Irrationality of Macroeconomic Forecasts and Behavioral Characteristics of Forecasters

Paulina Ziembińska[1]

Abstract

Authors seek to determine the characteristics of professional macroeconomic forecasts. They find some explanations for the conventionally accepted irrationality of professional expectations in the cognitive heuristics introduced by Kahneman and Tversky (1974) such as cognitive dissonance, avoidance of regret, overconfidence and anchoring in previously formed scenarios. They extend test of Campbell and Sharpe (2009) for a different type of anchor – previously formulated predictive scenarios. Although the high correlation between forecasts and realizations hinders analysis, authors find empirical evidence that corroborate the new anchoring hypothesis. These findings change the interpretation behind behavioral reasons of forecaster irrationality, i.e. forecasters are overconfident. The anchoring effect accounts for up to 50% of the final forecast value and is highly significant for more than half of the most relevant macroeconomic variables for financial market participants.

Keywords: Evaluating forecasts; Forecasting profession; Judgmental forecasting; Macroeconomic forecasting

JEL Classification: C53, G02

[1] Faculty of Economic Sciences, University of Warsaw, pziembinska@wne.uw.edu.pl
Authors gratefully acknowledge Fulbright Junior Advanced Research Award and support from the University of Warsaw.

1 - Introduction

Professional forecasting data are one of the few available datasets regarded as the closest to the "true" expectations. They are influential for all economic agents decisions [e.g. Carroll (2003) shows that households expectations derive from news reports on professional forecasters views], for financial market movements [there is vast literature on macroeconomic announcements impact on asset prices – see e.g. Andersson at.al. (2009)] and for many other processes in the economy [e.g. central bank forecasts and portfolio allocation decisions as described in Elliot and Timmerman (2008)]. The focus of this article is to point to the behavioral similarities between professional forecasters and financial markets participants, which provides foundation for equivalent explanations of deviations from widely assumed forecasts rationality and financial markets efficiency. The latter is done by evaluation of macroeconomic forecasts prepared by professional forecasters available in Bloomberg system[2] and description of factors influencing the forecasting procedure, with the main focus on unmeasurable factors like judgmental adjustments and its behavioral foundations.

Although seeking for reasons of irrationality and inefficiency in behavioral characteristics is not new in the literature, this study sheds new light on the interpretation of these factors. In particular authors seek not only for theoretical, but also for quantitative justification for the new interpretation of professional macro forecasts cyclical errors. Following literature on both decision making and forecasts evaluation, they extend statistical test proposed by Campbell and Sharpe (2009) to find that professional macroeconomic forecasters are overconfident, anchor their predictions in the previously formulated predictive scenarios and strength of the anchoring effect is dependent on the economic conditions. These results open the door to deeper analysis of the impact of behavioral characteristics of forecasters on all the processes in the economy, which are influenced by professional macro forecasts.

The article proceeds as follows: section 2 provides theoretical background to forecasting procedures and behavioral heuristics, section 3

[2] Most studies involve data from various Surveys of Professional Forecasters (SPF) – in United States the longest series are provided by Philadelphia Fed, in Eurozone by European Central Bank. Bloomberg and Thomson Reuters compete in providing macroeconomic forecasts directly linked to financial markets (see e.g. Adams (2011)). Using Bloomberg data enables further analysis of impact of forecaster overconfidence on financial markets.

presents methodology for testing influence of behavioral characteristics on forecasts errors and is followed by the main results and conclusions.

