Discussion Paper No. 09-042

Higher-Order Beliefs Among Professional Stock Market Forecasters: Some First Empirical Tests

Jesper Rangvid, Maik Schmeling,
and Andreas Schrimpf

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 09-042

Higher-Order Beliefs Among Professional Stock Market Forecasters: Some First Empirical Tests

Jesper Rangvid, Maik Schmeling,
and Andreas Schrimpf

Non-Technical Summary

A sizeable literature reports that financial market analysts and forecasters herd for reputational reasons (See e.g. Devenow and Welch (1996) and Hirshleifer and Teoh (2003) for surveys). As an example, Lamont (2002) argues that if forecasters are "punished" for being wrong, forecasters could have an incentive to mimic other forecasters, i.e. to herd. Using new data from a large survey of professional forecasters' expectations about stock market movements, we find strong evidence that the expected average of all forecasters' forecasts (the expected consensus forecast) influences an individual forecaster's own forecast.

This looks like herding. In our survey, forecasters do not herd for reputational reasons, however. Instead of herding, we suggest that forecasters form higher-order expectations in the spirit of Keynes (1936) who compared the setting of prices in financial markets with a newspaper beauty contest. In this contest, competitors were invited to pick the six prettiest faces from among one hundred photographs. The winner of the contest was the person whose choice most closely corresponded to the average preference of all competitors. Hence, a competitor should not only pick the prettiest faces, but those he thought the other competitors would also view as the prettiest, and – at the same time – take into account that all other competitors would form expectations in a similar manner. Keynes' idea was that financial markets work in this way, too.

Keynes' description of asset pricing as a beauty contest has recently received renewed attention in the theoretical literature examining the *theoretical implications* of higher-order expectations for asset prices (e.g. Allen, Morris, and Shin (2006), Bacchetta and van Wincoop (2006, 2009), Nimark (2007), Banerjee, Kaniel, and Kremer (2009), and Makarov and Rytchkov (2008)). We follow up on those papers by providing some first *empirical evidence* on this form of expectation formation in this paper.

Our first result is that forecasters are influenced by the expected consensus forecast. This basic result is highly statistically significant, remains significant when adding different kinds of control variables, shows up in all our robustness tests, and, most importantly, is also of substantial economic significance. Finding that forecasters are influenced by the consensus forecast in this dataset of German forecasters both confirms findings from the U.S. to other countries, and, at the same time, thereby provide "out-of-sample" evidence on the results from the U.S. studies, as reported in, e.g., Graham (1999), Welch (2000), and Lamont (2002). We find that young forecasters and portfolio managers, who in previous studies have been reported to be those who in particular herd, rely more on the expected consensus forecasts than other forecasters. Given that forecasters have no incentive to herd in our study, we conclude that our results indicate that the incorporation of the expected consensus forecast into individual forecasts is most likely due to higher-order expectations.

Zusammenfassung (Non-Technical Summary in German)

Eine umfangreiche Literatur findet emprische Belege, dass Finanzmarktanalysten und Prognostiker aus Reputationsgründen zu Herdenverhalten neigen (vgl. Devenow and Welch (1996) sowie Hirshleifer and Teoh (2003) für Überblicksartikel). So weist unter anderem Lamont (2002) darauf hin, dass Finanzmarktanalysten – sofern sie durch ihre Fehlprognosen Nachteile auf Grund von Reputationsverlusten erleiden – einen expliziten Anreiz dazu haben, das Verhalten anderer Prognostiker zu imitieren. Dieses Phänomen wird typischerweise als Herdenverhalten bezeichnet. Unter Verwendung eines neuen Mikrodatensatzes aus einer umfassenden Umfrage unter professionellen Finanzmarktprognostikern zu deren künftigen Aktienmarkterwartungen, finden wir starke Evidenz dafür, dass der erwartete Durchschnitt der Prognosen aller Prognostiker (d.h. die erwartete Konsensus-Prognose) einen wichtigen Einfluss auf die individuelle Prognose der einzelnen Prognostikers aufweist.

Dieses Resultat legt als Interpretation Herdenverhalten nahe. Ein Merkmal unseres Datensatzes ist jedoch, dass bei den befragten Prognostikern keinerlei Gründe für Herdenverhalten aus Reputationsgesichtspunkten gegeben sind. An Stelle von Herdenverhalten, führen wir daher das empirisch beobachtete Verhalten der Prognostiker auf die Bildung so genannter Erwartungen höherer Ordnung (higher-order expectations) zurück. Erwartungen höherer Ordnung gehen auf Keynes (1936) zurück, welcher die Preisbildung auf Finanzmärkten mit einem seinerzeit typischen Schönheitswettbewerb (beauty contest) in Zeitungen verglich. In diesem Wettbewerb wurden die Teilnehmer aufgefordert, aus hundert Fotos die sechs hübschesten Gesichter auszuwählen. Der Gewinner des Wettbewerbs war derjenige dessen Wahl am ehesten mit der Durchschnittspräferenz der anderen Teilnehmer übereinstimmte. Demnach war es für den einzelnen Teilnehmer nicht optimal, das aus seiner Sicht hübscheste Gesicht zu wählen, sondern dasjenige, von welchem er davon ausging, dass auch die anderen Teilnehmer es als das hübscheste ansehen würden. Dabei musste er beachten, dass die anderen Teilnehmer auf die gleiche Art und Weise Erwartungen bilden würden. Keynes war der erste, der diese Idee eines Schönheitswettbewerbs auf die Preisbildung auf Finanzmärkten übertrug.

Die Erklärung der Bildung von Vermögenspreisen auf Basis der von Keynes beschriebenen Mechanismen eines Schönheitswettbewerbs hat kürzlich in der *theoretischen* Literatur eine erneuerte Aufmerksamkeit erfahren (z.B. Allen, Morris, and Shin (2006), Bacchetta and van Wincoop (2006, 2009), Nimark (2007), Banerjee, Kaniel, and Kremer (2009), and Makarov and Rytchkov (2008)). Wir folgen diesen Arbeiten, indem wir in dieser Studie erste *empirische Evidenz* zu dieser Form der Erwartungsbildung vorlegen.

Unser erstes Resultat ist, dass die Finanzmarktprognostiker in unserer Stichprobe durch die erwartete Konsenus-Prognose beeinflusst werden. Dieses Resultat ist statistisch hoch signifikant, bleibt signifikant, wenn für eine Reihe weiterer Einflussfaktoren auf die Aktienmarkterwartung kontrolliert wird, zeigt sich in allen Robustheitstests und ist ebenfalls, was von besonderer Bedeutung ist, von substanzieller öknomischer Signifikanz. Das Resultat, dass Prognostiker durch den Konsensus beeinflusst werden, unterstützt bisherige Resultate für die USA (vgl. u.a. Graham (1999), Welch (2000), and Lamont (2002)) und liefert auf diese Weise “out-of-sample” Evidenz. Ein weiteres zentrales Resultat ist, dass sich besonders junge Prognostiker sowie Portfolio Manager verstärkt am erwarteten Konsensus orientieren. Angesichts der Tatsache, dass die Prognostiker in unserem Datensatz keinen Anreiz zu Herdenverhalten aus Reputationsgesichtspunkten haben, führen wir unsere Resultate hinsichtlich der Bedeutung des erwarteten Konsensus für die Individualprognosen auf die Bildung von Erwartungen höherer Ordnung zurück.

Higher-Order Beliefs among Professional Stock Market Forecasters: Some first Empirical Tests*

Jesper Rangvid[**] **Maik Schmeling**[‡] **Andreas Schrimpf**[§]

This version: July 2009

[*]The authors would like to thank Lukas Menkhoff, Paul Söderlind, Peter Norman Sørensen, Robert Whitelaw, Annette Vissing-Jørgensen, and several seminar and workshop participants for helpful comments and suggestions.

[**]Corresponding author. Department of Finance, Copenhagen Business School, Solbjerg Plads 3, DK-2000 Frederiksberg, Denmark. Phone: (45) 3815 3615, fax: (45) 3815 3600, e-mail: jr.fi@cbs.dk.

[‡]Department of Economics, Leibniz Universität Hannover, Königsworther Platz 1, 30167 Hannover, Germany, e-mail: schmeling@gif.uni-hannover.de.

[§]Centre for European Economic Research (ZEW) Mannheim, P.O. Box 10 34 43, 68034 Mannheim, Germany, Tel.: +49 621 1235160, e-mail: schrimpf@zew.de.

Higher-Order Beliefs among Professional Stock Market Forecasters: Some first Empirical Tests

Abstract

A sizeable literature reports that financial market analysts and forecasters herd for reputational reasons. Using new data from a large survey of professional forecasters' expectations about stock market movements, we find strong evidence that the expected average of all forecasters' forecasts (the expected consensus forecast) influences an individual forecaster's own forecast. This looks like herding. In our survey, forecasters do not herd for reputational reasons, however. Instead of herding, we suggest that forecasters form higher-order expectations in the spirit of Keynes (1936). We find that young forecasters and portfolio managers, who in previous studies have been reported to be those who in particular herd, rely more on the expected consensus forecasts than other forecasters. Given that forecasters have no incentive to herd in our study, we conclude that our results indicate that the incorporation of the expected consensus forecast into individual forecasts is most likely due to higher-order expectations.

JEL-Classification: G10, G15, G17

Keywords: Higher-Order Expectations, Stock Market Forecasts, Forecaster Heterogeneity

1 Introduction

Herding in financial markets occurs when individual market participants bias their forecast away from their best, or unbiased, estimate of future outcomes towards the consensus expectation from the previous period. The literature on herding among financial analysts and forecasters "has focused on herding for rational reputational reasons" (Hirshleifer and Teoh, 2003, p. 45). As an example, Lamont (2002) argues that if forecasters are "punished" for being wrong, forecasters could have an incentive to mimic other forecasters, i.e. to herd.[1]

The first thing we do in this paper is to use a new dataset to test whether the average (or consensus) expectation matters for the expectations individual market forecasters form. We find strong and robust evidence that this is indeed the case. On the face of it, this looks very much like herding for reputational reasons.

However, an interesting feature of the data we use, is that we can a priori rule out herding for reputational reasons. The reason for this is that the *individual* forecasts of the forecasters in our survey are not published – it is only the consensus forecast that is published (we describe the data in more detail below). This particular feature of the data eliminates herding arising from reputational concerns, as the forecast error of the individual forecaster cannot per definition have any effect on his career or reputation, as it simply cannot be evaluated by outside observers whether the forecaster made a good or a bad forecast.[2] Hence, even if it looks like herding, we know that it is not herding. We think this provides an interesting new dimension to the empirical literature on herding in financial markets.

If the consensus expectation matters for the formation of individual expectations, but it cannot be because of herding by construction of our data, why does the consensus matter then? The hypothesis we pursue in this paper is that investors form higher-order

[1] Stickel (1992), Graham (1999), Welch (2000), Hong, Kubik, and Solomon (2000), and Ashiya and Doi (2001) all provide evidence indicating that analysts or forecasters herd for reputational reasons. For surveys, see Devenow and Welch (1996) and Hirshleifer and Teoh (2003).

[2] Similarly, since individual forecasts are not published, we can also rule out explanations based on information cascades (e.g. Bikhchandani, Hirshleifer, and Welch, 1992).

expectations.

Keynes (1936) was the first to introduce higher-order expectations as a description of the way expectations are formed in financial markets. It makes sense to briefly review Keynes' idea before we proceed. Keynes (1936) compared the setting of prices in financial markets with a newspaper beauty contest. In the contest, competitors were invited to pick the six prettiest faces from among one hundred photographs. The winner of the contest was the person whose choice most closely corresponded to the average preference of all competitors. Hence, a competitor should not only pick the prettiest faces, but those he thought the other competitors would also view as the prettiest. At the same time, the competitor should take into account that all other competitors would form expectations in a similar manner. Keynes' idea was that financial markets work in this way, too: An investor should not only buy the financial assets he expects to perform well. Instead, the investor should buy those assets he expects the other investors will choose to buy, at the same time taking into account that the other investors will act in a similar way, i.e. that they all try to guess what the other investors try to guess about the other investors. Keynes dubbed this form of expectation formation "higher-order expectations".

In Keynes' beauty contest, the consensus expectation does not contain new information about fundamentals: all competitors can see all faces, i.e. no new information about fundamentals (the pictures) is revealed if investors see the consensus expectation. In financial markets it might be, however, that forecasters look at the consensus because they believe that other forecasters have information about future outcomes that the individual forecaster does not have. In other words, it could also be that forecasters believe that the average forecast summarizes other forecasters' otherwise dispersed private informative signals (i.e. the average forecast aggregates other forecasters' private information) about an asset's fundamental value. In this case, a reliance on consensus expectations would arise because forecasters update their beliefs with new information.

