{(6b)}$$

The sequences of weights in each matrix $W \in \{W_{(1)}, W_{(2)}, W_{(3)}\}$ are assumed to be exogenous triangular arrays satisfying standard regularity conditions:

1. $w_{ii} = 0 \quad \forall i \in \{1, \ldots, n\}$, i.e. the main-diagonal elements of W are zero.
2. $w_{ij} \geq 0 \quad \forall i, j \in \{1, \ldots, n\}$, i.e. there are no negative spatial weights.
3. W has bounded row norm $\|W\|_\infty = \max_{j \in \{1,\ldots,n\}} \sum_{i=1}^{n} |w_{ij}| \leq c$, bounded column norm $\|W\|_1 = \max_{i \in \{1,\ldots,n\}} \sum_{j=1}^{n} |w_{ij}| \leq c$ with $c < \infty$.

Thus, W has bounded spectral norm[10] $\|W\| \leq c$. Furthermore, assume $\delta \in (\frac{1}{\min\{\lambda_{W_{(3)}}\}}, \frac{1}{\max\{\lambda_{W_{(3)}}\}})$ and $\eta \in (\frac{1}{\min\{\lambda_{W_{(1)}}\}}, \frac{1}{\max\{\lambda_{W_{(1)}}\}})$ where $\{\lambda_W\}$ denotes the sequence of real eigenvalues extracted from W. That is, δ and η are, in absolute value, smaller than the inverse of the largest eigenvalue of the corresponding weights matrix. If these conditions hold, $(I_n - \delta W_{(3)})$ and $(I_n - \eta W_{(1)})$ will be finite and invertible, and the stochastic process will be cross-sectionally weakly dependent (see Chudik/Pesaran/Tosetti, 2011).

In general, the matrices associated with the endogenous, the contextual and the correlated effect may but do not need to be different. Depending on the estimation technique, the parameters are still identified when using the same weights for only two or all three components. Eq. (6) may be estimated consistently[11] either by (Quasi) Maximum Likelihood[12] (ML) or by Generalized Method of Moments (GMM) whereas the OLS estimator is biased due to the endogeneity of WY (see Anselin, 1988, 2001). Here, we rely on GMM since various instrumentation tests allowing better judgement on identification are not available for the ML estimators (see the discussion on model identification in Gibbons/Overman, 2012). The spatially lagged dependent variable WY_t is instrumented by (first- and second-order) spatially lagged exogenous/predetermined explanatory variables WX_t, W^2X_t, WY_{t-1}. The error-component parameters are (with exception of fixed effects) estimated in a separate step (see Mutl/Pfaffermayr, 2011 as well as Kelejian/Prucha, 1998, Kapoor/Kelejian/Prucha, 2007). Utilization of the same weights for modeling the contextual effects and for constructing the instruments may affect the validity of the GMM estimator. However, a sufficiently high partial R^2 statistic of the excluded instruments in the first-stage regression (explaining WY) and insignificance (low significance) when testing for overidentifying restrictions (Hansen/Sargan test) indicate an appropriate instrumentation strategy. Given valid instruments the parameters η, β and γ are identified. δ is identified if there exists a consistent estimator for $\alpha, \beta, \gamma, \eta$ and ϕ; however, δ is of minor interest for us.

[10] The spectral norm of a matrix A is the square root of the largest eigenvalue of AA', i.e. $\|A\| = [\max\{\lambda_{(AA')}\}]^{\frac{1}{2}}$.

[11] Nickell-Bias (Nickell, 1981) in the estimate of ϕ is negligible when T is sufficiently large or when the variance of the time-constant error component converges to zero. Thus, we estimate the model in levels and treat the serially lagged dependent variable as weakly exogenous.

[12] The ML estimator relies on the n-dimensional (or nT-dimensional) multivariate Gaussian distribution and thus accounts explicitly for the simultaneity of the observations.

We let the weights used for the endogenous effects and the contextual effect vary across the specifications in order to establish robustness of the results: as first set of weights we use an indicator variable where regions are considered as peers if the distance between their centroids is smaller than 88.66 km (the mean between percentile 90 and percentile 95 of all pairs of inverse distances; this value is chosen such that any region has at least two peers). The binary information is row-normalized such that $\sum_{j=1}^{n} w_{ij} = 1 \quad \forall i$. Besides this truncated inverse-distance metric we consider contiguity (indicating that regions share a border) and affiliation with a benchmarking group, both as row-normalized information. We always employ the same set of weights for the endogenous effect and the excluded instruments. In our preferred specification, the contextual effect is constructed with a fourth metric: we directly use inverse distance between regions as continuous weight, rather than row-normalized indicator variables.

3.3 Identification in nonlinear regression models

Procedures for ordered categorial limited dependent variables (LDV) as an alternative estimation strategy come along with two advantages. First, the values of the dependent variable do not reflect equally sized intervals; a procedure like *ordered probit* treats that more adequately than a linear approach. Second, most models for LDVs employ the cumulative density function (CDF). The generated nonlinearity might allow identification under certain conditions (see e.g. Manski, 1993; Brock/Durlauf, 2001 and Brock/Durlauf, 2003 for the bivariate and multinomial-logit case).

However, Bajari/Krainer (2004) argue convincingly that continuous exclusion restrictions are still necessary for identification in the ordered-probit peer effects model. Appropriate exclusion restrictions can for example be derived from distinction between first and second-order neighbours in the peer-related effects while aggregating linearly over the CDFs at individual level.

Estimation of the spatially-autoregressive ordered probit with ML requires structural assumptions on the latent variable y_{it} and its joint distribution over $i = 1, \ldots, n$ and $t = 1, \ldots, T$. Let the endogenous peer effect and the autoregressive component refer to the latent variable such that they can move to the left hand. Further, the parameters on the right hand are identified only up to a variance-scaling parameter σ such that the stochastic process can be represented by a standardized distribution (with unit variance). Eq. (6a) becomes, in reduced form with $H_{(\phi,\eta)} = \left[I_{Tn} - \phi(L_T \otimes I_n) - \eta(I_T \otimes W_{(1)})\right]$

$$Y = H_{(\phi,\eta)}{}^{-1}\left[X\tfrac{1}{\sigma}\beta + (I_T \otimes W_{(2)})X\tfrac{1}{\sigma}\gamma\right] + H_{(\phi,\eta)}{}^{-1}\tfrac{1}{\sigma}(\alpha + e). \tag{7}$$

Let $\upsilon = H_{(\phi,\eta)}{}^{-1}(\alpha + e)$. Then, the observable LDV $\tilde{y}_{it}$ takes on the value j if $\upsilon_{it} \in (a_{j-1}, a_j]$. We relax the assumptions of fixed effects and serially and spatially autocorrelated disturbances in order to reduce the complexity of the estimator. Though, we need to integrate over a nT-dimensional probability density function (PDF) $f(\upsilon_{1,1}, \ldots, \upsilon_{n,T})$ assumed to be multivariate-Gaussian. Both the local scores $g_{it}(X, W; \theta) = \{H_{(\phi,\eta)}{}^{-1}\left[X\tfrac{1}{\sigma}\beta + (I_T \otimes W_{(2)})X\tfrac{1}{\sigma}\gamma\right]\}_{it}$ and the disturbances $\upsilon_{i,t}$ are cross-sectionally interdependent. Additionally, the local scores $g_{it}(X, W; \theta)$ entail data from all observations. When we abstract for notational simplicity from the time dimension and use only the n-dimensional PDF, and arrange the observations according to the values of the observed LDV $\tilde{y}_i$, the likelihood can be written as

$$\mathcal{L} = \underbrace{\int_{-\infty}^{a_1 - g_1(X,W;\theta)} \cdots \int_{-\infty}^{a_1 - g_{n_0}(X,W;\theta)}}_{i:\tilde{y}_i=1} \cdots \underbrace{\int_{a_{j-1}-g_{n_{j-1}+1}(X,W;\theta)}^{a_j - g_{n_{j-1}+1}(X,W;\theta)} \cdots \int_{a_{j-1}-g_{n_j}(X,W;\theta)}^{a_j - g_{n_j}(X,W;\theta)}}_{i:\tilde{y}_i=j}$$
$$\cdots \underbrace{\int_{a_{J-1}-g_{n_{J-1}+1}(X,W;\theta)}^{\infty} \cdots \int_{a_{J-1}-g_n(X,W;\theta)}^{\infty}}_{i:\tilde{y}_i=J} f(\upsilon_1, \ldots, \upsilon_n) d\upsilon_1 \cdots d\upsilon_n \qquad (8)$$