2 - Theory

Elliot and Timmerman (2008) point to four factors constituting forecasting procedure: variables of interest and the information set containing available data; loss function; the family of forecasting models and finally - type of the outcome. This very formal attitude should be complemented with some practical issues. In practice loss function and type of the prepared forecast depend mostly on the purpose of forecasting. E.g. the most popular mean squared error loss might be inaccurate for trading purposes where asymmetry plays significant role and forecasts prepared for institutional planning purposes are mainly point forecasts, while some SPFs require range or distribution forecasts. Also family of possible models, although obviously depend on the forecaster's knowledge, in general might be assumed to include mainly simplified models[3]. Most importantly, however, what this formal procedure does not take into account is judgmental forecasting, which according to many studies is the key factor that determines the final performance of professional forecasts. The most evident proof comes from Batchelor (1990) survey among Blue Chip survey participants, who confessed to 42-54% contribution of judgment in their forecasts. Judgment may be responsible for the well documented outperformance of survey forecasts over econometric procedures and institutional forecasts like those prepared by OECD or IMF staff [see e.g. Batchelor (2000), Fildes and Stekler (2002)]. Judgment also provides the link between direct and indirect expectations modelling, and proves that “true” expectations cannot be obtained from the single equation.

Forecast adjustments may arise from estimation problems [e.g. structural breaks, measurement errors – see Fildes and Stekler (2002)]; wider information set available to forecaster [so called domain knowledge[4] e.g. from more insight into entrepreneur activities or closer links to monetary authorities], different prior beliefs [see e.g. Patton and Timmermann (2010)],

[3] Fuster et.al. (2011) points to pragmatic [i.e. ease and speed of implementation], psychological and statistical [i.e. reduced risk of overfitting] reasons behind common use of simplified models not only by laypeople, but also professionals, economists or statisticians.
[4] For references and wide review of the literature including experiments see Lawrence et al. (2006).

specific unmeasurable factors in the loss function, but also from personal decisions. In general process of adjusting model-based forecasts may be seen as decision making under uncertainty. Kahneman and Tversky (1974) prove that in such situation "people rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations" [p. 1124]. Here authors claim that judgmental forecasting is highly influenced by only few heuristics such as cognitive dissonance, regret avoidance, anchoring and overconfidence [for broader references to behavioral theory in macroeconomics and finance see Shiller (1999)].

In the standard round of forecasting, macroeconomic analyst uses econometric model based on most recent data and then adjusts model-generated result. The set of the available information includes all information from the previous forecasting round plus new time series information, new domain knowledge and previous market forecasts [so called consensus]. Judgmental adjustment may be based on any of these information depending on the forecasting strategy. There is a wide literature defining various types of strategies [note that these might be treated as the unmeasurable factors in the loss function], but they all have a common point which is reputation. Ottaviani and Srensen (2006) analyze situation when strategy is determined by forecasting contest with pre-specified rules; Croushore (1997) defines two strategies – avoiding unfavorable publicity by moving toward the consensus [herding behavior] and inversely – making unusually bold forecasts to stand out from others[5]; similarly Laster et.al. (1999) reports that those whose wages depend most on publicity, produce forecasts that differ most from the consensus. Wide discussion on the conservative and extreme forecasting strategies [both in academic journals and press] is provided by Lamont (2002), but empirical evidence on this matter is mixed. The reputational concern might be seen also from another perspective, especially for forecasters like those reporting to Bloomberg, whose predictions potentially influence longer term trading strategies and institution budgeting plans. In such cases forecaster is responsible not only for single economic variable, but for the whole economic scenario. This scenario should be consistent and should not be changed too often – on the one hand from mentioned reputational matters [for both forecaster and institution], on the other due to technical problems i.e. strategy implementation, financial plans disturbances.

[5] Herding behavior is much more common on financial markets [e.g. technical analysis]. In macroeconomic forecasting an obvious obstacle is simply not knowing the consensus - most of the surveys are published after all the data have been collected.

General situation when forecaster avoids giving signal that the previous predictions were wrong and therefore does not revise them frequently was called by Deschamps and Ioannidis (2013) "rational stubbornness".