Hence, we need to distinguish between higher-order expectations and standard updating of beliefs with new relevant information. Following up on Keynes' idea, the identifying assumption we use to empirically distinguish between higher-order expectations and standard rational updating of beliefs is that standard rational updating of beliefs occurs when forecasters believe that the other forecasters receive informative signals about final outcomes that are different from the signals that the individual forecasters see (i.e. that the individual forecasters believe that the consensus expectation contains information about, as in our case, the fundamental value of the stock market). On the other hand, we say that investors form higher-order expectations when the forecasters do not necessarily believe that the signals that the other forecasters receive are informative, but they believe that the forecasts of the other forecasters nevertheless influence outcomes.

We provide two tests that indicate that our finding that an individual forecaster forms expectations with an eye towards the consensus forecast is due to the formation of higher-order expectations. In these tests we evaluate whether those forecasters who have been found to be more prone to herding in earlier studies also rely more on the consensus expectation when forming their own expectation in this data set. In particular, we evaluate whether portfolio managers and/or young forecasters rely more on the consensus than other forecasters.[3]

To test hypotheses such as those just outlined above, good (micro-)data are required. The data we use come from a survey conducted among professional German forecasters. The main advantage of these data is that they do not reveal the individual forecasts of the forecasters – only the consensus expectation is made public. An additional advantage of the data is that all forecasters are closely tracked, as are their background characteristics, such as the age of the forecaster, his level of education, his current job function, etc. In addition, the survey keeps the forecast of a forecaster even after the forecaster has left the survey, i.e. there is no survivorship bias in our data.[4] The data spans the period from December 1991 to October 2008. All in all, over the whole sample period, we have 453

[3]To be precise: If it is rational to incorporate the consensus expectation because it affects asset prices, all investors should do so. And, indeed, we also find that investors on average incorporate the consensus expectation. Our identifying assumption, though, is that investors should look at the consensus *per se*, if it is higher-order expectations. Consequently, we see if there are groups of investors who we a priori expect to look at the consensus *per se*. If these investors incorporate the consensus expectation even more, we interpret this as a sign of higher-order expectations.

[4]These features of our data make them advantageous compared to, e.g., the Survey of Professional Forecasters or the Livingston Survey (both of which are U.S. surveys). For instance, there are cases in these U.S. data where the identification code for an investment bank, say, stays constant over time regardless of who the actual forecaster is.

forecasters.

In our survey, respondents are asked to indicate whether they expect the prices on the U.S. and the German equity market to increase ("1"), decrease ("−1"), or remain unchanged ("0 ") over the next six months. The respondents are also asked about their expectations for the U.K., French, Italian, and Japanese stock markets. We focus on the expectations for the U.S. and the home (the German) markets in the main part of the paper and present results for the other countries as robustness checks. As we have a panel of ordered choices made by each forecaster (positive, negative, or unchanged stock market), we estimate random parameters ordered panel logit models. In our basic implementation, we use the lagged average forecast as our estimate of the expected average forecast. It is important to note here that when the forecasters make their forecast for the next period, last period's average forecast is known to all forecasters.

We now explain our empirical findings. Our first result is that forecasters are influenced by the expected consensus forecast. This basic result is highly statistically significant, remains significant when adding different kinds of control variables, shows up in all our robustness tests, and, most importantly, is also of substantial economic significance. In terms of economic importance, for instance, we find that when the average forecast increases by two standard deviations, i.e. when the average forecaster becomes more optimistic with respect to the performance of the U.S. stock market, an individual forecaster becomes 13.84 percentage point more positive towards the U.S. stock market. Given that the unconditional probability of an "up"-forecast is 38.90 percent, the marginal effect of 13.84 percentage points is indeed economically significant. Finding that forecasters are influenced by the consensus forecast in this dataset of German forecasters both confirms findings from the U.S. to other countries, and, at the same time, thereby provide "out-of-sample" evidence on the results from the U.S. studies, as reported in, e.g., Graham (1999), Welch (2000), and Lamont (2002).

Our next test is based on a sorting of forecasters into "relative forecasters", i.e. forecasters whose pay is related to their outperformance relative to a market benchmark, such as portfolio managers, and forecasters whose pay is not ("absolute forecasters") but rather depends on absolute success. When portfolio managers' pay is related to their outperformance relative to a market benchmark, portfolio managers have a clear incentive to look more towards the forecasts of other forecasters when forming their own forecast, regardless of whether the average forecast contains informative signals about the fundamental value of asset prices or not. We find strong evidence that the impact of consensus beliefs on individual beliefs is higher (about 10% higher) if a forecaster is a "relative" forecaster, compared to the forecasters in our control group. In contrast being in the group of "absolute forecasters" decreases the impact of consensus beliefs by about 20%. Given that those forecasters whose pay is related to their relative outperformance do not look at the expected consensus forecast because they want to be close to the average forecast for reputational or career concerns, as argued above, we interpret this as evidence that they look at the consensus expectation because they think that the consensus matters for asset prices *per se*, i.e. regardless of whether or not it contains fundamental information. Forecaster who are evaluated in terms of absolute success by contrast should rely more on fundamental information and our results do indeed suggest that this group of forecasters relies less on consensus expectations.

The final test we conduct is based on the empirical finding in the literature that younger finance professionals tend to follow market trends and general market consensus (i.e. herd) when they trade (e.g. List, 2003; Feng and Seasholes, 2005; Haruvy, Lahav, and Noussair, 2007; Greenwood and Nagel, 2009). Again, there is no reason for herding in our survey. Hence, if the young traders look towards the expected consensus, it is not because they fear that it will have consequences for their career if they are wrong. Rather, they think that the consensus expectation matters for asset prices, even if it does not necessarily contain new fundamental information about asset prices. We find clear evidence that young and less experienced forecasters incorporate the expected consensus forecast into their own forecast to a significantly larger extent than older forecasters.

Related literature. Keynes' description of asset pricing as a beauty contest has recently received renewed attention in the theoretical literature: Allen, Morris, and Shin (2006), Bacchetta and van Wincoop (2006, 2009), Nimark (2007), Banerjee, Kaniel, and

Kremer (2009), and Makarov and Rytchkov (2008) all examine the theoretical implications of higher-order expectations for asset prices. We follow up on those papers by providing some first empirical evidence in this paper.

Even if we have no knowledge of studies that directly *test* for higher-order expectations within a sample of individual market participants or finance professionals, experiments or simulations have been used to evaluate whether individuals form higher-order beliefs. Bosch-Domènech, Montalvo, Nagel, and Satorra (2002) for example find that there is evidence for the existence of higher-order expectations among students and readers of certain news magazines. Also, Biais and Bossaerts (1998) analytically investigate the effect of differences in beliefs related to beauty contests on asset prices and trading volume and provide simulation results.

Moreover, the papers that analyze the theoretical implications of higher-order expectations often contain interesting motivating verbal discussions of instances where higher-order beliefs seem to have played a role. For instance, Corsetti, Dasgupta, Morris, and Shin (2004) discuss how the existence of one big trader, who is not necessarily better informed than the other traders, can increase the likelihood of a currency attack. Similarly, Morris and Shin (2005) discuss how a central bank can help coordinating the beliefs of the agents in the market by publishing their own forecasts. But again, these discussions contain no direct tests for higher-order beliefs – something we attempt to provide in this paper.

There is a large literature that tests whether financial market participants or forecasters herd. Prominent empirical papers in this area include Stickel (1992), Graham (1999), Welch (2000), Lamont (2002), and Ashiya and Doi (2001) who find evidence of herding, and Zitzewitz (2001) and Bernhardt, Campello, and Kutsoatic (2006), who find that analysts do not herd. Generally, these papers deal with herding for reputational reasons. In contrast, we employ a dataset that is characterized by the absence of reputational concerns. This allows us to discard herding as a reason for why forecasters look at the expected consensus forecast when forming their own forecast.

Structure of paper. The remainder of this paper is structured as follows. In the next section, we briefly sketch the main theoretical implication of higher-order expectations and the main empirical hypothesis that we test. In section 3, we describe the data we use to test our hypotheses and in section 4 we lay out our empirical procedure. Sections 5 and 6 contain the main results. Section 7 contains robustness tests. Section 8 concludes.

2 Theoretical Motivation

Higher-order expectations imply that assets are priced under *average* market expectations (Bacchetta and van Wincoop, 2009):

$$P_t = \rho \overline{\mathbb{E}}_t[X_{t+1}] \tag{1}$$

where ρ is a discount factor, P_t denotes the time t price of an asset and $\overline{\mathbb{E}}[\cdot] \equiv \int_i \mathbb{E}^i[\cdot]di$ denotes the average market expectation regarding the future total payoff X_{t+1} of an asset. The main difference compared with the standard approach is the use of the average market expectation operator which need not equal each individual's expectation $\mathbb{E}^i[\cdot]$ when investors are heterogeneous. The point about heterogeneity here is that individuals have to start thinking about the aggregate expectation of the market and not only about their own expectation, since the asset price is determined by these aggregate beliefs. In such situations, higher-order beliefs can have a significant impact on asset prices in equilibrium. Most importantly, higher-order expectations drive a wedge between the price of the asset in a world of identical agents and the price of the asset in a world with heterogeneous agents and expectations (Bacchetta and van Wincoop, 2009). The reason for this is that when information is heterogeneous, the law of iterated expectations does no longer hold for average expectations.[5] Higher-order expectations also lead to an excessive reaction of asset prices to public information (Allen, Morris, and Shin, 2006).[6]

[5]"It is *not* the case that the average expectation today of the average expectation tomorrow of future payoffs is equal to the average expectation of future payoffs.", Allen, Morris, and Shin (2006, p. 720).

[6]More generally, higher-order beliefs play a role under short-sale constraints where the asset price exceeds its fundamental value due to an implicit option to sell the asset to an investor with a higher private valuation in the future (see e.g. Allen, Morris, and Postlewaite, 1993; Scheinkman and Xiong, 2003). Differences in beliefs, such as higher-order expectations affect the value of that option. However,

It is clear from the above equation that higher-order beliefs imply that:

$$\mathbb{E}_t^i[P_{t+1}] = \mathbb{E}_t^i\left[\rho\overline{\mathbb{E}}_{t+1}(X_{t+2})\right] \tag{2}$$

so that an individual forecaster i has to rely on consensus expectations to make up his individual forecast. Neglecting the information contained in consensus expectations would mean to ignore an important determinant of asset prices. The point of departure of our paper is therefore a test of whether consensus expectations, i.e. $\overline{\mathbb{E}}[X]$, are a significant determinant of individual expectations $\mathbb{E}^i[P]$. Evidence of such a relationship would imply that higher-order beliefs may be at work. We provide strong statistical evidence in favor of such a relationship in the remainder of this paper.[7]

After having verified that $\overline{\mathbb{E}}[X]$ is an important determinant of $\mathbb{E}^i[P]$, we make use of the richness of our data to conduct even more stringent tests of whether it really is higher-order expectations that are at play. In particular, we hypothesize that there are certain groups of investors (to be detailed below) who have a larger incentive *per se* to rely on $\overline{\mathbb{E}}[X]$ when forming their own forecasts. Given that we find that these groups of investors in fact do put more weight on the average expectation, we interpret this as suggesting that individual forecasters form higher-order expectations.

3 Data and Descriptive Statistics

This section describes our data and provides descriptive statistics. First we describe the general data set and provide details of our cross-section of forecasts. We then move on to describe consensus stock market expectations for our sample countries.

papers in this line of literature do not explicitly focus on the effect of pure higher-order expectations. In addition, there are of course other issues with investor heterogeneity and learning in financial markets which we do not deal with in this paper. A recent survey of work in this field is provided by Pastor and Veronesi (2009).

[7]Eq. (2) can be written in a number of alternative ways. In our main empirical analysis, we thus do not test Eq. (2) directly but rather test for a more general relation between individual expectations and expected consensus expectations.

3.1 General Features of the Data

Our data come from a monthly survey of the Centre for European Economic Research (ZEW) which is one of the largest economic research institutes in Germany. At the beginning of each month, approximately 350 professional forecasters from large German banks, institutional investors, or treasury departments of large corporations are asked whether they expect a specific stock market to go "up", "down", or remain "unchanged" over the next six months. Hence, we are dealing with qualitative data. The respondents are asked about their directional forecast of aggregate stock market indices in the U.S., Germany, U.K., France, Italy, and Japan. The respondents generally reply before the second Tuesday of a given month, i.e. the individual forecasts are not spread out over the whole month. In addition, when the forecasts are made for the following month, last month's average expectation for this month is well-known to all forecasters since consensus expectations are made public on the second Tuesday of a month.[8] On the other hand, the forecasts of the individual forecasters are not made public and, hence, cannot be evaluated by any outside observers. This last feature of the data is crucial for our tests of higher-order expectations, as it eliminates concerns about herding for reputational reasons.