In contrast to the case with independent observations, we are not able to split the likelihood into multiplicative-separable parts (or the log-likelihood into additive-separable). Hence, this likelihood is hardly solvable with conventional probabilistic (that is 'frequentistical') methods whereas Bayesian statistics may still provide a solution (see e.g. LeSage/Pace, 2009: Ch. 10; Franzese/Hays, 2009; Wang/Kockelman, 2009).[13]

A computationally simple approximation can be derived from a specification similar to eq. (4). Like in eq. (7), β and γ are identified only up to the scaling parameter σ. The serially and spatially lagged dependent (latent) variable is approximated by the serially lagged LDV and by the spatially weighted average[14] over the LDV, respectively. Since the re-scaled stochastic process is standard-normal by assumption, we divide these by the conditional standard deviation of $y_{i,t}$ estimated in the linear model (without fixed effects), such that both $\frac{1}{\sigma_{\tilde{y}}}\tilde{y}_{i,t-1}$ and $\sum_{j=1}^{n} w_{ij}\frac{1}{\sigma_{\tilde{y}}}\tilde{y}_{j,t}$ have unit variance. Then we estimate the latent process

$$y_{i,t} = \phi \tfrac{1}{\sigma_{\tilde{y}}}\tilde{y}_{i,t-1} + \eta \sum_{j=1}^{n} w_{(1)ij}\tfrac{1}{\sigma_{\tilde{y}}}\tilde{y}_{j,t} + x_{i,t}\tfrac{1}{\sigma}\beta + \sum_{j=1}^{n} w_{(2)ij}x_{j,t}\tfrac{1}{\sigma}\gamma + \tfrac{1}{\sigma}e_{i,t} \qquad (9)$$

with ordered probit for independent observations while instrumenting for the spatially lagged LDV with the same internal instruments as in the linear case.

The condition for the identification of peer effects – existence of nonlinear exclusion restrictions – is the same as in the linear case. Furthermore, nonlinear estimation requires additional assumption (e.g. on standardization) and computational challenges. Hence, we use the ordered probit estimations only as a robustness check with regard to the linear-form assumption.

[13] We adapt the MATLAB code for a spatial probit model provided by Jim LeSage and employ a Gibbs sampler to draw from a Truncated Multivariate Normal (TMVN) distribution when generating the latent variable. To achieve a positive definite multiplier $H_{(\phi,\eta)}$, we restrict ϕ to the interval $[-\frac{1}{\varphi}, \frac{1}{\varphi}]$ with φ the largest eigenvalue of $I_n - \eta W_n$; to keep it finite, we replace the multiplier matrix $H_{(\phi,\eta)}$ by $\tilde{H}_{(\phi,\eta)} = [(I_T - \phi L_T) \otimes (I_n - \eta W_n)]^{-1}$ if $|H_{(\phi,\eta)}| < 10^{-6}$. However, since $(\hat{\eta}, \hat{\phi})$ frequently end up at the joint frontier of the parameter space and the estimates are not stable, we do not present results here.

[14] We are aware that the (weighted) mean of the observed LDV is, in contrast to the median or any other quantile, not a well-defined statistic for categorial variables. Nevertheless, it might be a suitable approximation for the social expectation $E(y|p)$.

4 Expectation formation: Evidence for herding

4.1 Results

In the following we present results when estimating the linear model over the entire observation period. The focus of the analysis is the economic structure in expectation formation. Here we focus on the mere existence of herding and its distinction from other peer effects; enlightening the mechanism behind is postponed to a later section.

Table 3: Parameter estimates (Distance)

	Coefficient	Pure AR	X	X, MX	W_dY, X	W_dY, X, MX
α	cons	-0.013**	0.058***	0.120***	-0.008	0.005
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
ϕ	AR(1)	0.838***	0.707***	0.539***	0.444***	0.445***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β	U$_{t-1}$	—	-0.116*	0.595***	-0.070	0.311***
			(0.07)	(0.09)	(0.06)	(0.08)
	V$_{t-1}$	—	-0.639***	-0.151***	-0.087**	-0.091**
			(0.03)	(0.04)	(0.03)	(0.04)
γ	MU$_{t-1}$	—	—	-4.554***	—	-1.794***
				(0.23)		(0.26)
	MV$_{t-1}$	—	—	-2.617***	—	-0.647***
				(0.09)		(0.14)
η	SAR(1)	—	—	—	0.520***	0.459***
					(0.02)	(0.02)
δ	Fixed Eff.	yes^{a}	yes^{a}	yes^{a}	yes^{a}	yes^{a}
	Time Eff.	no	no	no	no	no
	SAR(1)ε_t			0.328	-0.373	-0.336
	AR(1)ε_t			-0.059	-0.032	-0.046
σ_e (RMSE)		.517	.501	.473	.451	.449
Sargan test					45.461***	9.763*
Partial R^2					0.753	.669
Wu-Hausman					5.414**	1.288

W_d: Binary distance, row st. – M: inverse distance – WY instrumented with WX, W^2X, WLY.
Standard errors in parenthesis – Asterisks mark standard significance levels (90%/ 95%/ 99%).
[a]: Region-specific fixed effects significant in ca. 10 regions.

The first four columns of Table 3 refer to estimations where the model components are included stepwise: we start with a pure autoregressive model, estimate then a model without peer effects (only conditioning on X and LY) and such where the peer-effect components (MX and WY) are included separately. The last column in the table shows parameter estimates for eq. (4) in which we model the contextual effect by a continuous inverse-distance decay matrix (denoted with M) whereas we employ a row-standardized binary cut-off distance matrix to describe the social-learning network. Later on, we will denote this specification as our preferred model. Estimations with alternative weights, modifications in the disturbance structure and nonlinear functional form accounting for the discreteness of Y (ordered probit) are reported in Table 4 in the robustness section.

A high AR term in general expresses that current disturbances or innovations have little effect on believes whereas there is long memory with regard to the history of innovations. The evidence for long memory with regard to errors can be interpreted twofold: as slow adaption to new information, or as strong effect of learning from personal mistakes, averaging over the entire history of errors. In the latter case, a moving-average process in the disturbance with a short lag length could be understood to describe CEOs repeating mis-believes without personal learning. We find that the parameter estimate associated with the AR term declines only little when we include additional information on the market fundamentals: including local information results in a parameter shift of -0.13 (the difference between $\hat{\phi}$ in the first and second column of Table 3), consideration of contextual information lets the parameter shift by an additional -0.168. If we include the endogenous peer effect which accounts for social learning, the parameter estimate decreases again by an additional -0.095 down to 0.445. This AR parameter value is robust across most specifications which account for the endogenous effect (for social learning) and the local market fundamentals, regardless of the further information included in the estimation (only the estimates in the random-effects and in the ordered probit IV model, see Table 4, deviate significantly from that value). This remainder of the serially-autoregressive process seems to reflect the influence of signals received up to the previous period. Serial correlation in the residual (shown in the rows denoted with δ: AR(1)ε_t) which supposedly captures the short-run effect of personal learning in expectation formation seems negligible.