Effectively forecaster avoids any evidence [newly published data or other information] that contradicts his beliefs or assumptions [*cognitive dissonance* resulting in *confirmation bias*] and does not change his scenario and predictions not to feel *regret* [similarly to investor holding for too long lost positions]. Inversely, when hardly interpretable information [in terms of previously stated scenarios] is revealed, forecaster has a tendency to see patterns in random data or dependence in irrelevant factors – depending on his ability to confirm previous forecasts (*representativeness*). Sticking to the previously formed scenarios (*anchoring*) can be therefore seen as the sign of *overconfidence* - excessive belief in ones abilities and correctness. These examples are consistent with the literature evidence [6] that overconfidence implies over- or under-reactions to news. Forecaster over-interprets data in line with his scenario and under-interprets the relevant, but not supporting information. The above reasoning proves that described unmeasurable factors, which in fact influence forecaster loss function, make him follow simple cognitive *heuristics* and *systematic biases* introduced by Kahneman and Tversky (1974)[7]. What should be stressed is that presented hypothesis is not only valid due to institutional requirements, but also supported by psychological profile of individuals working on high level positions in financial institutions - similar to financial markets participants [what is more - these two sets of individuals have a significant intersection].

Most of the research on the cross-section of economic forecasting and cognitive heuristics has been done experimentally, so it is only weakly linked to the professional, most influential data. Research on macroeconomic forecasts meeting rational expectations has produced mixed results [for references see e.g. Croushore (1997), Keane and Runkle (1990), Fildes and Stekler (2002)] with some empirical evidence of the heuristic-driven

[6] Mostly from financial markets studies, but behavioral patterns in many cases seem to be similar to economic forecasting– for overview see e.g. Shiller (1999).

[7] Some remarks on the pointed heuristics: although according to the first definition from Kahneman and Tversky (1974) anchoring means being influenced by suggestions [in their primary study – by initially presented values], here the past forecasts are treated as self-imposed suggestions, enhanced by cognitive dissonance and overconfidence; overconfidence is mostly tested in experiments, but many authors point to obvious overconfidence among investors, whose risky work relies on clients trust and confidence – the case of professional forecasters is very similar; representativeness in general is a tendency to categorize events as typical or representative of the well-known class.

adjustments in macro-finance forecasts in e.g. Deschamps and Ioannidis (2013), Campbell and Sharpe (2009), Batchelor and Dua (1990), Fujiwara at.al. (2013).

Following Stekler (2007) to understand how forecaster incorporates his judgment, it is useful to reverse the question from creating a forecast to evaluating it, because error patterns are the main source of information on forecaster behavior. In the literature there is strong evidence of systematic character of macro-forecast errors resulting from inability to predict turning points [after Fildes and Stekler (2002): "errors occurred when the economy was subject to major perturbations, just the times when accurate forecasts were most needed" (p.442)]. Therefore authors test the hypothesis of dependence between error patterns and business cycle, corroboration of which opens the door to broader policy and business applications.

3 - Methodology

There are many methods to evaluate forecasts[8] based on ex post criterion functions connected with loss specification and rationality testing [in the literature there is no clear division between rationality, efficiency and optimality – all of these terms apply to the best forecast constructed under chosen loss function based on the all available information - for discussion see e.g. Nordhaus (1987)][9].

Before formal methods are introduced, it is worth reminding that authors assume MSE loss function and all factors which might influence its different form[10] are included in the judgmental adjustment term. This is based on the belief that similarly to choosing simplified models professional forecasters tend to measure their errors under MSE loss function, which enables easy computation [via least squares techniques] and evaluation, while formal econometric testing of more sophisticated loss functions is rather

[8] For overview of the theoretical results see e.g. Elliott and Timmermann (2008) and empirical: Fildes and Stekler (2002) or Heilemann and Stekler (2010).

[9] It should be noted that Bloomberg dataset does not include forecast revisions, which are necessary for many evaluation tests.

[10] The reasons behind more sophisticated loss function may stem from e.g. financial reasons, different utilization of forecasts [e.g. asymmetric or directional loss function for trading purposes], psychological features or strategic behavior [especially in not anonymous surveys presented publicly e.g. in Bloomberg]. For references see e.g. Elliott et.al. (2008). Note that most of these factors are unmeasurable and their inclusion in the loss function might be inaccurate.

academic exercise than work done in financial institutions. In this framework the optimality assessment should be done carefully as standard tests depend on the actual loss function [for discussion see e.g. Patton and Timmermann (2003)] and unmeasurable incentives might blur the picture of the definition.