Forecasts are collected in a micro-panel of forecasters and data in this study are available from December 1991 (the start of the survey) to October 2008. Therefore, our sample spans the major bull and bear markets in the late 1990s and early 2000s as well as the current 2007/2008 market meltdown.

Our dataset includes individual stock market forecasts for each forecaster and their corresponding socio-economic background characteristics. The panel of forecasters has been quite stable over time with relatively few exits and new entries. Our sample fortunately includes data for all forecasters even if some forecasters have left the panel at some point in time during the sample period. There are thus no biases arising from survivorship issues. Furthermore, the panel generically identifies individuals and not institutions. A forecaster who changes his job and moves from one bank to another, for example, keeps

[8] The consensus expectation (i.e. the average forecast) is actually followed closely in Germany and gets quite some attention in the financial news media.

his identifying code so that our panel is not distorted by these issues. This is a major advantage over other panels (such as the Survey of Professional Forecasters) where it is often less clear whether a specific "forecaster" actually remains the same over time.

3.2 Individual Forecasters

What kind of forecasters are we dealing with in this study? To answer this question, Table 1 provides descriptive statistics on the socio-economic background characteristics of the 453 individual forecasters in our (unbalanced) panel data set.[9]

Gender is a dummy variable equal to one for male forecasters and zero otherwise. Our sample mainly consists of men which account for more than 90% of all individuals. Age is measured in years and the table shows that our average forecaster is in his mid-forties and that we also have very young (minimum age of 27 years) and old (up to 70 years) forecasters.[10] Moreover, we have information about the professional experience of forecasters which is broken down into general work experience (on any full-time job) and experience relating to a specialization in financial markets. However, the latter two do not differ much and our average forecaster has a significant experience of about 20 years in the financial sector. Notice that a treasurer of an industrial firm, for instance, is characterized as having experience in financial markets, as a treasurer also manages financial risks (interest rate, exchange rate risk etc.). This last point becomes important later in the paper when we split forecasters into different groups based on their current job functions.

TABLE 1 ABOUT HERE

Turning to the educational background we see that our sample mostly consists of individuals having a training in economics or business (labeled "economics education" here).[11]

[9]On average, approximately 300 forecasters participate in the survey each month. Our sample, however, is smaller, since background characteristics are not available for each single forecaster.

[10]We have taken the age variable as of 2008 to make forecasters comparable.

[11]This includes any degree in Economics and Business Administration, such as Bachelor and Master degrees in Business, Finance, Economics, etc.

Furthermore, about 77% have a university degree and almost 10% have completed a doctoral degree. We are thus dealing with a well-trained and highly educated sample of forecasters.

3.3 Consensus Stock Market Expectations

Consensus stock market expectations in month t are calculated by the balance statistic (share of "up" minus "down" forecasts in percent) based on all individual forecasts in month t. Therefore, our consensus expectations are bound between minus and plus one by constructions. Figure 1 shows time-series plots of consensus expectations. It can be seen that consensus forecasts are quite variable, especially during some sharp short-term declines in consensus forecasts e.g. in March – May 1998, February – March 2000, or December 2007 – February 2008. It can also be seen that there are essentially two regimes during our sample, one from 1991 – 2001 where consensus expectations tend to rise and a second regime from 2002 – 2008 where consensus expectations are falling. We will look at these two regimes more closely in the robustness section.

FIGURE 1 ABOUT HERE

Table 2 shows descriptive statistics for consensus stock market forecasts for the six countries in our sample. As in Strong and Xu (2003), there is a clear tendency for forecasters to be more optimistic for their home (the German) equity market: The average, median, minimum and maximum consensus forecasts are largest for the German market.[12] Also note that the volatility of consensus expectations is lowest for the home market as in Strong and Xu (2003). Therefore, forecasters are most optimistic for their home market and they do not adjust their forecasts as much as for other countries.

Table 2 also shows first order autocorrelation coefficients (ρ_{-1}) and results from unit-root tests. Autocorrelations are very high so that consensus expectations are persistent, but

[12] Strong and Xu (2003) use the Merrill Lynch survey of professional investors.

unit-root tests suggest that the time-series of consensus forecasts are stationary. The ADF test suggests stationarity at least at the 10% level for all six countries whereas more powerful panel unit root tests (shown in the last two rows in the table) clearly reject the null of a unit-root.[13]

Table 2 about here

In an Appendix that is available on our webpages, we show that there is a positive correlation between all six consensus stock market forecasts. Some of these correlations (e.g. France/US, Japan/US) are quite low, however, so that there is quite some cross-country variation in stock market expectations (which is also evident from Figure 1). In the Webappendix, we also provide transition probabilities. We find that expectations are quite persistent, as also indicated by the first-order autocorrelations and the unit root tests in Table 1. For example, there is 62% probability that a "down" forecast for the U.S. in month t is followed by "down" forecast in month $t+1$ but only a 12% chance that the forecast in $t+1$ will be "up".

In the following, we will focus on the results we get for the U.S. and the German stock market forecasts. Choosing the U.S. and Germany actually covers most information in our sample as there are tight relationships between the U.S. and U.K. forecasts and the forecasts for Germany, France, and Italy, respectively. To provide the full picture also, though, we discuss results for the other countries in the robustness section.

[13]The high persistence makes sense from an economic perspective, since forecasts are six months ahead and thus overlap at our monthly frequency. Stationarity also seems reasonable, since balance statistics are bound between -1 and +1 by construction.

4 Empirical Implementation

This section details our estimation procedure. We are principally interested in regressions in the general form of:

$$e^*_{i,t} = \alpha_i + \mathbb{E}[\bar{\mathbf{e}}_t]\theta_i + x'_{i,t}\beta + +e^*_{t-1,i}\gamma + \varepsilon_{i,t} \tag{3}$$

where e^* is the stock market return forecast of forecaster i ($i = 1, ..., N$) at time t ($t = 0, ..., T_i$), $\mathbb{E}[\bar{\mathbf{e}}_t]$ is the *expected* consensus expectation for period t, and x is a vector of control variables which is possibly individual-specific. We focus our attention on the estimate of θ_i which quantifies the degree to which individual forecasts are influenced by movements in expected consensus expectations. In our empirical applications below, we will allow for heterogeneity across forecasters by making both the intercept α_i and the slope coefficient θ_i individual-specific, i.e. we allow for cross-sectional randomness in these parameters.

Unfortunately, direct estimation of this regression by OLS is not possible due to the qualitative nature of our data (we do not observe the quantitative return forecast $e^*_{i,t}$ but only the qualitative analogue), forecaster heterogeneity (via α_i and θ_i), and the need to include lagged dependent variables ($e^*_{i,t-1}$) as controls in several of our regression specifications. Our empirical strategy takes these important features of our data into account. First, our data are discrete and ordered so that we resort to ordered choice models. Second, we have to deal with biases arising from lagged variables entering a panel regression specification which allows for cross-sectional heterogeneity. We detail our econometric approach of tackling these challenges below and in the Appendix.

4.1 Ordered Choice Approach

Since our data consist of individual *qualitative* stock market forecasts, we make use of ordered choice regression models. Ordered choice models are well-known in finance (e.g. Hausman, Lo, and MacKinlay, 1992) and our implementation of the methodology is straightforward. For this reason, the basic points will only be discussed briefly.

Ordered choice models are based on the idea that an observed variable (e – qualitative stock forecasts) is related to the underlying but unobserved variable (e^* – quantitative stock market expectations) in the following way:

$$
\begin{aligned}
e &= -1 \;\text{ if }\; e^* \leq 0, \\
&= 0 \;\text{ if }\; 0 < e^* \leq \mu_1, \\
&= 1 \;\text{ if }\; \mu_1 < e^*
\end{aligned}
$$

where $e = -1, 0, 1$ correspond to forecasts of "down", "unchanged", and "up", respectively, and μ_1 is a threshold parameter to be estimated along with all other regression parameters of the model. If one assumes that ϵ in a regression of the latent variable e^* on a set of regressors $e^* = x'\beta + \epsilon$ has a logistic distribution, then the probability of observing outcome j ($j = -1, 0, 1$) is given by $\Pr[e = -1] = L(-x'\beta)$, $\Pr[e = 0] = \mathrm{L}(\mu_1 - x'\beta) - \mathrm{L}(-x'\beta)$, and $\Pr[e = 1] = 1 - \mathrm{L}(\mu_1 - x'\beta)$ where L is the cumulative distribution function of the logistic distribution. Since all probabilities have to be positive, we also have to impose the restriction $\mu_1 > 0$. The formulation here only estimates one cut-off parameter (μ_1) such that the intercept in the ordered logit regression is identified.

4.2 Cross-Sectional Heterogeneity

As indicated in Eq. (3) above, we allow the intercept α_i and slope coefficient θ_i to be random across forecasters. This specification allows us to take heterogeneity across forecasters into account since it cannot be expected that all forecasters share the same average level of optimism/pessimism (via α) or the same reliance on consensus expectations (via θ). Our random parameters model directly quantifies the degree to which individual forecasters differ in these dimensions.

In its simplest form, we employ the following specification for cross-sectional randomness:

$$\begin{pmatrix} \alpha_i \\ \theta_i \end{pmatrix} = \begin{pmatrix} \overline{\alpha} \\ \overline{\theta} \end{pmatrix} + \begin{pmatrix} \sigma_{\overline{\alpha}} & 0 \\ 0 & \sigma_{\overline{\theta}} \end{pmatrix} \begin{pmatrix} v_i^{\alpha} \\ v_i^{\theta} \end{pmatrix} \tag{4}$$

where $v_i^j \sim \mathcal{N}(0,1)$ are normally distributed and uncorrelated with each other. This formulation allows for an estimation of the overall (or average) level of optimism (via $\overline{\alpha}$) and reliance on consensus expectations (via $\overline{\theta}$) while also quantifying the dispersion across forecasters (via $\sigma_{\overline{\alpha}}$ and $\sigma_{\overline{\theta}}$). It is important to note, however, that randomness is purely cross-sectional (hence the i subscript) and not over time.

In the empirical applications below, we will also estimate regressions with more random parameters, e.g. interaction terms of consensus expectations with other conditioning variables. In these cases, the general structure of cross-sectional heterogeneity is directly analogous to the two-parameter case shown above and reads:

$$\theta_i = \overline{\theta} + \Lambda \mathbf{v}_i \tag{5}$$

where θ_i and $\overline{\theta}$ now denote column vectors (including the intercept α_i) and Λ is a diagonal matrix collecting the cross-sectional dispersion parameters (e.g. $\sigma_{\overline{\alpha}}$ and $\sigma_{\overline{\theta}}$ in the above case). Throughout the paper, we only consider the case of uncorrelated shocks in this paper, i.e. Λ is diagonal in all applications to follow. The covariance matrix of the random parameters, thus, is $\Lambda'\Lambda = \Sigma_\theta$ where Σ_θ is always diagonal.

4.3 Lagged Dependent Variables and Estimation

Lagged dependent variables. In principle, the above models can be estimated via a simulated maximum likelihood approach (see below). However, the presence of lagged dependent variables on the right hand side of the regression ($e_{i,t-1}$ in our case) may be problematic since the lagged dependent variable interferes with the random parameters. This can lead to biased and inconsistent estimates.

This problem is well known for typical panels with a small number of time periods but large cross-sections (see e.g. Greene, 2003) and can be handled by instrumental variables procedures in linear models. In our case, we are dealing with a long panel (many time periods) relative to the number of cross-sectional units so that our approach is unlikely to suffer from this "small-T" problem, just as the estimation of a simple, univariate autoregressive model is unlikely to yield biased estimate in a long time-series. In order to be conservative, however, we account for possible biases arising from the presence of lagged dependent variables by relying on a procedure proposed by Heckmann (1981).

This procedure basically treats the initial period ($t = 0$) as an equilibrium without effects from lagged dependent variables. These effects only enter the regression for subsequent periods $t \geq 1$ thereby circumventing the danger of inconsistent estimates. Details on the implementation of this procedure can be found in Appendix A.I of this paper. In the empirical section and tables of this paper, we only show estimation results for periods $t > 0$ and postpone the less interesting and less relevant results for $t = 0$ to the Web Appendix of this paper.

Estimation. Due to the multiple sources of heterogeneity induced by allowing for more than one random parameter, the likelihood function generally involves multi-dimensional integrals which render simple estimation by maximum likelihood infeasible. We therefore estimate our models via Simulated Maximum Likelihood (SML) based on 250 Halton draws. Halton draws speed up the estimation by spreading random draws more evenly across the support of the target distribution. More details on the simulation estimator can be found in the Appendix A.II.