We would like to briefly discuss the signals which can be drawn from observable market fundamentals.[15] An increase in local unemployment seems throughout most models correspond to expectations of unemployment rising further on; with exception of the model including only local information and the own history, its effect is insignificant or significantly positive. With regard to the effect associated with vacancies we can state that the sign is plausible insofar that unemployment is expected to rise when vacancies decline; furthermore the effect's height is stable across the models if we account at least for either the contextual or the endogenous peer effect. Vacancies in surrounding regions have an impact on unemployment sentiments which shows in the same direction. The size of the effect seems stronger at the first glance. However, the variation of this spatial average is much smaller than the variation of vacancies themselves, thus the larger size of the coefficient does not imply a relatively stronger impact. Contextual unemployment has, alike vacancies and in contrast to recent local unemployment, a negative sign. Declining unemployment in surrounding regions increases the probability that a CEO expects local unemployment to rise. This suggests a kind of crowding-out amongst the unemployed in close regions.

However, our central interest is on the endogenous peer effect, that is on the estimate for η; we interpret this as the effect of social learning (or herding). The estimate $\hat{\eta}$ in our preferred specification is 0.459. The corresponding estimates in the robustness checks are always significantly positive, mounting to approximately 0.4 to 0.5 throughout most specifications presented in Tables 3 and 4. A herding parameter of 0.459 corresponds to a social multiplier of $\frac{1}{1-\eta} = 1.85$ (when abstracting from mis-specification due to negligence of the limited definition range of the dependent variable). The spectral norm of the corresponding spatial multiplier matrix $(I - \eta W)^{-1}$ has an almost identical value; it is 1.86. That is, the direct signal extractable from variation in each variable in our preferred model is approxi-

[15] In an earlier version of the paper with data until May 2011, we found parameter estimates that could be translated into a matching function with an elasticity with respect to labour-market tightness of approximately 0.3. The current estimates do not support such clear evidence of a matching function.

mately doubled due to social interaction as follows from eq. (3); without social learning, the impulse of a market fundamental would require twice its amount to cause the same effect on the expectations.

Note that the parameter can be considered identified in terms of IV estimation in our preferred specification since the partial R^2 of the excluded instruments is relatively high whereas the Sargan statistics is just weakly significant. The Sargan test statistics declines even further when we do not use the second-order lags of unemployment and vacancies which are only week instruments in our preferred specification and which we only include for better comparison with the models in the robustness section. In contrast, the Sargan test rejects validity of the instruments in the model without contextual effect. This emphasizes that the contextual variables should be included in the second stage equation explaining Y and not only in the first stage explaining WY.

4.2 Robustness

In the following, we present a number of robustness checks. Table 4 report estimations of the linear model according to eq. (6a) where we use alternative networks (spatial structures) in the endogenous and the contextual effect, modify the disturbance structure, or employ an ordered probit rather than the linear specification.

The regressions in the first five columns differ from the preferred model with regard to the spatial weights matrices. The parameter estimates do in general not deviate strongly from the corresponding estimates in the last column of Table 3. The parameters in the contextual effect have smaller size when we use a row-normalized binary matrix (W_c, W_d, W_p) instead of the continuous distance-decay matrix M. This result is rather intuitive since the discrete weighting schemes result in less smooth network averages, i.e. show higher variation than those generated with m_{ij} as weights (though still less variation than the corresponding variable itself). The autoregressive parameters in these five models are not significantly different from our preferred specification. In contrast to this, the herding parameters show some deviations, even though the values are not far apart, in the range between 0.381 and 0.512. The major difference across the first five models is in the validity of instrumentation. The Sargan tests tend to be rejected if we use discrete geographical relations for both the endogenous and the contextual effect.

Columns six and seven show estimates in which we used alternative disturbance structures. The two-way error-component model with individual and time-specific fixed effects controls for cross-sectional error correlation by ruling out the time-specific average disturbance or the average factor dependence (since $\hat{\mu}_t = \sum_{i=1}^{n} \lambda_i f_t$). Significance of fixed effects in eight to ten from 176 regions – that is region-specific effects not deviating significantly from the average in more than 90% of the regions – suggests that excluding region dummies from estimation, or estimation with random effects will not cause serious bias. Indeed, most parameters are not too different. The major difference can be observed in the herding parameter which is significantly smaller than 0.459 in both estimations; nevertheless, both estimates for η are still significantly positive. Even if we control for cross-

Table 4: Parameter estimates under alternative weights matrices and disturbance structures

		geographical contiguity		Shared benchmarking class		Distance	Ind.+Time Eff.	Rand. Eff.	Ord. Probit (IV)^b
	Coefficient	$W_c y, X, MX$	$W_c y, X, W_c X$	$W_p y, X, MX$	$W_p y, X, W_p X$	$W_d y, X, W_d X$	$W_d y, X, MX$	$W_d y, X, MX$	(bootstrap s.e.)
ϕ	AR(1)	0.458***	0.463***	0.455***	0.461***	0.445***	0.437***	0.517***	0.599***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β	U_{t-1}	0.336***	0.192*	0.307***	0.154*	0.272***	0.336***	0.376***	0.079**
		(0.08)	(0.10)	(0.08)	(0.09)	(0.10)	(0.09)	(0.08)	(0.04)
	V_{t-1}	-0.099***	-0.116***	-0.086**	-0.102***	-0.095***	-0.087**	-0.088***	-0.109***
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)
γ	WU_{t-1}	-2.317***	-0.648***	-2.155***	-0.572***	-0.632***	-1.335***	-2.028***	-0.440***
		(0.26)	(0.13)	(0.25)	(0.13)	(0.14)	(0.46)	(0.25)	(0.07)
	WV_{t-1}	-1.020***	-0.235***	-0.915***	-0.172**	-0.109	-0.624***	-0.713***	0.569***
		(0.13)	(0.07)	(0.13)	(0.08)	(0.08)	(0.19)	(0.13)	(0.10)
η	SAR(1)	0.381***	0.453***	0.410***	0.480***	0.512***	0.287***	0.381***	0.542***
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.02)	(0.02)
δ	Fixed Eff.	yes^a	yes^a	yes^a	yes^a	yes^a	yes^a	no	—
	Time Eff.	no	no	no	no	no	yes	no	—
	SAR(1)ε_t	-.276	-.351	-.150	-.229	-.376	-.336	-.186	—
	AR(1)ε_t	-.039	-.039	-.039	-.04	-.035	-.046	-.107	—
σ_e (RMSE)		0.454	0.456	0.447	0.449	0.451	0.441	0.461	—
N		7216	7216	7216	7216	7216	7216	7216	7216
Partial R^2		.480	.562	.610	.623	.613	.301	.599	.554
Sargan test		2.134	15.340***	5.041	5.959*	22.418***	10.091*	4.677	
Wu-Hausman		6.076**	13.514***	11.976***	9.811***	3.434*	21.878***	0.176	—

W_c: Binary contiguity, row st. – W_p: Binary, same SGB3-benchmark type, row st. – W_d: Binary distance, row st. –M: inverse distance (continuous) – Wy instrumented with WX, W^2X, WLy.

Standard errors in Parenthesis – Asterisks mark standard significance levels (90%/ 95%/ 99%).

[a]: Region-specific fixed effects significant in ca. 10 regions.

[b]: Threshold values between categories at -1.975, 1.249, 3.514 and 6.049.

sectional correlation in a very rigorous way by time-specific dummies the estimate for the herding parameter mounts to a significantly positive value of 0.287.

The ordered-probit estimation – which is identified under the assumption that the standard-deviation of the discrete dependent variable conditional on X is 0.449 (the estimate for $\sigma_{\tilde{y}|X} = \sigma_e$ in our preferred model) – deviates from the linear estimations with regard to two points. First, we find the highest estimates for both the serially and the cross-sectionally autoregressive parameters, ϕ and η, amongst all models accounting for the entire information about local and contextual market fundamentals. Second, the sign associated with vacancies in surrounding regions reverses; for us, it seems implausible that increasing vacancies should be associated with unemployment expected to rise.