Optimality testing under MSE loss function is the starting point for this analysis. Standard testing procedure incorporates unbiasedness and orthogonality tests introduced by Minzer and Zarnowitz (1969) based on the following regression:

$$x_{it} = a_i + b_i f_{it} + \varepsilon_{it}, \qquad (1)$$

where x_i denotes release of the i-th variable and f_{it} its one-period forecast [prepared at time (t-1) for the announcement at time t], t=1,…,T denotes sample size. The test of $H_0 : a_i = 0, b_i = 1$ combined with the analysis of errors autocorrelation is the most popular rationality test[11]. Additionally one may check forecasts efficiency - characteristic equivalent to effective utilization of the available information, e.g. with following regression:

$$s_{it} = x_{it} - f_{it} = \alpha_i + \beta_i x_{it-1} + \delta_i DBC_t + \theta_i z_{t-1} + \varepsilon_{it} \qquad (2)$$

where s denotes surprise [i.e. forecast error], DBC is dummy equal to one in recession period as determined by NBER and z is OECD composite leading indicator [chosen on purpose from the source different than Bloomberg to avoid overlapping variables].

Rationality relies on a very strict assumptions that cannot be met in reality [e.g. knowledge of the true model of the economy, access to all available information and ability to process it]. This together with adjustment effect constitute reasons behind systematic errors. For professional forecasters, who are the closest to meet rational expectations assumptions due to experience and substantive preparation, the judgmental factor seems to be of highest importance. Factors that determine how experts adjusts prediction from the econometric model may be divided into two categories: wider unmeasurable information set available to forecaster, which would lead to his

[11] Examples of empirical papers include Baghestani and Kianian (1993), Zarnowitz (1985). In some settings also condition of uncorrelated forecast revisions might be added, but here data contain only one-period forecasts.

systematically better performance[12] or specific factors influencing his loss function, which beside other reasons result in behavioral heuristics as described in the previous section. Most of the literature on heuristics introduced by Kahneman and Tversky (1974) focuses on experimental testing procedures [for review see e.g. Lawrence et.al (2006)], while only few tests were introduced for macroeconomic forecasting data.

One, which does not incorporate forecast revisions, is the anchoring test proposed by Campbell and Sharpe (2009) [CS hereafter]. They decompose forecast into forecaster's unbiased prediction and anchor [which in their original paper is the average value of the forecasted series over previous months] allowing weights to indicate strength of the anchoring effect:

$$f_{it} = \lambda_i E(x_{it} \mid I_{t-1}) + (1-\lambda_i)\cdot anchor_{it} \qquad (3)$$

This leads to testing regression of the form (for details see the CS paper):

$$s_{it} = \gamma_{0i} + \gamma_i (f_{it} - anchor_{it}) + \varepsilon_{it} \qquad (4)$$

where $H_0 : \gamma_i > 0$ is equivalent to significantly positive anchoring effect with relative weight: $1-\lambda_i = \frac{\gamma_i}{1+\gamma_i} \in (0,1)$. To check strength of the anchoring effect in the business cycle, interaction with the recession dummy variable is added to the above regression:

$$s_{it} = \gamma_{0i} + \gamma_{1i}(f_{it} - anchor_{it}) + \gamma_{2i}(f_{it} - anchor_{it}) \times DBC_t + \varepsilon_{it}. \qquad (5)$$

CS test was criticized by Hess and Orbe (2013), who argue that it incorporates only univariate time series information and therefore is biased to the superior information processing abilities. The basic idea of this critique is based on comparisons of forecasts generated by ARIMA models with survey errors and proof that they incorporate information correlated with additional macroeconomic series. It seems obvious [which is also argued in this paper] that forecaster incorporates much wider information set than univariate series of the forecasted variable, which under rationality assumption should only

[12] Literature proves that this is not the case and most often consensus outperforms individual forecasts, see e.g. Zarnowitz (1984).