5 Do Individual Forecasters Rely on the Expected Average Forecast when Forming Their own Expectations?

This section shows that individual forecasts are significantly affected by the expected consensus forecast: When forecasters expect consensus forecasts to be high, they also raise their individual expectation for the stock market, and vice versa. This result consistently holds even when we control for a number of other determinants of expected stock returns.

Our main result is based on the following regression specification (in the latent variable formulation):

$$e^*_{i,t} = \alpha_i + \overline{\mathbf{e}}_{t-1}\theta_i + x'_{i,t}\beta + e^*_{t-1,i}\gamma + \varepsilon_{i,t} \tag{6}$$

where θ_i is given by Eq. (4) above.

The difference between Eq. (3) and Eq. (6) is that we have replaced $\mathbb{E}[\overline{\mathbf{e}}_t]$ in Eq. (3) with $\overline{\mathbf{e}}_{t-1}$. We use lagged consensus expectations $\overline{\mathbf{e}}_{t-1}$ as our proxy for expected consensus expectations since our summary statistics in Table 2 revealed that consensus forecasts are significantly autocorrelated. This serial correlation makes it easy for forecasters to infer current expected consensus forecasts from past consensus forecasts.[14] We run different variants of this regressions: With and without control variables in x and with and without lagged individual forecasts $e_{i,t-1}$. Note (again) that both the intercept (α) as well as the regression coefficient on the consensus forecast (θ) are allowed to have a random element in the cross-section of forecasters.

Table 3 shows estimation results for these random parameters ordered logit models. The left panel of this table shows results for U.S. forecasts whereas the right panel shows results for German forecasts. We show three regression specifications for both countries. The first specification regresses individual expectations on a constant and on consensus expectations ($\overline{\mathbf{e}}$), the second controls for lagged individual expectation (e_{-1}), and the third specification includes even further control variables which, as we describe below, proxy for the major determinants of expected stock returns and which are part of the public information set of forecasters.

The main result from these regressions is that lagged consensus forecasts – as a proxy for expected consensus expectations – is significantly and positively linked to individual stock market forecasts in all specifications and across forecasts for the U.S. and German stock markets. The effects are also of economic significance, as demonstrated by the marginal effects discussed below.[15] Therefore, individual forecasts are influenced by consensus expectations even when controlling for lagged individual forecasts and further controls.

[14]We discuss an alternative approach – which leads to very similar results – in the robustness section.
[15]We provide similar findings for the other countries in Table A.4 in the robustness section.

Table 3 about here

In the second specification we include lagged individual expectations. It is only natural that (e_{-1}) is significant since forecasters issue six month forecasts, i.e. there is a natural degree of overlap in the forecast series. The interesting point to notice from Table 3, though, is that consensus forecasts ($\overline{\mathbf{e}}$) remain significant after including lagged individual expectation (e_{-1}), i.e. even if forecasts overlap, forecasting relies on lagged consensus expectations in addition to the own lagged forecasts.[16]

Furthermore, we find that cross-sectional heterogeneity in intercepts and the slope coefficients of consensus forecasts ($\overline{\mathbf{e}}$) matters quite substantially. For example, in specification (iii) of Table 3 we find a standard deviation of the heterogeneity in the intercept of $\sigma_{\overline{\alpha}} = 0.824$ compared to the estimate of the mean ($\overline{\alpha}$) of the intercept of "only" 0.512. Therefore, the general level of optimism (or pessimism) varies considerably across forecasters. Similarly, we find a standard deviation for heterogeneity in the coefficient on $\overline{\mathbf{e}}$ of $\sigma_{\overline{\mathbf{e}}} = 0.030$, which is large compared to the mean slope coefficient estimate of 0.014. These estimates imply that the effect of consensus expectations on individual expectations varies considerably across forecasters and that there are several forecasters who do not form higher-order beliefs. In Section 6, we use this last finding to specify more stringent tests of whether it really is higher-order beliefs that are at play.

Control Variables and Expected Returns. We control for usual determinants of expected stock returns mainly to account for time-varying risk premia in stock returns. The control variables we apply in the full model specifications (iii) and (vi) are: Lagged values of price-earnings ratios (PER), term spread (TS), annual industrial production growth (IP), annual percentage changes in the CPI, six-months changes of the OECD index of leading indicators (LD6), and the short rate (IR3M).[17] We furthermore include

[16] We have also estimated the model without overlapping observations by using only observations from January and July. Results remain qualitatively unchanged.

[17] There is a large literature on the impact of valuation ratios and macro variables on expected stock returns. Cochrane (2005) reviews the link between stock returns and price-earnings ratios and term spreads, Cooper and Priestley (2006) and Rangvid (2006) investigate the impact of production (or output) on expected returns, Campbell and Vuolteenaho (2004) and Cohen, Polk, and Vuolteenaho (2005) show

(lagged) aggregate stock market returns over the previous months (R_{-1}) and the return over the last six months (R_{-6}) in our regressions. We include these two measures of past returns to control for possible trend-chasing or contrarian behavior of forecasters (see e.g. Brown and Cliff, 2004, 2005; Dominitz and Manski, 2005, for the impact of lagged returns on stock market expectations). Controlling for past returns seems especially important to disentangle simple reliance on past returns from reliance on consensus expectations, i.e. higher-order beliefs. Since forecasters are surveyed at the beginning of each month, we use lagged values from the end of the previous month for all control variables. Also, all control variables are country-specific here. We look at global control variables in the robustness section and find very similar results.

Apart from the general effect of consensus expectations on individual forecasts, we document several interesting results for our additional control variables in the full specifications (iii) and (vi). As in Vissing-Jørgensen (2004) we find that price-earnings ratios, as a measure of stock market valuation, are positively linked to expected returns. This seems interesting because price-earnings ratios (or price-dividend ratios) are known to negatively forecast stock returns (Campbell and Shiller, 1988; Fama and French, 1988; Lamont, 1998), so that one would also expect the coefficient on PER to be negative in our regressions. It is not, however, which suggests that forecasters in our sample do not view high valuations as an indication of low subsequent returns. Vissing-Jørgensen (2004) finds the same result in a sample of *individual investors* and argues that this positive impact of valuation on stock market forecasts is evidence of investor overreaction. Our result is corroborative of her finding and it seems rather interesting to detect a similar effect in a sample of *professional forecasters* as well.

Measures of macroeconomic activity employed in these regressions as control variables – industrial production growth, inflation, and leading indicators – are found to have only weak effects that are mostly insignificant. Short-term interest rates (IR3M) are significantly negatively related to expected stock market movements, consistent with recent evidence from predictive regressions in Ang and Bekaert (2007) or Lioui and Rangvid

that inflation impacts stock markets, and Ang and Bekaert (2007) and Lioui and Rangvid (2008) look at the importance of short-term interest rates for future stock returns.

(2008). Past market returns also affect stock market expectations. Short-term returns over the past month (R_{-1}) lead to higher return expectations whereas past returns over intermediate horizons of six months (R_{-6}) lower expected returns. Therefore, forecasters have extrapolative expectations over the short but mean reverting expectations over intermediate horizons.[18]

All in all, there is some evidence that forecasters use information contained in valuation ratios and current macroeconomic conditions. These effects are mostly in line with earlier evidence from the literature but we also find some new evidence regarding the impact of price-earnings ratios on expectations of professional forecasters and the different impact of lagged short and intermediate horizon market returns.

Economic Significance and Marginal Effects. Table 4 investigates marginal effects for the full regression specifications (iii) and (vi) of Table 3 to evaluate the economic significance of higher-order expectations on individual forecasts. Marginal effects for a certain determinant are computed by holding all other determinants at their unconditional mean and by raising the explanatory variable under consideration by two standard deviations (from their mean minus one standard deviation to their mean plus one standard deviation).[19] Results are shown for the three forecast categories "down", "unchanged", and "up" and suggest that higher-order beliefs also matter in economic terms. The result for the U.S. stock market, for example, suggests that a two standard deviation shock to consensus forecasts increases the probability of forecasting "up" by 13.84%. This increase in the probability of an optimistic forecast is quite large relative to the unconditional probability of an "up" forecast of only 38.90%. Less pronounced but still significant is the increase of 6.05% for the German market relative to an unconditional "up" forecast of 62.2%. More generally, the marginal increase in the probability of a certain forecast category is always larger than 10% relative to the unconditional probability for all three categories and both

[18] In their classic papers, Frankel and Froot (1987, 1990) find similar effects for a set of professional exchange rate forecasters. In their data, forecasters also tend to predict that short-run market movements will continue but that exchange rates mean-revert to fundamental levels over intermediate to long horizons. Frankel and Froot have coined the expression "expectation twist" for this empirical phenomenon and our results also point towards the existence of such a twist in equity markets.

[19] For the impact of individual lagged expectations (e_{-1}), we use an increase of one category since using standard deviations for this ordered, discrete variable does not make much sense.

countries.

Table 4 about here

Another strong effect shown in Table 4 is due to lagged individual forecasts. An increase in one category in the lagged forecast (e.g. e_{t-1} increases from "unchanged" to "up") raises the probability of an optimistic US forecast by 23.11% for example. This points towards a strong impact of past expectations on current forecasts and documents the high persistence of individual forecasts as discussed above. Finally, looking at the other control variables we find less dominant effects. The strongest effects are visible for the term spread, the interest rates, and the past six-months market returns. The first two of these results are in line with previous findings in the return-predictability literature since the term spread is a classic forecasting device for future output movements and since the interest rate is highly important for discount factors. Regarding past returns, it is interesting to note that forecasters seem to put a relatively large weight on mean-reverting returns although the empirical evidence on mean reversion in stock returns over horizons of a few months is not very strong.

6 Is It Higher-Order Expectations?

A necessary condition for forecasters forming higher-order expectations is that they incorporate the expected average forecast of other forecasters into their own forecast: If an individual forecaster's forecast is not influenced by what he expects the average forecast to be, then there is no role for higher-order expectations. We have shown that there is strong empirical evidence that individual forecasters look towards the forecasts of others when forming their own forecast. We have also noted that there is no reason to herd in our dataset, as the individual forecasts are not published. Hence, we conclude that individual forecasters do not incorporate the consensus expectation because they want to be "close to the consensus" for reputational reasons.

It could be, though, that forecasters rely on the consensus expectation because they think that the average forecast summarizes otherwise dispersed private information that is relevant for the determination of the fundamental values of stocks. In such a case, the reliance on the average expectation might simply be standard information updating. In essence, if forecasters receive public signals (where the consensus expectation is part of those) and private signals, forecasters should use both when updating their beliefs. As mentioned above, the empirical strategy we pursue to differentiate between higher-order beliefs and standard updating of beliefs is as follows. Standard rational updating of beliefs occurs when the forecasters believe that the other forecasters receive signals that are different from the signal that the individual forecasters see. However, higher-order expectations are at play when the forecasters do not necessarily believe that the signals other forecasters receive are informative, but that the other forecasters nevertheless have an influence on the outcomes. Hence, to evaluate whether it is higher-order expectations that are at play, we evaluate whether forecasters look at the forecasts of others *per se*, i.e. look at the expected consensus expectation regardless of whether they think the consensus forecast summarizes useful new information that is otherwise not publicly available or not.[20]

The first step in this regard is that we split the sample of forecasters into two groups:[21] Those forecasters who are likely to be paid more if they do better than the market and those forecasters who are not. Our prior here is that those forecasters who are paid more if they outperform relative to a market trend look more towards the average expectation, as these forecasters are more interested in where the market is expected to move (and, hence, where the others expect the market to move), regardless of whether the average forecast contains informative signals about the fundamental value of stocks or not. Again, we stress that the forecasters have no reputational reason for doing so as their individual forecasts are not published. Hence, if they incorporate the consensus into their own forecasts, they

[20]A simple check of whether rational updating is a likely explanation of our finding can also be conducted by regressing future stock returns on current consensus expectations, since a rational updating makes sense only if consensus beliefs are actually informative for future stock returns. Running such regressions (not reported for the sake of brevity) in a variety of different specifications does not suggest that consensus expectations forecast future stock returns in a way consistent with rational updating. If anything, our results indicate that consensus expectations are *negatively* related to future stock returns.

[21]We are grateful to Annette Vissing-Jørgensen for suggesting this way of testing for higher-order expectations.

do so because they believe that the consensus affects the outcome, regardless of whether it contains new information or not.

The second step is to investigate whether young forecasters rely more on the consensus expectation. Based on the results of Scharfstein and Stein (1990) and Hong, Kubik, and Solomon (2000) that young market participants tend to herd due to career concerns and the result in Greenwood and Nagel (2009) that young mutual fund managers were "riding the technology bubble" as they exhibited trend-chasing behavior in their investments in technology stocks, we hypothesize that young forecasters might be more prone to forming higher-order expectations. We remember also here that young forecasters have no incentive to herd in our data.