5 Social learning or joint adaption to news?

Carroll (2003) presents a model where the average sentiment at time t on unemployment in $t+h$, $\bar{Y}_t[u_{t+h}]$ in a survey of interest depends to a fraction λ on current-period public information $N_t[u_{t+h}]$ (e.g. published in newspapers) and to an amount $(1-\lambda)$ to prospective information which has already been available in the previous period. Recursion leads to the average expectation

$$\begin{aligned}\bar{Y}_t[u_{t+h}] &= \lambda N_t[u_{t+h}] + (1-\lambda)\left(\lambda N_{t-1}[u_{t+h}] + (1-\lambda)\left(\lambda N_{t-2}[u_{t+h}] + \ldots\right)\right)\\ &= \lambda N_t[u_{t+h}] + (1-\lambda)\bar{Y}_{t-1}[u_{t+h}]. \qquad (10)\end{aligned}$$

Carroll (2003) interprets λ as the fraction of a population that receives news. However, the model can be adapted to individual sentiments wherein λ reflects an individual's probability to get new information rather than the informed-population's share; Carroll's population model can be derived by averaging across independent individuals. Public information can be integrated into a simple model of expectation formation in addition to market fundamentals and other variables. This allows us to identify the partial contribution of public information to individuals' sentiments.

Hence, in order to compare the effect of public information and social learning, we estimate two equations (both two times, either with public forecasts released in the current or with those published in the previous month): In the first, we explain current unemployment expectations by the local log unemployment stock and log vacancies observed in the previous period, lagged expectations and by the news regarding unemployment in the future.[16] In the second, we augment this model by the contextual and the endogenous peer effect, that is by spatially lagged market fundamentals and the spatially lagged expectations; with regard to public information, there won't be a contextual peer effect since it is in principle

[16] The institutes do not forecast unemployment at the same horizon and the same frequency that we have in the FEA management survey. Hence, we construct our 'public information variable' as follows: We first average forecasts published in the same month; in each month, we then consider only those forecasts published latest (if no new forecast is available, we extrapolate the average forecast from the previous month). Since most institutes publish forecasts for the current and the subsequent year, we use in the first six months of a year the forecast for the respective year and from July to September a weighted average between the forecasts regarding current (linearly declining weight) and next year (linearly increasing weight).

observable to all persons in the survey. Note that in this model the parameter associated with the public-information variable will only reflect the probability that a person will directly notice the new information; this effect may be amplified by social learning from informed persons. Results are reported in Table 5; for means of comparison we add the estimates of our preferred model from Table 3.

Table 5: Estimation: Public information vs. social learning

Param.	Variable	Herding (SL)	Public Info (PI)		Combined (PI+SL)	
α	cons	0.005	-0.724***	-0.412***	-0.107	0.091
		(0.01)	(0.12)	(0.12)	(0.12)	(0.11)
ϕ	AR(1)	0.445***	0.526***	0.529***	0.444***	0.445***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β	U_{t-1}	0.311***	0.594***	0.617***	0.315***	0.306***
		(0.08)	(0.09)	(0.09)	(0.08)	(0.08)
	V_{t-1}	-0.091**	-0.138***	-0.146***	-0.090**	-0.091**
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
γ	MU_{t-1}	-1.794***	-4.379***	-4.731***	-1.816***	-1.750***
		(0.26)	(0.23)	(0.23)	(0.26)	(0.27)
	MV_{t-1}	-0.647***	-2.274***	-2.497***	-0.631***	-0.643***
		(0.14)	(0.10)	(0.09)	(0.14)	(0.14)
η	SAR(1)	0.459***			0.451***	0.463***
		(0.02)			(0.03)	(0.03)
λ	N_t		0.251***		0.036	
			(0.04)		(0.04)	
	N_{t-1}			0.157***		-0.017
				(0.03)		(0.03)
δ	Fixed Eff.	yes	yes	yes	yes	yes
	SAR(1)	-.319	.324	.326	-.295	-.317
	AR(1)	-.037	-.049	-.054	-.036	-.037
RMSE		0.449	0.472	0.473	0.449	0.449
Part. R^2		.554	—	—	.525	.533
Sargan		9.763**	—	—	9.486*	9.891**
Wu-Hausman		1.288	—	—	0.862	1.512

Standard errors in parenthesis – Asterisks mark standard significance levels (90%/ 95%/ 99%).
The spatially lagged dependent variable WY is instrumented with WX, W^2X, WLY.

We find that the effect of contemporaneous public information is in general larger than the effect of the one-month lagged information; however, both are significantly positive. The size of the estimate $\hat{\lambda} = 0.251$ in the second column of Table 5 – the estimation which is most similar to eq. 10 in Carroll (2003) – is even smaller than his estimate for the effect of public information (which mounts to 0.36), and far from the corresponding estimate (mounting to 0.851) in Curtin (2003) which refers as well to individual rather than averaged expectations.

The effect vanishes completely when we account for herding. Note that the two parameters η and λ reflect only the partial effects of social learning and public information. The estimates for the herding parameter have a similar size as in our preferred model from Section 4.2 (without public information); in contrast, $\hat{\lambda}$ is not significantly different from zero. If we use lagged news rather than contemporaneous, we get the same result. This indicates that most of the (direct) effect which had been attributed to public information before is in fact due to social learning. However, the spillover of information through social networks even works as an amplifying device for public information.

6 Is there an informational cascade?

So far we have found evidence for a significant amount of herding in the unemployment expectations reported in the FEA management survey. The exact nature is however unclear: Does social learning guide a herd towards the correct outcome as it could be expected because market fundamentals form continuous signals (Smith/Sørensen, 2000) or because repeated announcements reduce the information set (Manski, 2004)? Or does it behave like an informational cascade (Bikhchanandi/Hirshleifer/Welch, 1992, 1998; Chamley, 2004) in which a herd may converge not only to a correct but even to an incorrect herd belief (or at least remain at an incorrect expectation for a long time and converge just slowly)? The two are different insofar that private information becomes at most irrelevant within a cascade, and that a cascade is fragile:[17] It can be broken by a person with strong believe in her private information or by public information made available after the cascade has started. We will employ this relation – public announcement that the economy develops in a new direction which supposedly causes a drastic change of individual believes with regard to future unemployment (without sharp changes in real unemployment/vacancies data) – to investigate the existence of an informational cascade.

The theoretical concept of an information cascade can be modeled empirically as a pair of structural breaks: Before the arrival of new external information, private information is of little importance whereas social learning should be dominant (that is, η should be high). During the period τ_1 in which, supposedly, public information becomes available and the subsequent adaption takes place, social learning pales in comparison to unobservable but even observable private information. The adjustment process presumably lasts for more than a month but not too long, somewhat between a quarter or half a year. In the final stage (starting in τ_2) a new informational cascade establishes. Social learning again dominates the direct influence of a CEO's own private information. However, the level of the social equilibrium, that is the average expectation, should have changed between the initial and the new cascade. In addition, the residual dispersion should be wider in the period when public information arrives since it reflects not only the innovations themselves but also the uncertainty about information. Thus, we model the two structural breaks explicitly and estimate the following equation without fixed effects:

$$
\begin{aligned}
y_{i,t} = e_{i,t} &+ \mathbf{I}(t < \tau_1)_{i,t} \left(\alpha_1 + x_{i,t}\beta_1 + \sum_{j=1}^{n} m_{ij} x_{j,t} \gamma_1 + \sum_{j=1}^{n} w_{ij} y_{j,t}\, \eta_1 + \phi_1 y_{i,t-1} \right) \\
&+ \mathbf{I}(\tau_1 \leq t < \tau_2) \left(\alpha_2 + x_{i,t}\beta_2 + \sum_{j=1}^{n} m_{ij} x_{j,t} \gamma_2 + \sum_{j=1}^{n} w_{ij} y_{j,t}\, \eta_2 + \phi_2 y_{i,t-1} \right) \\
&+ \mathbf{I}(\tau_2 \leq t) \left(\alpha_3 + x_{i,t}\beta_3 + \sum_{j=1}^{n} m_{ij} x_{j,t} \gamma_3 + \sum_{j=1}^{n} w_{ij} y_{j,t}\, \eta_3 + \phi_3 y_{i,t-1} \right) \qquad (11)
\end{aligned}
$$

[17] Bikhchanandi/Hirshleifer/Welch (1992) write that "*conceptually, [their] paper differs from Welch's and Banerjee's in emphasizing the fragility... cascades can explain not only uniform behavior but also drastic change such as fads.*"

We consider January 2010 as the month when new information is published, since at this date the first institute revised its unemployment forecast from *rising* to *remaining equal*, see Section 2. Results for eq. (11) estimated analogously to our preferred specification (albeit with a reduced set of instruments) are shown in Table 6.