create independent forecast errors. This however is not the case as errors from survey forecasts reveal cyclicality. Here authors argue that the reason behind such pattern is anchoring in previous scenarios, which causes forecasts underestimation in growing periods and overestimation in recession periods – behavior similar to macroeconomic variables, which Hess and Orbe (2013) use as a contrarian proof for validity of the CS test. It should be also noted that CS test does not include any assumptions about the data generating process and information set – the only problem with reasoning [as pointed also by Hess and Orbe (2013)] is the anchor variable, for which CS choose average value of the forecasted series over few previous periods [they argue that forecasters do not rely on their personal information set and therefore are anchored in the recently published data]. Here authors claim the opposite - that forecasters are overconfident, stick to previously formed scenarios and do not adjust their views to recently published data in a sufficient way to form rational forecasts. Therefore here the averages of past forecasts are used as anchoring variable [which also bypasses the above critique of CS procedure[13]].

Authors use ordinary least squares [OLS] with Newey-West standard errors with three lags to adjust for serial correlation of the error term resulting from the situations when forecaster might not know the forecast error from the previous period [due to the probable different timing of forecast submission or announcement release].

4 – Data

Proposed dataset includes 23 macroeconomic variables for United States regarded as most relevant for Bloomberg users [i.e. relevance index published by Bloomberg measures popularity of each variable as scaled number of alerts set for its announcement relative to all alerts set for all events in the country]. The chosen dataset consist mostly of monthly data [the only exceptions are quarterly GDP data and weekly initial jobless claims]. By one-period ahead forecast authors mean prediction of the nearest announcement in

[13] It might be shown formally that decomposition of γ coefficient introduced by Hess and Orbe (2013) is not valid when anchoring process is based on forecasting data [for brevity calculations are not included].

the respective frequency[14]. Following Zarnowitz (1984) data are transformed to stationary series.

Table 1. Results of the rationality and efficiency tests [equations (1) and (2)]

#	relevance	variable	Rationality test									Efficiency test	
			a	std.err.(a)	b	std.err.(b)	R2(adj)	DW	H0:a=0	H0:b=1	H0:a=0,b=1	R2adj	F-test p-value
1	98,1	**Change in Nonfarm Payrolls**	-18,32	7,17	0,97	0,04	81,5%	**0,09**	**0,01**	0,45	**0,01**	2,5%	0,25
2	99,1	Initial Jobless Claims	2,78	3,90	0,99	0,01	94,5%	**0,00**	0,42	0,56	0,43	1,5%	**0,01**
3	96,3	GDP QoQ (Annualized)	0,03	0,06	1,00	0,02	96,4%	0,95	0,57	0,80	0,82	0,7%	0,77
4	95,4	Consumer Confidence	-0,40	1,10	1,00	0,01	97,8%	0,44	0,75	0,77	0,95	0,8%	0,73
5	94,4	ISM Manufacturing	1,01	1,44	0,98	0,03	89,0%	0,62	0,49	0,53	0,65	0,5%	0,84
6	93,5	**Consumer Price Index (MoM)**	-0,05	0,01	1,21	0,05	86,6%	0,85	**0,00**	**0,00**	**0,00**	5,5%	**0,03**
7	92,6	U. of Michigan Confidence	1,26	0,90	0,98	0,01	96,2%	0,79	0,21	0,15	0,24	1,2%	0,29
8	90,7	**Durable Goods Orders**	-0,13	0,17	1,42	0,11	46,6%	0,38	0,52	**0,00**	**0,00**	4,6%	**0,06**
9	89,8	New Home Sales	-8,89	7,52	1,02	0,01	96,8%	0,30	0,52	0,26	0,28	0,4%	0,87
10	88,9	**Advance Retail Sales**	-0,03	0,04	1,23	0,13	68,1%	**0,01**	0,59	**0,00**	**0,01**	9,6%	**0,01**
11	88,3	**Unemployment Rate**	0,04	0,04	0,99	0,01	100,0%	0,33	0,35	**0,11**	**0,05**	6,0%	**0,02**
12	88,0	Housing Starts	-4,55	9,66	1,01	0,01	98,4%	**0,01**	0,80	0,43	0,32	1,7%	0,44
13	87,0	Existing Home Sales	-0,09	0,11	1,02	0,02	95,0%	0,13	0,56	0,54	0,81	0,8%	0,90
14	86,1	**Industrial Production**	-0,08	0,03	1,23	0,08	69,1%	0,34	**0,01**	**0,00**	**0,00**	3,5%	0,13
15	85,2	**Producer Price Index (MoM)**	-0,08	0,05	1,52	0,15	74,8%	0,24	**0,03**	**0,00**	**0,00**	2,2%	0,32
16	84,3	**Leading Indicators**	-0,01	0,01	1,17	0,04	85,6%	0,83	0,66	**0,00**	**0,00**	8,7%	**0,00**
17	83,3	**Personal Spending**	-0,05	0,03	1,15	0,06	88,5%	0,66	**0,00**	**0,00**	**0,00**	5,2%	**0,04**
18	83,3	Personal Income	0,04	0,02	0,99	0,03	68,5%	0,88	**0,10**	0,92	0,12	2,4%	0,27
19	82,4	**Factory Orders**	0,03	0,05	1,06	0,03	91,6%	0,18	0,54	**0,02**	**0,05**	1,1%	0,63
20	81,5	Trade Balance	-0,79	0,56	0,98	0,01	96,2%	**0,06**	0,28	0,27	0,54	5,9%	**0,02**
21	80,6	Empire Manufacturing	1,19	1,11	0,93	0,05	65,0%	0,57	0,33	0,30	0,53	1,5%	0,66
22	79,6	**Wholesale Inventories**	0,03	0,04	1,14	0,09	52,9%	**0,00**	0,55	0,12	**0,07**	9,2%	**0,00**
23	79,1	**Chicago Purchasing Manager**	0,16	1,97	1,01	0,04	75,0%	**0,01**	0,95	0,82	**0,10**	2,6%	0,25