6.1 Relative Forecasters

As just mentioned, we divide the sample of forecasters into two groups: One group containing forecasters who are expected to be paid more if they outperform relative to a benchmark and another group containing forecasters who are more likely paid in terms of absolute success. In the first group we mainly collect portfolio managers and in the second group we mainly collect treasurers and CFO's from industrial firms. The two groups account for roughly 25% ("relative" forecasters) and 30% ("absolute" forecasters), respectively.

Our basic reasoning is as follows: A portfolio manager is often evaluated in terms of outperformance in relation to a certain benchmark. For instance, if an index a portfolio manager is following is down ten percent, but the portfolio manager is down only five percent, the portfolio manager has outperformed the benchmark and has in this sense been successful. Most likely, his pay will also be influenced by his success. Hence, a portfolio manager has a reputational reason for following the benchmark closely, if his outperformance relative to the trend can be observed.

In earlier studies, portfolio managers have been shown to herd, i.e. bias their forecast towards the consensus. If it is also portfolio managers who in this dataset (where they have no incentive to herd) rely particularly much on the consensus expectation, this would indicate that they overweight public information, i.e. form higher-order expectations.

We want to test this assertion. To do so, we augment our baseline specification by two interaction variables: One where we interact the lagged consensus forecast with a dummy variable (d_{rel}) equal to one if a forecaster has a job where it is likely that he is paid in terms of outperformance in relation to a market trend (and zero otherwise) and another interaction variable that interacts the lagged consensus forecast with a dummy equal to one if a forecaster works in a position where it is important whether the firm does well in an absolute sense.

Table 5 shows the results.[22] We find that the coefficients to the $\overline{\mathbf{e}} \times d_{rel}$ interaction term are consistently estimated to be positive and significant whereas the coefficients to the $\overline{\mathbf{e}} \times d_{abs}$ interaction term are consistently estimated to be negative. This indicates that those forecasters who are likely to be paid in terms of outperformance relative to a market trend incorporate consensus expectations to an even larger extent, compared to the reference group which contains forecasters not assigned to any of our two other groups, whereas those forecasters who are not paid in terms of relative outperformance care less about the average expectations.

In Panel B, we show the marginal effects. The marginal effects for the U.S. are such that being in the group of relative forecasters increases the impact of consensus beliefs on individual beliefs by about 10% (2.13% increase relative to the overall effect of 20.52%) whereas being in the group of absolute forecasters decreases the impact of consensus beliefs by about 20% (decrease of 4.13% relative to the overall effect of 20.52%). The effect for the German market forecasts are about 15% (relative group) and -30% (absolute group) relative to the general impact of consensus expectations on individual expectations. These effects clearly seem economically significant, in particular because it is likely that we are underestimating the economic significance of being in the relative or absolute group of forecasters, since we cannot perfectly identify relative and absolute forecasters.

[22]We do not show estimation results for the control variables here to save space and only indicate in the row “Controls” whether control variables are included or not.

Table 5 about here

6.2 Young Forecasters

Several papers (e.g. Scharfstein and Stein, 1990; Hong, Kubik, and Solomon, 2000; Greenwood and Nagel, 2009) argue that younger and less experienced financial market participants are more prone to consensus-orientated behavior.[23] These papers generally find that young market participants *follow* the trades of other traders. The question we ask here is whether young market participants also rely more on what they *expect* the other forecasters to forecast *per se*. The latter case is higher-order expectations. The hypothesis we test here is whether those forecasters (the young) who have been shown to simply follow the herd when trading display the same characteristics when they make forecasts, taking into account that the forecasters in the survey we use here do not have to fear any consequences for their reputation from the forecasts they submit to the survey.

Table 6 shows estimates of regressions where we augment our baseline specification by an interaction variable of the lagged consensus forecast with the age of forecaster i. For symmetry, we also include an interaction variable of lagged individual expectations and age. In this specification we allow cross-sectional randomness in the intercept and the coefficients on consensus forecasts and the interaction term ($\overline{\mathbf{e}}$ $\times$ age). We use age as a broad encompassing proxy of experience although we have direct evidence on job-related experience in financial markets. Experience and age are highly correlated ($> 80\%$) and our results are not affected by using experience instead of age.

Estimates for both countries show a negative coefficient on the interaction term in Panel A of Table 6, indicating that older and more experienced forecasters care significantly less about consensus expectations. Higher-order beliefs thus do play a greater role for young and inexperienced forecasters. Moreover, we find that including age reduces the cross-sectional variation in intercepts and slope coefficients considerably (especially in $\sigma_{\overline{\mathbf{e}}}$),

[23]Corroborating this, Chevalier and Ellison (1999) and Menkhoff, Schmidt, and Brozynski (2006) show that younger fund managers in the U.S. and Germany, respectively, tend to engage more in herding than their older peers.

which suggests that age (or experience) captures a large share of cross-sectional differences between forecasters.

Table 6 about here

Looking at marginal effects in Panel B of that table, we also find that including age leads to a much more pronounced effect of consensus expectations on individual expectations and that being older (and more experienced) significantly reduces the formation of higher-order expectations.

7 Robustness

In this section, we evaluate along different dimensions whether our results are robust. First, we show that those forecasters whose forecast deviated more from the consensus forecast of the previous period tend to incorporate the expected consensus forecast even more when forming their own forecast for the current period. Next, we show that our overall results are not influenced by the general state of the market. Finally, we find that our results hold for the remaining countries in our sample as well as for using other control variables or proxies for the expected consensus forecast.

7.1 Do Individuals Adjust to Consensus Beliefs?

We have shown that forecasters take into account the expected average of individual forecasts. This is a necessary condition for higher-order expectations. Another way to illustrate that investor really adapt to the consensus expectation is to evaluate whether those forecaster's whose forecast was "far away" from the consensus forecast last period adapt more to the consensus than do those forecasters whose forecast was not that far away from the consensus. To do this, we run regressions of individual expectations on the difference between lagged individual expectations and consensus beliefs. We call this variable $DEV_{i,t} = e_{i,t-1} - \overline{\mathbf{e}}_{t-1}$. The results are shown in Table 7.

Table 7 about here

The effect of DEV on own forecasts is consistently estimated to be negative. This makes intuitive sense when interpreted within a higher-order setting: The larger is DEV, the more optimistic was the forecaster last period in comparison with the consensus expectation. The forecaster takes this into account, and adjusts his expectation for this period downwards. This is another strong finding revealing that the average forecast matters for individual forecasts.

7.2 Dependence on Market States

We investigate whether different states of the market such as bull versus bear markets and high versus low volatility have an impact on the degree to which forecasters form higher-order beliefs. To this end, we split our sample into sub-samples of up and down markets and into sub-samples of high and low volatility and re-estimate our baseline models on these samples. We consider both trends and volatility in returns as well as in consensus expectations. Results for these exercises are shown in Table 8.

Bull versus Bear Markets. We first look at sub-samples formed on lagged six-months market returns. Bull markets correspond to months where returns over the last six months are positive whereas bear markets are defined to have negative past returns (over the last six months).[24] We generally do not find different effects regarding the formation of higher-order beliefs during bull and bear markets. Higher-order beliefs thus do not seem to depend on the direction of the general market.

Table 8 about here

[24] Using three or twelve months instead does not change our qualitative findings.

Market Volatility. We form two sub-samples depending on high versus low market return volatility (proxied for by lagged six-months absolute returns of the aggregate stock market). Looking at the results in Table 8 we find that the coefficient estimate of consensus expectations is almost twice as large for the low market volatility sub-sample compared to the high volatility sub-sample for both the U.S. and Germany. That means that forecasters tend to rely more on consensus expectations when the stock market is calm.

Times of Optimism and Pessimism. The charts in Figure 1 indicate two major regimes in our sample: the period from 1991 to 2001 when consensus expectations were rising and a pessimistic regime from 2002 to 2008 when expectations were decreasing. We therefore split the sample into these two sub-samples and re-estimate our base regression for these two regimes. We find that higher-order beliefs matter more in times of upward trending consensus expectations (and matter less in times of downward trends in consensus expectations. In fact, the coefficient on consensus expectations is negative for the pessimistic regime (but insignificant) but positive and highly significant for the optimistic regime. A potential explanation is that the distortions of asset prices resulting from higher-order beliefs (Bacchetta and van Wincoop, 2009) are higher when the consensus estimate is rising since it is relatively easier to buy than to sell short.

Volatility in Consensus Expectations. Finally, we split our full sample into two sub-samples depending on whether the lagged absolute change in consensus forecasts is above or below the median value of absolute consensus changes. The resulting sub-samples are used to investigate whether higher-order beliefs are more important when the volatility in consensus expectations is low or high. Our results suggest that higher-order beliefs are more important in times of large changes in consensus forecasts. Individuals tend to rely more on consensus forecasts when up- or downward revisions of consensus beliefs are large. This result seems intuitive since the information contained in consensus forecasts should be more valuable when consensus expectations change rapidly. It is interesting to note, however, that the *volatility of consensus expectations* has an opposite effect on the formation of higher-order beliefs compared with the effect of *market volatility* documented above. It is thus important to differentiate between the two concepts.

7.3 Results for Other Countries

We show results of our baseline regression specification for the U.K., France, Italy, and Japan – i.e. all other countries with available data – in Table A.4 in the Webappendix. It can be inferred that our main result is stable across these other countries as well. The coefficient estimate of $\overline{\mathbf{e}}$ is significantly positive and highly significant for all four countries. We also find very similar results (compared to the benchmark countries U.S. and Germany) for the impact of lagged individual expectations and the other control variables. Again, the price-earnings ratio is significantly positive (similar to the findings in Vissing-Jørgensen (2004)), the term spread tends to be significantly negative, the short rate (IR3M) is significantly negative, and we find a positive impact of lagged short-term returns and a negative impact of lagged six-months returns (except for Japan).

7.4 Global Control Variables

We have also estimated our regressions with additional global control variables. More specifically, we have included not only country-specific determinants (such as the German price-earnings ratio for the German market for example) but also global determinants, such as global money market interest rates, output movements, inflation etc. This procedure seems to make sense since stock markets are highly integrated nowadays (Bekaert, Harvey, Lundblad, and Siegel, 2008). We find (not reported for the sake of brevity) that the inclusion of global controls does not alter our main findings about higher-order expectations. Regarding the effect of global controls on individual expectations, we find that global short rates are significant in many countries even when including jointly with country-specific rates and that global price-earnings ratios are sometimes negative when included jointly with national price earnings ratios (although not when included alone). Results for other controls are significant for some countries but there is no clear and consistent pattern. We therefore conclude that our results on higher-order expectations are not due to omitted global control variables.

7.5 Expected Consensus Forecasts

Finally, we have also experimented with a different proxy for forecasters' expectations about consensus forecasts. While the results documented in this paper are based on using *lagged* consensus forecasts as a proxy for *expected* consensus forecasts, we have also used time-series forecasts of consensus forecasts as proxies. More specifically, we have computed forecasts of consensus expectations using simple ARMA(1,1)-models. These forecasts are then used as proxies for expected consensus beliefs, i.e. $\mathbb{E}_{t-1}[\overline{\mathbf{e}}_t]$ is used instead of $\overline{\mathbf{e}}_{t-1}$ in our regressions. Our (unreported) results show that using forecasted instead of lagged consensus stock market expectations yields extremely similar results. This result seems reasonable since consensus forecasts are significantly autocorrelated so that the information contained in lagged and predicted consensus expectations is nearly identical.

8 Conclusion

We have shown that individual forecasters tend to be more optimistic for the stock market when they can reasonably assume the general consensus opinion to be optimistic and vice versa. This finding is directly in line with the classic beauty-contest argument by Keynes (1936).

While recent contributions (Allen, Morris, and Shin, 2006; Bacchetta and van Wincoop, 2006, 2009; Nimark, 2007; Banerjee, Kaniel, and Kremer, 2009; Makarov and Rytchkov, 2008) have examined the *theoretical* implications of these higher-order beliefs for asset pricing, we are the first to present *empirical* tests of this kind of expectation formation and learning in financial markets. Using data from a large and reliable survey of professional forecasters, we find strong evidence that the average of expected forecasts matters much for individual forecasters' expectations about future stock market movements. This finding is robust when controlling for a number of factors known from earlier research to be related to expected stock returns and it is robust when controlling for lagged individual stock market expectations. We also find that young forecasters and forecasters who are likely to be paid more if they outperform the market tend to rely more on the consensus expectations, which we interpret as evidence that forecasters form higher-order expectations.

An alternative way to test for higher-order expectations would be to look directly at asset prices. Indeed, if investors form higher-order expectations, a wedge will arise between the fundamental price of an assets and its market price (Bacchetta and van Wincoop, 2009). Also, higher-order expectations lead to an excessive reaction of asset prices to public information (Allen, Morris, and Shin, 2006). Even if both of these implications of higher-order expectations would be interesting to study, they also both require the formulation and estimation of a model giving the "fundamental" price of assets, something that is inherently difficult to specify and implement. For this reason also, we have looked at the formation of expectations among professional forecasters and leave the test of higher-order expectations using market prices to future work.