Table 6: Informational cascade – Structural break estimation

Parameter	Variable	$t < 01/2010$	$01 - 06/2010$	$07/2010 \leq t$
$\alpha_{I(t\in[t_1,t_2))}$		0.079	-0.087	-0.088
		(0.09)	(0.08)	(0.08)
ϕ	AR(1)	0.442***	0.412***	0.446***
		(0.02)	(0.02)	(0.02)
$\beta_{I(t\in[t_1,t_2))}$	U_{t-1}	0.575***	0.436*	-0.040
		(0.14)	(0.23)	(0.15)
	V_{t-1}	-0.063	0.005	-0.139**
		(0.06)	(0.09)	(0.06)
$\gamma_{I(t\in[t_1,t_2))}$	MU_{t-1}	-1.842***	2.042	-1.457**
		(0.44)	(1.52)	(0.66)
	MV_{t-1}	-0.421	-1.426***	-0.509**
		(0.26)	(0.64)	(0.22)
$\eta_{I(t\in[t_1,t_2))}$	WY	0.414***	0.110	0.393***
		(0.06)	(0.14)	(0.05)
IV statistics (for instrumenting WY with WX, WLY):				
Partial R^2		0.561	0.161	0.446
Sargan test		J = 7.638	$\sim \chi^2_6$	(p-val. = 0.266)
Wu-Hausman		H = 5.975	$\sim F_{3,7017}$	(p-val. = 0.000)

Standard errors in parenthesis. Asterisks mark standard significance levels (90%/ 95%/ 99%). Estimation deviates from baseline model with regard to instruments and error components: carried out without fixed effects and without second-order spatial lags of exogenous variables.

We find that information/belief persistence does not change over time: we can not reject that $\hat{\phi}_1 = \hat{\phi}_2 = \hat{\phi}_3$ at reasonable significance levels. The influence of market fundamentals is at most equal between period one (before the first break) and period three (after the second break); only the influence of local unemployment changes substantially. In contrast, most parameters associated with unemployment or vacancies are significantly different between periods two and three. The parameter estimates determining the average sentiment ($\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3$) are not significantly different from zero. However, equality of $\hat{\alpha}_2, \hat{\alpha}_3$ with 0.079, the point estimate of α_1, can be rejected at least weakly. In accordance with our theoretical considerations regarding an informational cascade, social learning is less important in the break period: $\hat{\eta}_1$ and $\hat{\eta}_3$ are significantly positive, in contrast to $\hat{\eta}_2$. Moreover, $\hat{\eta}_2$ is smaller than the further at the 90% confidence level. Thus, we observe social learning both before the collapse and a while after once a new cascade might have established, but find hardly evidence for herding at the time of the potential collapse and short after.

To illuminate the behaviour of unobservable information, we look at the distribution of innovations (or disturbances) resulting from estimation of eq. (11). The left panel in Fig. 3 plots the disturbances by month, with the first and third quartile highlighted. The right panel reports the innovations' time-specific standard deviation. Both figures show that the distribution becomes slightly wider dispersed in the first half-year of 2010: The standard deviation exceeds 0.5 in more than a single month only from February till May 2010. I.e., unexplained variation is stronger, private (unobserved) signals are more outstanding in this period than before January 2010 or after June 2010.

Our findings give in general support towards the fragility of herding amongst CEOs, that is, towards the existence of informational cascades. At the time when new information regarding oppositely directed future development of the labor market is published, social learning becomes insignificant. Then, the herding parameter is significantly smaller than in the periods both before that date and some months later. As well, the dispersion of (only ex-post measurable) private signals becomes wider.

However, it might be difficult to identify (or reject) the existence of a cascade in real-world data because of various reasons. On the one hand, achieved statistical significance is affected by a longer observation period since parameters converge over both n and T. Hence, the parameters corresponding to the short period between two cascades are necessarily estimated relatively imprecise.

On the other hand, shifts in the expectations might not only be due to the arrival of public *information* on a trend reversal and the subsequent shift in social beliefs but also due to the trend reversal itself, that is due to a shift of the world's *true state* (e.g. the shift from a recession to an upturn). Then, breaks in the parameters would result from the new correct value and not from the new information. Nonetheless, it is unlikely that a new correct value of unemployment has formed exactly in early 2010 because of three reasons: first, unemployment declined, seasonally adjusted, at a smooth rate in the time between Spring 2009 and January 2012 – that is, we do not observe a trend reversion in realized unemployment. Second, the quarterly GDP growth rate has been in a range between 0.5 and one from the second quarter 2009 till 2011, with exception of the second quarter 2010 where it mounted to approximately two.[18] And third, published GDP growth rate forecasts were revised substantially in and after June 2010, from values between one and two percent to more than three percent. Research on Okun's Law in Germany from the 1990s (though admittedly outdated) assigns the unemployment threshold of the output gap to GDP growth rates between 1.5 and two percent. According to this theory, unemployment should have remained stable in the first half of 2010, a further decline in unemployment expectable only in late 2010. If the reversion of sentiments would be due to a change in the real world, the direction of herding would either have followed the correct

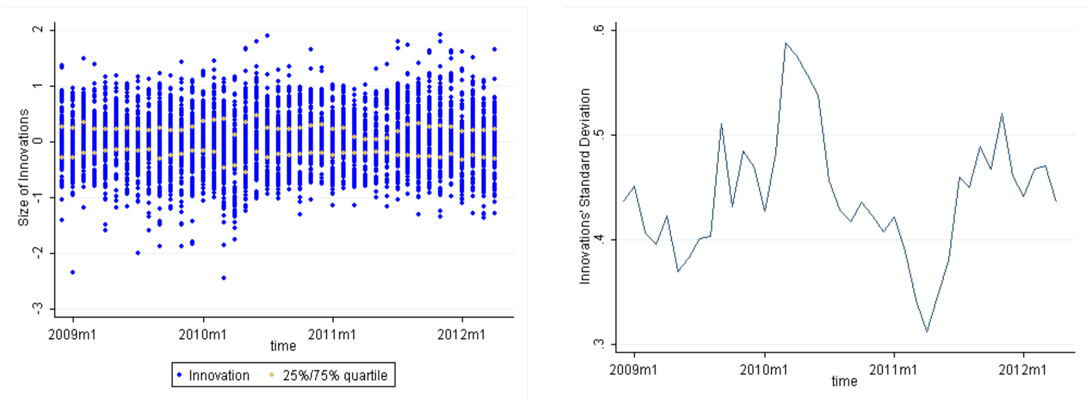

Figure 3: Dispersion of innovations over time

[18] See the national accounts of Germany published by the Federal Statistical Office .

value's shift with a delay of three quarters, or it would have anticipated this shift. Hence, a causation of the sentiments' reversion by a shift in the published unemployment forecasts, that is by information, is more plausible.