Last four columns in ratonality part present p-values from Durbin-Watson test and tests of presented hypothesis H0.

Bold selection of variables mean rejection of rationality hypothesis.

As mentioned before professional forecasting surveys introduce only one measurement error connected with announcements revisions [i.e. forecast error is determined by which announcement was truly predicted]. The other data-related problem is choice of proper summary statistic for consensus forecast. Analysis of root mean square error statistic (RMSE) and Diebold and Mariano (2002) statistic to compare predictive accuracy of forecasts revealed no significant difference between median and mean [Bloomberg system favors median aggregation – this is the benchmark prediction for people who

[14] Although Bloomberg system enables forecasters to update their forecasts in real time, we can assume that they are prepared with different purposes and are rarely updated in the system after the survey was sent.

observe publications e.g. traders] and slightly better performance of forecasts compared to initial announcement rather than to finally revised data. Therefore error is defined as difference between median consensus (or individual forecast) and initial announcement [15] (due to only one-period horizon and taking as benchmark value of initial announcement authors avoid any overlapping data problems).

Table 2. Results of the anchoring test [equation (4)]

	1-month anchor					3-months anchor					6-months anchor					12-months anchor				
#	γ		1-λ	R2(adj)	BIC	γ		1-λ	R2(adj)	BIC	γ		1-λ	R2(adj)	BIC	γ		1-λ	R2(adj)	BIC
1	-0,16			36,2%	960,8	-0,26			62,6%	856,7	-0,29			72,6%	812,2	-0,31			72,6%	755,6
2	-0,13			1,0%	-2278,2	-0,17			1,5%	-2251,1	-0,15			0,7%	-2213,5	-0,13			0,7%	-2173,1
3	0,02			0,3%	249,3	0,04			1,6%	239,1	0,03	*	0,03	0,4%	232,9	0,01			0,4%	220,1
4	0,36	***	0,26	5,1%	-360,3	0,71	***	0,42	15,7%	-373,0	0,82	***	0,45	21,4%	-369,7	0,92	***	0,48	21,4%	-354,9
5	-0,11			0,5%	-589,8	0,08			0,3%	-580,8	0,13			1,0%	-568,3	0,16			1,0%	-550,1
6	0,07	**	0,06	1,8%	-186,6	0,16	***	0,14	11,5%	-199,7	0,15	***	0,13	13,7%	-191,4	0,18	***	0,15	13,7%	-190,7
7	0,19	*	0,16	0,7%	-1138,1	0,47	**	0,32	3,2%	-1128,6	0,69	***	0,41	6,5%	-1113,2	0,69	**	0,41	6,5%	-1057,9
8	0,21	***	0,18	3,8%	787,9	0,38	***	0,28	7,1%	759,7	0,47	***	0,32	7,8%	713,3	0,42	***	0,30	7,8%	659,0