References

Allen, F., S. Morris, and A. Postlewaite (1993): "Finite Bubbles with Short Sale Constraints and Asymmetric Information," *Journal of Economic Theory*, 61, 206–229.

Allen, F., S. Morris, and H. S. Shin (2006): "Beauty Contests and Iterated Expectations in Asset Markets," *Review of Financial Studies*, 19, 719–752.

Ang, A., and G. Bekaert (2007): "Stock Return Predictability: Is it There?," *Review of Financial Studies*, 20(3), 651–707.

Ashiya, M., and T. Doi (2001): "Herd Behavior of Japanese Economists," *Journal of Economic Behavior and Organization*, 46, 343–346.

Bacchetta, P., and E. van Wincoop (2006): "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle," *American Economic Review*, 96, 552–576.

——— (2009): "Higher Order Expectations in Asset Pricing," *Journal of Money, Credit, and Banking*, forthcoming.

Banerjee, S., R. Kaniel, and I. Kremer (2009): "Price Drift as an Outcome of Differences in Higher Order Beliefs," *Review of Financial Studies*, forthcoming.

Bekaert, G., C. R. Harvey, C. T. Lundblad, and S. Siegel (2008): "What Segments Equity Markets?," Manuscript, Columbia Business School.

Bernhardt, D., M. Campello, and E. Kutsoatic (2006): "Who Herds?," *Journal of Financial Economics*, 80, 657–675.

Biais, B., and P. Bossaerts (1998): "Asset Prices and Trading Volume in a Beauty Contest," *Review of Economic Studies*, 65, 307–340.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades," *Journal of Political Economy*, 100, 992–1026.

Bosch-Domènech, A., J. G. Montalvo, R. Nagel, and A. Satorra (2002): "One, Two, (Three), Infinity, ... : Newspaper and Lab Beauty-Contest Experiments," *American Economic Review*, 92, 1687–1701.

Brown, G. W., and M. T. Cliff (2004): "Investor Sentiment and the Near-Term Stock Market," *Journal of Empirical Finance*, 11, 1–27.

——— (2005): "Investor Sentiment and Asset Valuation," *Journal of Business*, 78, 405–440.

Campbell, J. Y., and R. J. Shiller (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1(3), 195–228.

Campbell, J. Y., and T. Vuolteenaho (2004): "Inflation Illusion and Stock Prices," *The American Economic Review (Papers & Proceedings)*, 94, 19–23.

Chevalier, J., and G. Ellison (1999): "Career Concerns of Mutual Fund Managers," *Quarterly Journal of Economics*, 114, 389–432.

Cochrane, J. H. (2005): *Asset Pricing (Revised Edition)*. Princeton University Press, NJ.

Cohen, R., C. Polk, and T. Vuolteenaho (2005): "Money Illusion in the Stock Market: The Modigliani-Cohn Hypothesis," *Quarterly Journal of Economics*, 120, 639–668.

Cooper, I., and R. Priestley (2006): "Time-Varying Risk Premia and the Output Gap," *Review of Financial Studies*, forthcoming.

Corsetti, G., A. Dasgupta, S. Morris, and H. S. Shin (2004): "Does One Soros Make a Difference? A Theory of Currency Crises with Large and Small Traders," *Review of Economic Studies*, 71, 87–114.

Devenow, A., and I. Welch (1996): "Rational Herding in Financial Economics," *European Economic Review*, 40, 603–615.

Dominitz, J., and C. Manski (2005): "Measuring and Interpreting Expectations of Equity Returns," NBER Working Paper No. 11313.

Fama, E. F., and K. R. French (1988): "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22(1), 3–25.

Feng, L., and M. S. Seasholes (2005): "Do Investor Sophistication and Trading Experience Eliminate Behavioral Biases in Financial Markets?," *Review of Finance*, 9, 305–351.

Frankel, J., and K. Froot (1987): "Using Survey Data to Test Standard Propositions Regarding Exchange Rate Expectations," *American Economic Review*, 77(1), 133–153.

Frankel, J. A., and K. A. Froot (1990): "Chartists, Fundamentalists, And Trading In The Foreign Exchange Market," *American Economic Review, Papers and Proceedings*, 80(2), 181–185.

Graham, J. R. (1999): "Herding among Investment Newsletters: Theory and Evidence," *Journal of Finance*, 54, 237–268.

Greene, W. H. (2003): *Econometric Analysis*. Prentice Hall, Upper Saddle River, fifth edn.

Greenwood, R., and S. Nagel (2009): "Inexperienced Investors and Bubbles," *Journal of Financial Economics*, forthcoming.

Haruvy, E., Y. Lahav, and C. Noussair (2007): "Traders' Expectations in Asset Markets: Experimental Evidence," *American Economic Review*, 97, 1901–1920.

Hausman, J. A., A. W. Lo, and A. C. MacKinlay (1992): "An Ordered Probit Analysis of Transaction Stock prices," *Journal of Financial Economics*, 31, 319–379.

Heckmann, J. (1981): "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. F. Manski, and D. McFadden. MIT Press, Cambridge.

Hirshleifer, D., and S. H. Teoh (2003): "Herd Behavior and Cascading in Capital Markets: A Review and Synthesis," *European Financial Management*, 9, 25–66.

Hong, H., J. D. Kubik, and A. Solomon (2000): "Security Analysts' Career Concerns and Herding of Earnings Forecasts," *RAND Journal of Economics*, 31, 121–144.

Keynes, J. M. (1936): *The General Theory of Employment, Interest and Money*. MacMillan, London.

Lamont, O. (1998): "Earnings and Expected Returns," *Journal of Finance*, 53, 1563–1587.

——— (2002): "Macroeconomic Forecasts and Microeconomic Forecasters," *Journal of Economic Behavior and Organization*, 48, 265–280.

Lioui, A., and J. Rangvid (2008): "Stock Return Predictability in a Monetary Economy," Working paper, Copenhagen Business School.

List, J. A. (2003): "Does Market Experience Eliminate Market Anomalies?," *Quarterly Journal of Economics*, 118, 41–71.

Makarov, I., and O. Rytchkov (2008): "Forecasting the Forecasts of Others: Implications for Asset Pricing," Working Paper, MIT Sloan School of Management.

Menkhoff, L., U. Schmidt, and T. Brozynski (2006): "The Impact of Experience on Risk Taking, Overconfidence and Herding of Fund Managers: Complementary Survey Evidence," *European Economic Review*, 50, 1753–1766.

Morris, S., and H. S. Shin (2005): "Central Bank Transparency and the Signal Value of Prices," *Brookings Papers on Economic Activity*, 2, 1–66.

Nimark, K. P. (2007): “Dynamic Higher-Order Expectations,” Manuscript, UPF Barcelona.

Pastor, L., and P. Veronesi (2009): “Learning in Financial Markets,” Working Paper, University of Chicago.

Rangvid, J. (2006): “Output and Expected Returns,” *Journal of Financial Economics*, 81, 595–624.

Scharfstein, D. S., and J. C. Stein (1990): “Herd Behavior and Investment,” *American Economic Review*, 80, 465–479.

Scheinkman, J., and W. Xiong (2003): “Overconfidence and Speculative Bubbles,” *Journal of Political Economy*, 111, 1183–1220.

Stickel, S. E. (1992): “Reputation and Performance among Security Analysts,” *Journal of Finance*, 48, 1811–1836.

Strong, N., and X. Xu (2003): “Understanding the Equity Home Bias: Evidence from Survey Data,” *Review of Economics and Statistics*, 85, 307–312.

Train, K. (1999): “Halton Sequences for Mixed Logit,” Manuscript, UC Berkeley.

Vissing-Jørgensen, A. (2004): “Perspectives on Behavioral Finance: Does ”Irrationality” Disappear with Wealth? Evidence from Expectations and Actions,” *NBER Macroeconomics Annual 2003*, pp. 139–194.

Welch, I. (2000): “Herding among Security Analysts,” *Journal of Financial Economics*, 58, 369–396.

Zitzewitz, E. (2001): “Measuring Exaggeration and Herding by Equity Analysts,” Working Paper, Stanford GSB.

Appendix

A.I Lagged Dependent Variables

This section of the Appendix illustrates the procedure proposed by Heckmann (1981) which we adapt to our panel setting. We directly consider the case of multiple random parameters and several lagged endogenous variables (we will use lagged dependent variables and interactions thereof with other explanatory variables in our applications).

To this end, consider the following panel regression (which we write in terms of the latent quantitative forecast e^* to ease notational burden)

$$e^*_{i,t} = z'_{i,t}\theta_i + x'_{i,t}\beta_i + y'_{t-1,i}\gamma + \varepsilon_{i,t} \tag{7}$$

for $t = 0, 1, \ldots, T_i$ and $i = 1, \ldots, N$. In this general notation, the vector z collects all explanatory variables (including the intercept) that are allowed to have random parameters (the consensus expectation and interactions of consensus expectations with other variables in our case) and the parameter vector θ is given by

$$\theta_i = \overline{\theta} + \Lambda v_i$$

as in the main text in section 4.2. Furthermore, the vector x collects control variables with fixed slope coefficients and y denotes the vector related to lagged dependent variables (e.g. the lagged dependent variable $e^*_{i,t-1}$ itself and interaction terms). β and γ are parameter vectors and $\varepsilon_{i,t}$ is a normally distributed error term.

Heckman's general procedure treats the first period $t = 0$ as an equilibrium without effects of lagged dependent variables. In $t = 0$ we therefore have

$$e^*_{i,0} = z'_{i,0}\tau_i + x'_{i,0}\delta_i + \varepsilon_{i,0} \tag{8}$$

where the vector of random parameters now reads

$$\tau_i = \overline{\tau} + \Omega v_i$$

so that the mean parameters ($\overline{\tau}$) and standard deviations (Ω) are different from the case $t > 0$. Note, however, that the random vector v_i is the same across *all* time periods t.

For periods $t > 0$ we have the same specification as shown in Eq. (7) above. Now, let ζ_t^0 denote a dummy variable that takes on the value 1 if $t = 0$ and let $\zeta_t = 1 - \zeta_t^0$, then the encompassing model is

$$e_{i,t}^* = \zeta_t^0 z_{i,0}' \tau_i + \zeta_t z_{i,t}' \theta_i + \zeta_t^0 x_{i,0}' \delta_i + \zeta_t x_{i,t}' \beta_i + \zeta_t y_{t-1,i}' \gamma + \varepsilon_{i,t} \tag{9}$$

which can be estimated via Simulated Maximum Likelihood, as we will briefly discuss next.[25]

A.II Details on the Simulated Maximum Likelihood Procedure

Denote the likelihood of observing a qualitative forecast $e_{i,t}$ of forecaster i in month t as

$$P(e_{i,t}|w_{i,t}, \Theta) = g(w_{i,t}, \Theta) \tag{10}$$

where $w_{i,t}$ contains all explanatory variables on the right-hand side of our regressions and Θ collects all parameters. The function $g(w_{i,t}, \Theta)$ is short-cut for the ordered logit model discussed in section 4.1 above. We suppress the differentiation into period $t = 0$ and $t > 0$ (see above) to ease the notational burden.

Remember that we employ random parameters of the form $\mathbb{E}[\theta_i] = \overline{\theta}$ such that $\theta_i = \theta + \Lambda \mathbf{v}_i$ and let $\mathbb{V}[\theta_i] = \Sigma_\theta$ be a diagonal covariance matrix, i.e. we have uncorrelated random parameters. θ contains both the constant as well as all other random slope coefficients.

The true log-likelihood is $\log L = \sum_i \log L_i$ where L_i is the likelihood contribution of forecaster i to the total likelihood. Conditional on the random vector v_i it is straightforward to find this likelihood contribution:

$$L_i | v_i = \prod_{t=1}^{T_i} g(w_{i,t}, \Theta) \tag{11}$$

[25]Note that the initial period $t = 0$ is different across forecasters since we have an unbalanced panel where some forecasters enter the sample later than December 1991. Therefore, δ is identified even if determinants in $x_{i,0}$ are not forecaster-specific, i.e. $x_{i,t} = x_t$.

for $t = 1, ..., T_i$ because we have an unbalanced panel. Since v_i is not observable, one has to integrate it out of the likelihood function by taking the expectation over the distribution of v_i which yields a multivariate integral (the dimension of which depends on the number of random elements in θ_i and, in turn, v_i)

$$L_i = \mathbb{E}_{v_i}[L_i|v_i] = \int_{\mathcal{R}_{v_i}} g(v_i) \left(\prod_{t=1}^{T_i} g(w_{i,t}, \Theta) \right) dv_i \tag{12}$$

which generally cannot be solved for analytically. In the simulated maximum likelihood estimation one thus replaces the unknown integral by an approximate integral obtained via Monte Carlo Simulation. This approximation can be done by drawing a large number of random vectors $v_{i,s}$ which can in turn be used to replace the analytically intractable expression in Eq. (12) above

$$\mathbb{E}_{v_i}[L_i|v_i] \simeq \frac{1}{S} \sum_{s=1}^{S} L_i|v_{i,s} \tag{13}$$

i.e. the average of the likelihood function conditional on v_i where the unobserved vector v_i is replaced by the simulated (and thus observable) random vector $v_{i,s}$. The total log likelihood is then simply the sum over all forecasters i, i.e. $\log L = \sum_i \log L_i$ as mentioned above.