7 Are sentiments more rational due to herding?

Another topic of interest is whether the expectations are rational (or realistic), and whether social learning contributes to make sentiments more realistic: Despite public conviction, herding might contribute to the rationality of expectations through social aggregation of information. In order to analyse if the forecasts deviate systematically from the bisector in a prediction-realization diagram – as it has been proposed by Mincer/Zarnowitz (1969) besides several other criteria for the evaluation of expectations and forecasts – the realized values u_{t+h} are regressed on the forecasts $Y_t[u_{t+h|\mathcal{I}_t}]$ and a constant (maintaining our previous notation):

$$u_{t+h} = \vartheta_1 + \vartheta_2\, Y_t[u_{t+h|\mathcal{I}_t}] + \upsilon_{t+h} \tag{12}$$

The estimated slope $\hat{\vartheta}_2$ is tested against the value 1, the intercept $\hat{\vartheta}_1$ against 0. Here, we rely on a similar concept. A first graphical inspection of the relation between sentiments as reported in the survey and realized unemployment development has been provided in Section 2. The estimates from a Mincer-Zarnowitz (MZ) regression will be shown subsequently.

However, for answering the question in this section, this test will have only limited use when applying it directly to the sentiments in the survey. The results in the previous sections have provided evidence for social learning in the formation of sentiments. Hence, in order to analyze whether social learning really contributes to more rational expectations we need to construct counterfactual sentiments in which we assume isolation of agents. In these hypothetical sentiments the effect of social learning is eliminated. To get a direct counterpart to this counterfactual, we construct as well hypothetical sentiments with social learning. For both, we use the estimates of eq. (4) in our preferred specification, and in both we include even the estimated innovations $\hat{e}_{i,t}$ as a measure for the private signals. Thus, we consider the following prediction as our hypothetical sentiments under social learning:

$$Y^{(s)}_{i,t}[u_{t+h|X,MX,WY}] = \left\{ (H_{\hat{\eta},\hat{\phi}})^{-1} (\iota_{Tn}\hat{\alpha} + X\hat{\beta} + (M_n \otimes I_T)X\hat{\gamma} + \hat{e}) \right\}_{i,t} \tag{13}$$

The counterfactual under isolation uses the same parameter estimates. The social multiplier $(I - \hat{\eta}W)^{-1}$ which represents the simultaneous interaction between responses is however replaced by its spectral norm, a representation of the scale but not the interaction of the social multiplier. That is, the remaining score (which still accounts for a contextual peer effect) is multiplied with a constant such that the variance of the predictor under isolation is comparable to the predictor under social learning:

$$\begin{aligned} Y^{(i)}_{i,t}[u_{t+h|X,MX}] = \Big\{ &\|(I_n - \hat{\eta}W_n)^{-1}\| \cdot \left(I_n \otimes (I_T - \hat{\phi}L_T)^{-1} \right) \\ &\times (\iota_{Tn}\hat{\alpha} + X\hat{\beta} + (M_n \otimes I_T)X\hat{\gamma} + \hat{e}) \Big\}_{i,t} \end{aligned} \tag{14}$$

In the presented results, both hypothetical expectations $Y^{(i)}[u_{t+h|X,MX}]$ and $Y^{(s)}[u_{t+h|X,MX,WY}]$ are rounded to the next integer in $[-2, 2]$; that is, they have the same ordinal scale as the expectations from the survey. The two ordinal hypothetical expectations are then used as regressors in MZ regressors. We will consider social learning as improving forecast accuracy if the slope in the regression with the 'hypothetical herding' is significantly closer to 1 than the slope coefficient under 'hypothetical isolation' (without the intercept significantly more distinct from 0), or vice versa.

The question in the survey refers to unemployment's development within the next three months (besides the typical seasonal development). Thus, we use the three-month difference between annual growth rates of unemployment three months ahead and today, $u_{t+3} = \Delta[\frac{(U_{t+3}-U_{t-12+3})}{U_{t-12+3}} - \frac{(U_t-U_{t-12})}{U_{t-12}}] \cdot 100$ to construct the realization variable for the MZ regression. However, since the expectation variables – both the survey responses and the hypothetical expectations – are ordinal, we use an ordinal scale (likewise coded from -2 to $+2$) where we define a growth-rate difference (in percentage points) in the interval $(-2.5, +2.5)$ as no-change, in the interval from $(-12.5, 2.5]$ and $[2.5, 12.5)$, respectively, as moderate change (decline or growth), and growth-rate differences up to -12.5 (or above 12.5 percentage points) as strong change.

Table 7: Realized development vs. hypothetical sentiments

$\tilde{u}$	Social Learning $Y^{(s)}[u\vert_{X,MX,WY}]$					Isolation $Y^{(i)}[u\vert_{X,MX}]$					Sum
	-2	-1	0	1	2	-2	-1	0	1	2	
2	1	34	5	106	**143**	6	30	6	113	**134**	289
1	66	802	254	**435**	276	104	750	289	**411**	279	1833
0	135	1135	**580**	449	77	154	1059	**682**	383	98	2376
-1	50	**594**	766	397	46	73	**564**	793	352	71	1853
-2	**0**	21	205	102	9	**2**	26	180	100	29	337
	252	2586	1810	1489	551	339	2429	1950	1359	611	

Table 7 crosses the two hypothetical expectation with the realized values; the values on the bisector are in bold letters. At the first glance, there is hardly any relationship detectable between hypothetical expectations and realized values – whereas we can replicate the survey responses with our predictions (under herding) pretty well. If at all, the cross-tables hint at a more or less horizontal pattern, that is sentiments which are uncorrelated with realized unemployment development.

The results of the corresponding MZ regression are shown in Table 8. The sentiments – both the survey responses and the two hypothetical sentiments – are far from perfect as can be seen from Table 8 and as we have already supposed for the survey responses in Sec. 2. The MZ hypothesis is rejected in each case. The slope in the two regressions with the discrete dependent variable is positive but closer to zero than to one. However, the slope under hypothetical social learning is significantly steeper than the slope under hypothetical isolation.

A horizontal line would express that there isn't, on average, any relation between realized values and expectations at all; a slope of one would indicate unbiased (that is rational) expectations. Since the slope of the hypothetical sentiments under social learning is closer

Table 8: Mincer-Zarnowitz regressions

Coefficient	Survey responses	Hypothetical social learning	Hypothetical isolated
	Realized growth (ordinal)		
Intercept $\hat{\vartheta}_1$	-.024	-.009	-.011
	(.01)	(.01)	(.01)
Slope $\hat{\vartheta}_2$	.045	.105	.085
	(.01)	(.01)	(.01)

Robust standard errors in parentheses.

to one than the hypothetical sentiments under isolation we conclude that social learning improves the rationality of sentiments.

8 Conclusion

In this paper we have analysed unemployment expectations with a particular focus on social learning in expectation formation. A novel survey amongst the CEOs of the local departments of the German Federal Employment Agency allows us to employ geographical structure and organizational networks in order to discriminate between close and less-close peers in communication and social learning. Thus, we have been able to deal with the reflection problem in the empirical analysis of social networks which allowed us to identify the effect of social learning and herding in the formation of unemployment expectations.

We have presented evidence for the influence of socially aggregated expectations in individual's expectation formation in Section 4; the results have been robust across various specifications. The estimate for the social multiplier in our preferred specification mounts to 1.86; that is, social interaction approximately doubles the impact on the average expectation which we have assessed directly to observable or private information. We have found that this effect does persist with a similar size when we account for public information on unemployment forecasts. The social multiplier still mounts to approximately 1.5 if we account for other contemporaneous effects in a rigorous way.

We have only to some extent been successful in detecting the nature of social information aggregation inherent in the survey's responses. First, we have rejected that the CEOs's sentiments only mimic unemployment forecasts published by professional forecasters in Germany; the estimated contribution of herding in sentiment formation is robust against controlling for the latter. Second, we have found a shift in believes accompanied by higher private uncertainty (or wider dispersion of unobservable private signals) and increased impact of observable information in the first half-year of 2010 – in succession to the economic research institutes' predictions that Germany has passed most of the crisis. This pattern hints at the existence of an informational cascade in the CEOs' announced unemployment expectations. However, our last result – that expectations do at least weakly become more realistic due to herding – is not systematically conformable with cascading of sentiments. The latter provides good news for economic tendency surveys amongst experts insofar that, despite the threat of a misleading cascade, they seem to aggregate information efficiently.