9	0,14			0,8%	-360,7	0,28	*	0,22	2,0%	-350,6	0,38	**	0,28	3,9%	-336,7	0,43	**	0,30	3,9%	-321,3
10	0,18	**	0,15	8,7%	266,4	0,20	*	0,17	6,7%	267,0	0,06			1,9%	209,5	0,10	**	0,09	1,9%	196,3
11	-0,19			1,1%	-723,9	-0,19			0,9%	-714,7	-0,07			0,1%	-702,1	0,05			0,1%	-676,4
12	0,15	*	0,13	1,3%	-436,9	0,34	***	0,26	4,3%	-436,1	0,34	***	0,25	4,8%	-424,9	0,39	***	0,28	4,8%	-408,7
13	-0,12			1,5%	-287,6	0,16	*	0,14	2,4%	-280,4	0,27	**	0,22	6,3%	-270,0	0,30	**	0,23	6,3%	-245,0
14	0,23	**	0,19	5,6%	137,4	0,36	***	0,26	10,6%	127,4	0,30	***	0,23	11,1%	122,6	0,30	***	0,23	11,1%	115,2
15	0,44	***	0,31	20,1%	199,4	0,45	***	0,31	20,9%	196,5	0,39	***	0,28	20,7%	200,1	0,43	***	0,30	20,7%	191,6
16	0,14	***	0,12	9,2%	-106,6	0,22	***	0,18	17,2%	-111,3	0,21	***	0,17	13,8%	-95,1	0,20	***	0,17	13,8%	-77,4
17	0,10	***	0,09	10,7%	-127,7	0,14	***	0,12	12,4%	-129,8	0,16	***	0,14	12,0%	-126,1	0,15	***	0,13	12,0%	-116,4
18	0,00			0,0%	41,3	0,01			0,0%	42,5	0,01			0,0%	44,8	0,00			0,0%	48,8
19	0,03	**	0,03	1,9%	339,5	0,04	**	0,04	2,1%	337,0	0,06	**	0,05	2,9%	332,0	0,06	**	0,05	2,9%	319,5
20	0,03			0,1%	-359,7	0,08			0,4%	-355,6	0,04			0,7%	-354,2	0,12			0,7%	-348,6
21	0,77	**	0,43	20,4%	484,9	1,00	**	0,50	23,7%	473,5	1,09	**	0,52	32,1%	444,7	1,22	**	0,55	32,1%	420,2
22	-0,49			3,7%	222,2	-0,21			0,8%	225,9	0,00			0,9%	225,4	0,12			0,9%	219,7
23	0,15	*	0,13	1,1%	-357,1	0,25	**	0,20	2,1%	-358,2	0,25	**	0,20	2,5%	-350,9	0,26	**	0,21	2,5%	-337,0

*** denotes 1% significance level, ** 5% and * 10% - for one sided tests: H1: $\gamma>0$.

For bolded viariables anchoring in past forecasting scenario (6-month) is significant factor in explaining forecast errors.

[15] For robustness authors check all configurations of errors available in this dataset – estimation results are fairly indifferent. In general only for six variables revisions were proved to be significant. For brevity these statistics are not presented.

Results

For more than half of the variables authors reject rationality and efficiency [see Table 1]. Past information can explain up to 10% of forecast errors variation. This results prove that for many variables forecasters do not incorporate all available information or change it depending on their interpretation of facts. On average they are over-optimistic about economic conditions.