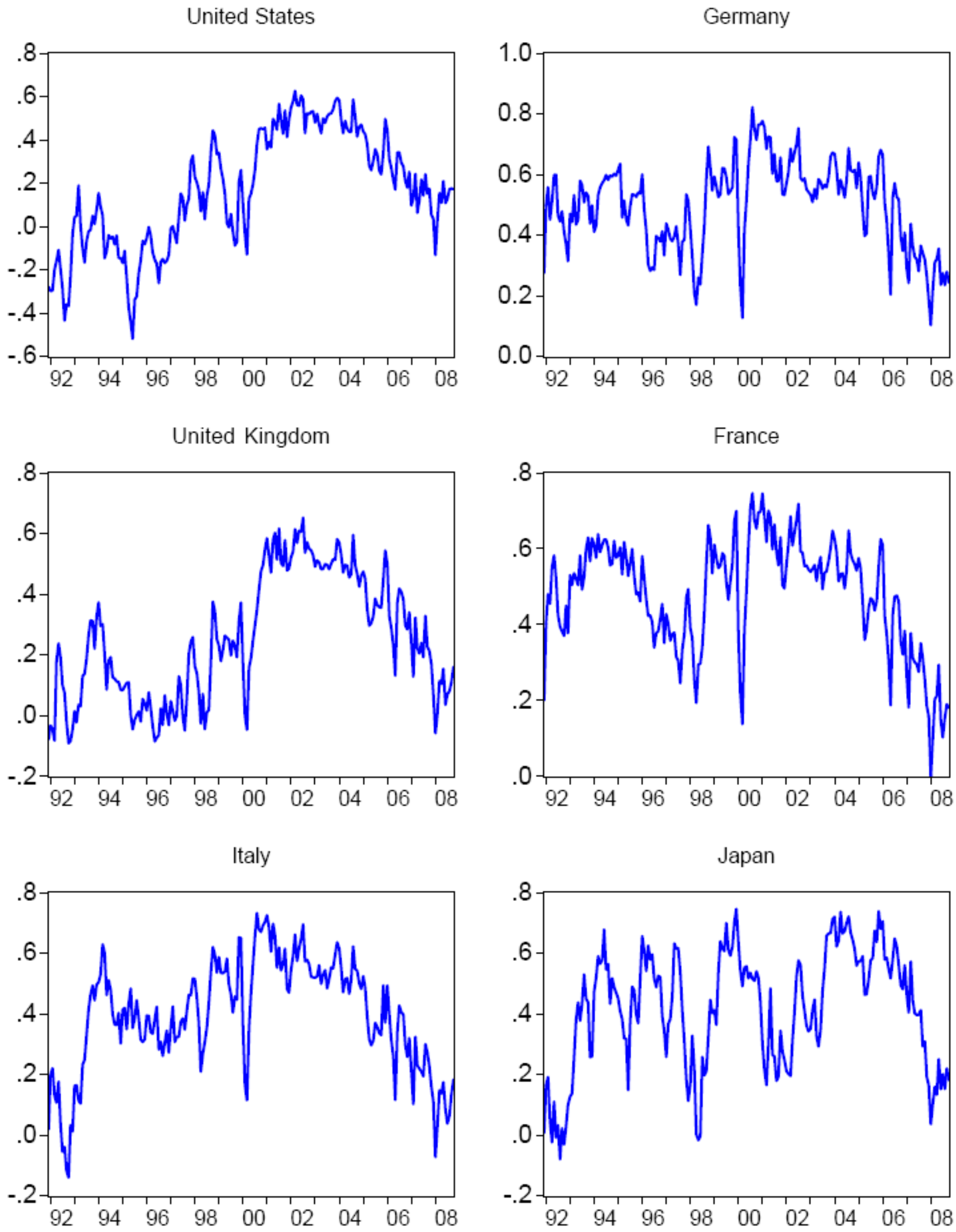
For estimation purposes, it is necessary to choose a large number of simulations S to obtain stable approximations. Choosing a large S, however, also slows down computation, especially in a non-linear ordered choice panel setting as encountered here. We therefore use 250 Halton draws to approximate the unknown expectation $\mathbb{E}_{v_i}[L_i|v_i]$. Halton draws can speed up computational performance considerably by spreading random numbers more effectively over the unit interval than standard random number generators so that the number of draws necessary can be reduced by as much as 90% (see Train, 1999).

Table 1: Descriptive statistics: Individual forecasters

	mean	median	max	min	std
Gender	0.94	1	1	0	0.24
Age	46.27	46	70	27	8.63
Job experience					
General	23.52	22	51	5	9.82
Financial Markets	20.48	19	50	5	9.19
Economics education	0.73	1	1	0	0.49
University degree	0.77	1	1	0	0.42
Doctoral degree	0.09	0	1	0	0.29

Notes: This table shows descriptive statistics for characteristics of roughly 450 individual forecasters in our sample. Gender is a dummy variable equal to one for a male forecaster. "Age" and "job experience" are measured in years and are reported as of 2008 to make forecasters comparable. "Economics education", "University degree", and "PhD" are dummy variables indicating a degree in economics, finance, or business, a degree granted from a university, or the completion of a doctoral degree, respectively.

Figure 1: Consensus forecasts



Notes: The figure shows time-series plots of monthly consensus stock market forecasts for the U.S., Germany, U.K., France, Italy, and Japan. The sample period is 12/1991 – 10/2008.

Table 2: Descriptive statistics: Consensus stock market forecasts

	US	GER	UK	FR	IT	JP
mean	0.17	0.50	0.25	0.48	0.39	0.42
median	0.17	0.53	0.23	0.52	0.41	0.45
max	0.63	0.82	0.65	0.75	0.73	0.75
min	-0.52	0.10	-0.09	0.00	-0.14	-0.08
std	0.27	0.14	0.21	0.15	0.18	0.20
skew	-0.22	-0.37	0.14	-0.73	-0.55	-0.50
kurt	2.13	2.62	1.78	3.00	2.84	2.38
ρ_{-1}	0.94	0.83	0.93	0.86	0.90	0.89
ADF	-2.57	-4.12	-2.83	-3.62	-3.36	-3.40
	(0.10)	(0.00)	(0.05)	(0.01)	(0.01)	(0.01)
Levin, Lin, Chu t				-4.43	(0.00)	
Im, Pesaran, Shin W				-5.08	(0.00)	

Notes: This table shows descriptive statistics for consensus stock market forecasts. ρ_{-1} denotes first order autocorrelations and ADF denotes the Augmented Dickey-Fuller test. The panel unit-root test of Levin, Lin, and Chu tests for a common unit root, whereas the test by Im, Pesaran, and Shin tests for individual unit-roots in a panel. P-values are in parentheses. The sample period is 12/1991 – 10/2008 on a monthly frequency.

Table 3: Random parameters panel ordered logit models – Results for $t > 0$

	U.S.			Germany		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
const.	0.923	-0.456	0.512	1.087	-0.175	-0.006
	[2.76]	[-2.97]	[4.25]	[13.02]	[-2.73]	[-0.09]
$\overline{\mathbf{e}}$	0.022	0.009	0.014	0.026	0.015	0.012
	[16.35]	[13.30]	[11.94]	[18.29]	[12.28]	[11.31]
e_{-1}		1.675	1.280		1.357	1.359
		[23.64]	[17.38]		[21.05]	[18.41]
PER			0.010			0.010
			[3.88]			[1.11]
TS			-0.064			-0.052
			[-3.71]			[-3.79]
IP			-0.004			-0.007
			[-0.60]			[-1.65]
CPI			-0.033			-0.025
			[-1.60]			[-0.76]
LD6			0.007			0.014
			[1.26]			[3.19]
IR3M			-0.097			-0.007
			[-4.68]			[-0.36]
R_{-1}			0.016			0.008
			[4.08]			[2.84]
R_{-6}			-0.010			-0.009
			[-4.84]			[-6.87]
σ (const.)	0.071	0.821	0.824	1.081	0.737	0.746
σ ($\overline{\mathbf{e}}$)	0.017	0.027	0.030	0.014	0.019	0.023
Log L	-31.47	-28.48	-28.33	-27.96	-26.13	-25.06
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pseudo R^2	0.11	0.18	0.19	0.12	0.17	0.18
obs	34,195	34,195	34,195	34,798	34,798	34,798

Notes: This tables shows regressions of individual stock market forecasts (e) on different sets of determinants for the U.S. and German stock market. "$\overline{\mathbf{e}}$" denotes the (lagged) consensus forecast of all forecasters, e_{-1} is the lagged individual forecast. PER denotes the aggregate market's price-earnings ratio, TS the term spread, IP the annual growth in industrial production, CPI the annual inflation rate. LD6 denotes the OECD's Index of leading indicators (6-month change), IR3M denotes the three months money market interest rate, R_{-1} and R_{-6} are lagged market returns over the last month and last six months, respectively. All determinants are measured at the end of the month just prior to the forecast. $\sigma()$ denotes the standard deviation of a random parameter, Log L reports the log likelihood (scaled by 10^{-4}). Estimates are for $t > 0$. Numbers in brackets show t-statistics.

Table 4: Marginal effects

	U.S.			Germany		
	down	unch.	up	down	unch.	up
$\overline{\mathbf{e}}$	-9.58	-4.26	13.84	-2.54	-3.67	6.50
e_{-1}	-15.87	-7.24	23.11	-10.57	-14.81	25.38
PER	-1.11	-0.51	1.62	-0.52	-0.74	1.33
TS	2.24	1.01	-3.25	0.94	1.35	-2.29
IP	0.34	0.17	-0.45	0.35	0.49	-0.91
CPI	0.63	0.29	-0.91	0.42	0.59	-1.01
LD6	-0.61	-0.30	0.91	-1.06	-1.54	2.60
IR3M	3.97	1.80	-5.77	0.20	0.26	-0.43
R_{-1}	-1.65	-0.74	2.40	-0.76	-1.01	1.89
R_{-6}	2.04	1.02	-3.07	2.11	3.01	-5.12
$P(e_i = j)$	23.7	37.3	38.9	13.4	24.2	62.2

Notes: Marginal effects of explanatory variables corresponding to specification (iii) and (vi) in Table 3 for the three forecasting categories “down”, “unchanged”, and “up”. The last row shows unconditional probabilities of forecasts being in one of the three categroies.

Table 5: Relative versus absolute evaluation

Panel A: Coefficient estimates

	U.S.			Germany		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
const.	0.174	0.310	0.544	0.009	-0.056	-0.142
	[3.13]	[3.42]	[4.63]	[-0.13]	[-0.87]	[-1.73]
$\bar{\mathbf{e}}$	0.021	0.018	0.017	0.017	0.016	0.014
	[13.29]	[10.04]	[9.76]	[11.15]	[9.32]	[7.70]
$\bar{\mathbf{e}} \times d_{rel}$	0.003	0.003	0.003	0.003	0.003	0.002
	[3.22]	[2.89]	[3.01]	[3.41]	[3.30]	[2.61]
$\bar{\mathbf{e}} \times d_{abs}$	-0.003	-0.004	-0.004	-0.006	-0.004	-0.004
	[-3.75]	[-4.12]	[-3.38]	[-4.86]	[-3.74]	[-4.51]
e_{-1}		1.160	1.154		1.199	1.304
		[18.09]	[17.91]		[15.66]	[14.20]
$e_{-1} \times d_{rel}$		-0.056	-0.051		-0.083	-0.052
		[-3.08]	[-3.22]		[-1.99]	[-1.56]
$e_{-1} \times d_{abs}$		0.144	0.137		0.111	0.156
		[6.32]	[7.20]		[2.71]	[3.48]
Controls	NO	NO	YES	NO	NO	YES
σ (const.)	0.643	0.939	0.926	0.712	0.573	0.650
σ ($\bar{\mathbf{e}}$)	0.019	0.017	0.016	0.012	0.010	0.009
σ ($\bar{\mathbf{e}} \times d_{rel}$)	0.006	0.004	0.004	0.009	0.007	0.007
σ ($\bar{\mathbf{e}} \times d_{abs}$)	0.010	0.013	0.011	0.011	0.008	0.006
Log L	-30.16	-28.71	-28.03	-29.01	-25.12	-24.47
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pseudo R^2	0.14	0.21	0.22	0.14	0.21	0.22
obs	34,195	34,195	34,195	34,798	34,798	34,798

Panel B: Marginal effects for specification (iii) and (vi)

	U.S.			Germany		
	down	unch.	up	down	unch.	up
$\bar{\mathbf{e}}$	-11.34	-9.18	20.52	-3.52	-6.44	9.96
$\bar{\mathbf{e}} \times d_{rel}$	-1.32	-0.81	2.13	-0.19	-1.04	1.43
$\bar{\mathbf{e}} \times d_{abs}$	2.70	1.62	-4.32	2.02	0.92	-2.94
e_{-1}	-14.25	-11.03	25.28	-8.11	-21.02	29.13
$e_{-1} \times d_{rel}$	0.64	0.49	-1.13	0.32	0.84	-1.16
$e_{-1} \times d_{abs}$	-1.69	-1.31	3.00	-0.97	-2.51	3.48
$P(e_i = j)$	23.70	37.30	38.90	13.40	24.20	62.20

Notes: The setup in Panel A is the same as in Table 3 but we also include interaction terms of consensus expectations $\bar{\mathbf{e}}$ and lagged individual expectations e_{-1} with two dummy variables indicating forecasters which are evaluated more in terms of relative performance (d_{rel}) and absolute performance (d_{abs}), respectively. Panel B shows marginal effects for specifications (iii) and (vi).