In future research – if the data cover more than one business-cycle turn and more than one expectation reversal – it might become possible to reject the cascading hypothesis, to provide more robust evidence for the informational efficiency of social learning, or to identify the origin of a herd: be it by pursuing a similar strategy to ours, or (if the data is much longer) by making a more detailed use of the timing of events. The evidence for social learning in unemployment expectations could be strengthened on the one hand if we can ascertain a similar endogenous peer effect in the further questions enclosed in the survey; supposedly, social learning amongst the same agents covers more than one issue. On the other hand, it would be interesting to verify that the responses are honest, i.e. that the expectations are followed by the adequate actions. However, both is beyond the scope of this paper. Finally, when setting up other surveys on expectations, in particular amongst experts or other small groups with a high probability of interaction, the information extracted from their responses may be improved by adequately accounting for social learning. For this, it is however necessary not only to ask for their expectations but even to collect (and provide) some information on their networks.

References

Acemoglu, Daron; Dahleh, Munther A.; Lobel, Ilan; Ozdaglar, Asuman (2011): Bayesian Learning in Social Networks. In: Review of Economic Studies, Vol. 78, p. 1201–1236.

Acemoglu, Daron; Ozdaglar, Asuman (2011): Opinion Dynamics and Learning in Social Networks. In: Dynamic Games and Applications, Vol. 1, p. 3–49.

Anselin, Luc (2001): Spatial Econometrics. In: Baltagi, B. (Ed.) A Companion to Theoretical Econometrics, Companions to Contemporary Economics, chap. 14, Malden (MA): Blackwell, p. 310–330.

Anselin, Luc (1988): Spatial Econometrics: Methods and Models. Dordrecht: Kluwer Academic Publishers.

Bajari, Patrick; Krainer, John (2004): An Empirical Model of Stock Analysts' Recommendations: Market Fundamentals, Conflict of Interest, and Peer Effects. Working Paper 10665, NBER.

Baltagi, Badi H. (2006): Random Effects and Spatial Autocorrelation with Equal Weights. In: Econometric Theory, Vol. 22, p. 973–984.

Baltagi, Badi H.; Liu, Long (2009): Spatial Lag Tests with Equal Weights. In: Economics Letters, Vol. 104, p. 81–82.

Banerjee, Abhijit V. (1992): A Simple Model of Herd Behavior. In: Quarterly Journal of Economics, Vol. 107, No. 3, p. 797–817.

Bewley, Ronald; Fiebig, Denzil G. (2002): On the Herding Instinct of Interest Rate Forecasters. In: Empirical Economics, Vol. 27, p. 403–425.

Bikhchanandi, Sushil; Hirshleifer, David; Welch, Ivo (1998): Learning from the Behavior of Others: Conformity, Fads, Informational Cascades. In: Journal of Economic Perspectives, Vol. 12, p. 151–170.

Bikhchanandi, Sushil; Hirshleifer, David; Welch, Ivo (1992): A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades. In: Journal of Political Economy, Vol. 100, p. 992–1026.

Blien, Uwe; Hirschenauer, Franziska; Thi Hong Van, Phan (2010): Classification of Regional Labour Markets for Purposes of Labour Market Policy. In: Papers in Regional Science, Vol. 89, p. 859–880.

Bramoullé, Yann; Djebbari, Habiba; Fortin, Bernard (2009): Identifiction of Peer Effects through Social Networks. In: Journal of Econometrics, Vol. 150, p. 41–55.

Brock, William A.; Durlauf, Steven N. (2003): Multinomial Choice with Social Interactions. Technical Working Paper 288, NBER.

Brock, William A.; Durlauf, Steven N. (2001): Discrete Choice with Social Interactions. In: Review of Economic Studies, Vol. 68, p. 235–260.

Carroll, Christopher D. (2003): Macroeconomic Expectations of Households and Professional Forecasters. In: Quarterly Journal of Economics, Vol. 118, p. 269–298.

Çelen, Boğaçhan; Kariv, Shachar (2004): Distinguishing Informational Cascades from Herd Behavior in the Laboratory. In: American Economic Review, Vol. 94, No. 3, p. 484–498.

Chamley, Christophe P. (2004): Rational Herds: Economic Models of Social Learning. Cambridge (UK), New York (NY): Cambridge University Press.

Chroushore, Dean (2010): An Evaluation of Inflation Forecasts from Surveys Using Real-Time Data. In: The B.E. Journal of Macroeconomics, Vol. 10, No. 1.

Chroushore, Dean (1998): Evaluating Inflation Forecasts. working paper 98-14, Federal Reserve Bank of Philadelphia.

Chudik, Alexander; Pesaran, M. Hashem; Tosetti, Elisa (2011): Weak and Strong Cross Section Dependence and Estimation of Large Panels. In: Econometrics Journal, Vol. 14, No. 1, p. C45–C90.

Curtin, Richard T. (2003): Unemployment Expectations: The Impact of Private Information on Income Uncertainty. In: Review of Income and Wealth, Vol. 49, No. 4, p. 539–554.

De Giorgi, Giacomo; Pellizari, Michele; Redaelli, Silvia (2010): Identification of Social Interactions through Partially Overlapping Peer Groups. In: American Economic Journal: Applied Economics, Vol. 2.

Franzese, Robert J. Jr.; Hays, Jude (2009): The Spatial Probit Model of Interdependent Binary Outcomes: Estimation, Interpretation, and Presentation, presented at the Annual Meeting of the Public Choice Society, 6 Mar 2009.

Gibbons, Stephen; Overman, Henry G. (2012): Mostly Pointless Spatial Econometrics? In: Journal of Regional Science, Vol. 52, No. 2, p. 172–191.

Hey, John D. (1994): Expectations formation: Rational or adaptive or ...? In: Journal of Economic Behavior & Organization, Vol. 25, p. 329–349.

Kapoor, Mudit; Kelejian, Harry H.; Prucha, Ingmar R. (2007): Panel Data Models with Spatially Correlated Error Components. In: Journal of Econometrics, Vol. 140, No. 1, p. 97–130.

Keane, Michael P.; Runkle, David E. (1990): Testing the Rationality of Price Forecasts: New Evidence from Panel Data. In: American Economic Review, Vol. 80, No. 4, p. 714–735.

Kelejian, Harry H.; Prucha, Ingmar R. (2002): 2SLS and OLS in a Spatial Autoregressive Model with Equal Spatial Weights. In: Regional Science and Urban Economics, Vol. 32, p. 691–707.

Kelejian, Harry H.; Prucha, Ingmar R. (1998): A Generalized Spatial Two Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. In: Journal of Real Estate Finance and Economics, Vol. 17, p. 99–121.

Lamont, Owen (2002): Macroeconomic Forecasts and Microeconomic Forecasters. In: Journal of Economic Behavior & Organization, Vol. 48, No. 3, p. 265–280.

LeSage, James; Pace, R. Kelley (2009): Introduction to Spatial Econometrics. Boca Raton (FL), New York (NY): CRC Press.

Manski, Charles F. (2004): Social Learning from Private Experience: The Dynamics of the Selection Problem. In: Review of Economic Studies, Vol. 71, p. 443–458.

Manski, Charles F. (1993): Identification of Endogenous Social Effects: The Reflection Problem. In: Review of Economic Studies, Vol. 60, No. 3, p. 531–542.

Mayor, Matias (Sr.); Patuelli, Roberto (2012): Short-Run Regional Forecasts: Spatial Models Through Varying Cross-Sectional and Temporal Dimensions. Quaderni DSE working paper 835, University of Bologna.

Mincer, Jacob A.; Zarnowitz, Victor (1969): The Evaluation of Economic Forecasts. In: Mincer, Jacob A. (Ed.) Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance, chap. 1, Cambridge: NBER, p. 1–46, URL http://www.nber.org/chapters/c1214.

Mutl, Jan; Pfaffermayr, Michael (2011): The Hausman Test in a Cliff and Ord Panel Model. In: Econometrics Journal, Vol. 14, p. 48–76.