Table 6: Experienced versus inexperienced forecasters

Panel A: Coefficient estimates						
	U.S.			Germany		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
const.	0.165	-0.293	0.561	0.017	0.086	0.042
	[1.78]	[-2.56]	[3.76]	[2.12]	[2.18]	[0.63]
$\bar{\mathbf{e}}$	0.028	0.021	0.020	0.042	0.028	0.023
	[4.89]	[5.99]	[5.74]	[13.21]	[9.97]	[10.76]
$\bar{\mathbf{e}}$ × age	-0.013	-0.012	-0.011	-0.025	-0.025	-0.026
	[-2.74]	[-2.48]	[-2.52]	[-6.88]	[-6.32]	[-5.56]
e_{-1}		0.965	0.932		0.888	0.810
		[15.54]	[17.62]		[16.01]	[12.11]
e_{-1}× age		0.308	0.434		0.638	0.798
		[4.16]	[4.02]		[7.39]	[7.42]
Controls	NO	NO	YES	NO	NO	YES
σ (const.)	0.901	0.880	0.898	0.598	0.542	0.719
σ ($\bar{\mathbf{e}}$)	0.005	0.007	0.003	0.013	0.010	0.008
σ ($\bar{\mathbf{e}}$ × age)	0.029	0.026	0.033	0.011	0.012	0.007
Log L	-30.78	-28.99	-28.10	-29.79	-25.68	-25.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pseudo R^2	0.12	0.19	0.22	0.13	0.19	0.21
obs	34,195	34,195	34,195	34,798	34,798	34,798

Panel B: Marginal effects for specification (iii) and (vi)						
	U.S.			Germany		
	down	unch.	up	down	unch.	up
$\bar{\mathbf{e}}$	-13.31	-10.65	23.96	-5.09	-13.00	18.09
$\bar{\mathbf{e}}$ × age	4.62	0.99	-5.61	2.60	6.83	-9.43
e_{-1}	-11.38	-9.22	20.60	-5.06	-13.10	18.16
e_{-1}× age	-4.35	-3.53	7.88	-4.62	-11.97	16.59
$P(e_i = j)$	23.70	37.30	38.90	13.40	24.20	62.20

Notes: The setup in Panel A is the same as in Table 3 with an additional random parameter – the interaction of $\bar{\mathbf{e}}$ and e_{-1} with forecasters' age. Panel B shows marginal effects for specifications (iii) and (vi).

Table 7: Impact of individual deviations from lagged consensus beliefs

	U.S.			Germany		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
const.	0.301	-0.392	-0.721	0.975	0.372	0.110
	[1.88]	[-1.66]	[-3.23]	[6.36]	[2.99]	[1.02]
DEV	-2.205	-1.782	-1.567	-1.430	-1.195	-1.203
	[-17.54]	[-13.79]	[-12.39]	[-11.02]	[-8.33]	[-9.451]
e_{-1}		2.883	2.792		2.555	2.598
		[16.65]	[17.38]		[18.23]	[18.78]
Controls	NO	NO	YES	NO	NO	NO
σ (const.)	0.861	0.820	0.758	0.730	0.471	0.405
σ (DEV)	0.652	0.791	0.714	0.828	0.662	0.607
Log L	-32.49	-30.23	-29.61	-29.08	-25.45	-24.79
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pseudo R^2	0.10	0.18	0.20	0.12	0.17	0.19
obs	34,195	34,195	34,195	34,798	34,798	34,798

Panel B: Marginal effects for specification (iii) and (vi)						
	U.S. forecasters			German forecasters		
	down	unch.	up	down	unch.	up
DEV	14.33	10.32	-24.65	5.21	13.84	-19.05
e_{-1}	-15.40	-11.79	27.19	-9.21	-17.04	26.25
$P(e_i = j)$	23.70	37.30	38.90	13.40	24.20	62.20

Notes: The setup is the same as in Table 3 above but here we use *DEV* (the difference between lagged individual forecast and consensus forecast) as explanatory variable instead of lagged consensus forecasts $\overline{\mathbf{e}}$. We do not report coefficient estimates for the control variables to save space but indicate (in row “Controls”) whether or not they are included.

Table 8: Impact of market states

Panel A: US

	Market direction		Market volatility		Cons. trend		Cons. volatility	
	Bull	Bear	High	Low	Up	Down	High	Low
$\bar{\mathbf{e}}$	0.028	0.026	0.016	0.031	0.027	-0.004	0.033	0.016
	[8.04]	[2.56]	[3.23]	[8.71]	[7.42]	[-0.85]	[10.10]	[7.31]
e_{-1}	0.867	0.908	0.991	0.707	0.684	1.181	0.928	0.897
	[15.91]	[5.22]	[12.54]	[6.96]	[11.46]	[9.83]	[15.42]	[21.38]
Pseudo R^2	0.20	0.18	0.20	0.19	0.19	0.20	0.21	0.22
obs	26,811	7,263	16,865	17,233	22,719	11,476	17,065	17,112

Panel B: Germany

	Market direction		Market volatility		Cons. trend		Cons. volatility	
	Bull	Bear	High	Low	Up	Down	High	Low
$\bar{\mathbf{e}}$	0.029	0.022	0.019	0.036	0.031	-0.005	0.031	0.017
	[7.55]	[4.32]	[4.21]	[7.06]	[8.83]	[-0.88]	[9.97]	[7.02]
e_{-1}	0.854	0.994	0.908	0.799	0.588	1.449	0.898	0.813
	[10.38]	[6.35]	[11.77]	[5.93]	[6.13]	[11.25]	[11.01]	[8.94]
Pseudo R^2	0.21	0.16	0.18	0.23	0.24	0.21	0.21	0.20
obs	24,973	9,825	17,384	17,421	23,312	11,486	17,384	17,421

Notes: This table shows results for estimating the base regression (see Table 3, specifications (iii) and (vi)) on different sub-samples. “Market direction” shows results for splitting the full sample into bull and bear markets (positive or negative past 6-months aggregate market returns). “Market volatility” shows results for lagged six month market return volatility being above or below median volatility. “Cons. trend” divides the sample into a period of rising consensus forecasts (1992-2001) and falling consensus forecasts (2002-2008). “Cons. volatility’ divides the full sample into two sub-samples depending on whether lagged absolute consensus forecast changes are above or below median.

Web Appendix to accompany

Higher-order beliefs among professional stock market forecasters: Some first empirical tests

Jesper Rangvid Maik Schmeling Andreas Schrimpf

Table A.1: Correlation coefficients: Consensus stock market forecasts

	US	GER	UK	FR	IT	JP
US	1.00	0.46	0.88	0.33	0.62	0.29
GER		1.00	0.67	0.92	0.74	0.35
UK			1.00	0.57	0.71	0.34
FR				1.00	0.78	0.33
IT					1.00	0.49
JP						1.00

Notes: This table shows simple correlation coefficients for the consensus series of stock market forecasts.

Table A.2: Transition probabilities: Individual stock market forecasts

	USA				Germany		
	down	unch.	up		down	unch.	up
down	0.62	0.26	0.12	down	0.58	0.22	0.21
unch.	0.16	0.60	0.24	unch.	0.11	0.51	0.38
up	0.07	0.22	0.71	up	0.05	0.14	0.81
	UK				France		
	down	unch.	up		down	unch.	up
down	0.56	0.31	0.13	down	0.55	0.24	0.21
unch.	0.11	0.67	0.22	unch.	0.11	0.55	0.35
up	0.06	0.24	0.70	up	0.04	0.17	0.79
	Italy				Japan		
	down	unch.	up		down	unch.	up
down	0.53	0.28	0.18	down	0.51	0.30	0.18
unch.	0.11	0.58	0.31	unch.	0.12	0.56	0.33
up	0.05	0.17	0.78	up	0.05	0.20	0.75

Notes: This table shows transition probabilities for individual stock market forecasts. The transition is from a row to a column element, e.g. 0.26 in the upper left panel “USA” means that there is a 26% probability that a forecaster revises his expectation from “down” in period t to “unchanged” in period $t + 1$.

Table A.3: Random parameters panel ordered logit models – Estimates for $t = 0$

	U.S.			Germany		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
const.	0.451	0.341	0.409	1.087	-0.175	0.018
	[1.15]	[0.38]	[1.25]	[13.02]	[-2.73]	[1.38]
$\overline{\mathbf{e}}$	0.010	0.003	0.009	0.026	0.015	-0.009
	[2.01]	[1.40]	[1.89]	[13.29]	[12.28]	[-1.35]
PER			-0.032			-0.069
			[-0.94]			[-0.45]
TS			0.246			-0.029
			[1.22]			[-1.02]
IP			0.111			-0.029
			[1.68]			[-1.02]
CPI			-0.369			0.470
			[-1.59]			[2.61]
LD6			-0.128			0.024
			[-2.72]			[0.93]
IR3M			0.020			-0.294
			[0.12]			[-1.92]
R_{-1}			0.134			0.084
			[4.24]			[4.41]
R_{-6}			-0.024			-0.028
			[-1.59]			[-2.47]
σ (const.)	0.432	0.176	0.601	0.552	0.222	0.197
σ ($\overline{\mathbf{e}}$)	0.001	0.007	0.011	0.015	0.008	0.009

Notes: This table shows coefficient estimates of the initial period $t = 0$ for the regressions shown in Table 3.

Table A.4: Higher-order expectations for other countries

	UK	FR	IT	JP
const.	0.680	-0.188	-0.060	0.265
	[5.25]	[-1.67]	[-0.84]	[5.24]
$\overline{\mathbf{e}}$	0.031	0.035	0.021	0.010
	[9.12]	[13.85]	[7.58]	[4.75]
e_{-1}	0.984	0.868	1.010	1.020
	[16.73]	[14.13]	[16.37]	[18.05]
PER	0.004	0.015	0.013	0.004
	[0.86]	[3.18]	[3.81]	[3.31]
TS	-0.017	-0.029	-0.066	-0.072
	[-1.88]	[-1.96]	[-2.92]	[-2.97]
IP	-0.034	-0.007	-0.002	0.003
	[-3.84]	[-1.30]	[-0.32]	[0.96]
CPI	0.015	0.013	0.078	0.035
	[0.76]	[0.56]	[3.32]	[2.37]
LD6	0.012	0.008	0.022	0.028
	[2.13]	[2.13]	[2.25]	[6.88]
IR3M	-0.054	0.001	-0.044	-0.028
	[-3.75]	[0.12]	[-4.21]	[-2.72]
R_{-1}	0.015	0.008	0.009	0.016
	[4.35]	[3.30]	[3.72]	[6.51]
R_{-6}	-0.012	-0.008	-0.006	0.005
	[-6.96]	[-5.38]	[-5.80]	[3.44]
σ (const.)	0.746	0.796	0.788	0.690
σ ($\overline{\mathbf{e}}$)	0.018	0.009	0.021	0.005
Log L	-28.08	-24.31	-23.72	-25.67
	(0.00)	(0.00)	(0.00)	(0.00)
Pseudo R^2	0.24	0.20	0.21	0.21
obs	35,122	33,615	32,453	34,180

Notes: This table shows results analogous to specifications (iii) and (vi) in Table 3 for the United Kingdom, France, Italy, and Japan.