Nickell, Stephen J. (1981): Biases in Dynamic Models with Fixed Effects. In: Econometrica, Vol. 49, p. 1417–1426.

Ottaviani, Marco; Sørensen, Peter N. (2006): The Strategy of Professional Forecasting. In: Journal of Financial Economics, Vol. 81, No. 2, p. 441–466.

Pons-Novell, Jordi (2003): Strategic Bias, Herding Behaviour and Economic Forecasts. In: Journal of Forecasting, Vol. 22, p. 67–77.

Rangvid, Jesper; Schmeling, Maik; Schrimpf, Andreas (2012): What do Professional Forecasters' Stock Market Expectations tell us about Herding, Information Extraction and Beauty Contests? http://staff.cbs.dk/JRangvid/HOE.pdf.

Rüb, Felix; Werner, Daniel (2008): "Den Arbeitsmarkt" gibt es nicht – Arbeitsmarktregionen des SGB II im Vergleich. In: Jahrbuch für Regionalwissenschaft, Vol. 28, p. 93–108.

Rülke, Jan-Christoph; Tillmann, Peter (2011): Do FOMC members herd? In: Economics Letters, Vol. 113, No. 2, p. 176–179.

Schanne, Norbert (2010): Forecasting Regional Labour Markets with GVAR Models and Indicators. http://www-sre.wu.ac.at/ersa/ersaconfs/ersa10/ERSA2010finalpaper1044.pdf.

Schanne, Norbert; Wapler, Rüdiger; Weyh, Antje (2010): Regional Unemployment Forecasts with Spatial Interdependencies. In: International Journal of Forecasting, Vol. 26, p. 908–926.

Scharfstein, David S.; Stein, Jeremy C. (1990): Herd Behavior and Investment. In: American Economic Review, Vol. 80, p. 465–479.

Smith, Lones; Sørensen, Peter (2000): Pathological Outcomes of Observational Learning. In: Econometrica, Vol. 68, No. 2, p. 371–398.

Soetevent, Adriaan R. (2006): Empirics of the Identification of Social Interactions; An Evaluation of the Approaches and their Results. In: Journal of Economic Surveys, Vol. 20, p. 193–228.

Tortorice, Daniel L. (2012): Unemployment Expectations and the Business Cycle. In: The B.E. Journal of Macroeconomics, Vol. 12, No. 1:Topics.

Wang, Xiaokun; Kockelman, Kara M. (2009): Bayesian Inference for Ordered Response Data with a Dynamic Spatial Ordered Probit Model. In: Journal of Regional Science, Vol. 49, No. 5, p. 877–913.

Welch, Ivo (2000): Herding among Security Analysts. In: Journal of Financial Economics, Vol. 58, p. 369–396.

Zarnowitz, Victor; Lambros, Louis A. (1987): Consensus and Uncertainty in Economic Prediction. In: Journal of Political Economy, Vol. 95, No. 3, p. 591–621.

A Additional Information regarding the Survey

Table 9: Availability of Questions in the FEA Management Survey (April 2012)

Question	Availability	Items/Scale
■ How do you expect unemployment in your district to develop within the next three months? (besides the usual seasonal pattern)	11/2008 – 4/2012	5-item Likert
■ Have more or less lay-offs been announced by the employers ?[a]	11/2008 – 1/2011	5-item Likert
■ How do you expect employment in your district to develop within the next three months? (besides the usual seasonal pattern)	2/2011 – 4/2012	5-item Likert
■ Have more or less lay-offs of contingent workers been announced by the employers?[a]	11/2008 – 6/2010	5-item Likert
■ How do you expect employment in contingent work in your district to develop within the next three months (besides the usual seasonal pattern?)	2/2011 – 4/2012	5-item Likert
■ Do you observe an increase in the demand for contingent workers?	9/2009 – 6/2010	yes/no
■ Do you observe more lay-offs or more job creation in contingent work?[a]	7/2010 – 1/2011	5-item Likert
■ Do you observe lay-offs of workers subsequently to support by “reduced hours compensation”? If yes, how many?	3/2009 – 5/2010	yes/no (+ number)
■ Do you need to give more advise regarding “reduced hours compensation”?[a]	11/2008 – 6/2010 11/2011 – 4/2012	5-item Likert
■ Do you need to give more advise regarding “transitional companies”?[a]	11/2008 – 4/2012	5-item Likert
■ Do more or less employees in your district have contracts with “transitional companies”?[a]	11/2008 – 4/2012	5-item Likert
■ Do you observe excess demand for specialist workers? (If yes, in which occupations?)	7/2010 – 10/2011 01/2012	yes/no (+ text field in some waves)
■ Do you observe more intra-firm transitions from vocational training to regular work?[a]	7/2010 – 10/2010 2/2011 – 10/2011 2/2012 – 4/2012	5-item Likert

[a]: (compared to one year before)

The five item Likert scale is centered around zero; in general a value of -2 corresponds to the answer ‘declines strongly’, a value of $+2$ to the answer ‘increases strongly’. An exception is the question ‘Do you observe more lay-offs or more job creation in contingent work’ where we coded ‘much more job creation’ with a value of -2 and ‘much more lay-offs’ with $+2$.

Recently published

No.	Author(s)	Title	Date
11/2012	Deeke, A. Baas, M.	Berufliche Statusmobilität von Arbeitslosen nach beruflicher Weiterbildung: Ein empirischer Beitrag zur Evaluation der Förderung beruflicher Weiterbildung	4/12
12/2012	Nordmeier, D.	The cyclicality of German worker flows: Inspecting the time aggregation bias	5/12
13/2012	Kröll, A. Farhauer, O.	Examining the roots of homelessness: The impact of regional housing market conditions and the social environment on homelessness in North-Rhine-Westphalia, Germany	5/12
14/2012	Mendolicchio, C. Paolini, D. Pietra, T.	Asymmetric information and overeducation	6/12
15/2012	Poeschel, F.	Assortative matching through signals	6/12
16/2012	Dauth, W. Findeisen, S. Suedekum, J.	The rise of the East and the Far East: German labor markets and trade integration	7/12
17/2012	Münich, D. Srholec, M. Moritz, M. Schäffler, J.	Mothers and Daughters: Hereogeneity of German direct investments in the Czech Republic Evidence from the IAB-ReLOC survey	7/12
18/2012	Scholz, Th.	Employers' selection behavior during short-time work	8/12
19/2012	Osiander, Ch.	Selektivität beim Zugang in Weiterbildungsmaßnahmen: Die Bedeutung individueller und struktureller Faktoren am Beispiel der „Initiative zur Flankierung des Strukturwandels"	8/12
20/2012	Dauth, W. Südekum, J.	Profiles of local growth and industrial change: Facts and an explanation	9/12
21/2012	Antoni, M. Heineck, G.	Do literacy and numeracy pay off? On the relationship between basic skills and earnings	9/12
22/2012	Blien, U. Messmann, S. Trappmann, M.	Do reservation wages react to regional unemployment?	9/12
23/2012	Kubis, A. Schneider, L.	Human capital mobility and convergence: A spatial dynamic panel model of the German regions	9/12
24/2012	Schmerer, H.-J.	Skill-biased labor market reforms and international competitiveness	10/12

As per: 23.10.2012

For a full list, consult the IAB website
http://www.iab.de/de/publikationen/discussionpaper.aspx

Imprint

IAB-Discussion Paper 25/2012

Editorial address
Institute for Employment Research
of the Federal Employment Agency
Regensburger Str. 104
D-90478 Nuremberg

Editorial staff
Regina Stoll, Jutta Palm-Nowak

Technical completion
Jutta Sebald

Website
http://www.iab.de

Download of this Discussion Paper
http://doku.iab.de/discussionpapers/2012/dp2512.pdf

ISSN 2195-2663

For further inquiries contact the author:

Norbert Schanne
E-mail norbert.schanne@web.de