Table 3. Results of anchoring test with recession dummy [equation (5)]

variable	**γ0**		std.err.(γ0)	**γ1**		std.err.(γ1)	**γ2**		std.err.(γ2)	R2(adj)	F-test p-value
Change in Nonfarm Payrolls	0,16		0,19	0,01		0,05	-0,27	***	0,21	0,7%	0,58
Initial Jobless Claims	0,00	***	0,00	-0,16		0,06	0,03	**	0,18	1,0%	0,03
GDP QoQ (Annualized)	-0,02	***	0,04	**0,05**	***	0,02	**-0,07**	***	0,06	3,8%	0,05
Consumer Confidence	0,00	***	0,00	0,63		0,15	0,65		0,41	21,2%	0,00
ISM Manufacturing	0,00	***	0,00	-0,03	***	0,14	0,50		0,20	2,9%	0,10
Consumer Price Index (MoM)	0,00	***	0,01	**0,20**	***	0,04	**-0,08**	***	0,07	10,9%	0,00
U. of Michigan Confidence	0,00	***	0,00	**0,92**	***	0,24	**-0,63**	***	0,65	7,9%	0,00
Durable Goods Orders	-0,11	***	0,16	0,43	**	0,11	0,18		0,30	9,5%	0,00
New Home Sales	0,00	***	0,01	0,38		0,22	0,14		0,41	3,5%	0,07
Advance Retail Sales	0,00	***	0,04	0,05		0,07	0,03	**	0,14	0,9%	0,57
Unemployment Rate	0,00	***	0,00	-0,17	***	0,15	0,30		0,35	0,7%	0,60
Housing Starts	0,01	***	0,00	**0,34**	**	0,11	**0,02**	**	0,31	3,5%	0,06
Existing Home Sales	0,00	***	0,00	**0,29**	**	0,14	**-0,25**	***	0,40	5,9%	0,11
Industrial Production	-0,03	***	0,02	0,15	***	0,08	0,68		0,36	15,9%	0,00
Producer Price Index (MoM)	0,03	**	0,04	**0,65**	***	0,13	**-0,47**	***	0,17	24,5%	0,00
Leading Indicators	0,01	***	0,01	**0,20**	***	0,05	**0,03**	**	0,07	14,5%	0,00
Personal Spending	0,00	***	0,01	0,12		0,05	0,06	*	0,09	13,7%	0,00
Personal Income	0,05	*	0,02	-0,03		0,03	0,52		0,24	4,0%	0,04
Factory Orders	0,05	*	0,05	0,04		0,03	0,08	*	0,06	4,3%	0,03
Trade Balance	0,00	***	0,01	-0,12	**	0,17	0,23		0,23	0,7%	0,60
Empire Manufacturing	-0,15	***	0,20	1,06		0,60	2,48		1,18	31,2%	0,00
Wholesale Inventories	0,10		0,04	-0,35	**	0,18	0,92		0,26	6,1%	0,01
Chicago Purchasing Manager	0,01	**	0,01	**0,31**	***	0,16	**-0,17**	***	0,27	2,5%	0,15

*** denotes 1% significance level, ** 5% and * 10%. For H1: $\gamma0 \neq 0$, H1:$\gamma1 > 0$, H1:$\gamma2 \neq 0$.
F test is joint significance test in regression (5).
Bolded variables reveal significant anchoring effect dependend in magnitude on the business cycle.

For anchoring tests it should be noted that the difference in results for the two chosen anchors is very small as both series of forecasts and realizations have similar patterns [all information criteria slightly favor forecasting anchor]. However, there are two reasons to support anchoring in past forecasting scenarios rather than in past realizations [which is of less importance in econometric tests, but of highest in understanding judgmental forecasting process and its motives]. Firstly, from the theoretical perspective, as argued above it seems reasonable to assume that professional forecasters in financial institutions are overconfident similarly to financial market participants [who they often in fact are]. Secondly and more importantly – results from the empirical analysis [see Table 2] prove that the longer the anchor [i.e. the longer historical average], the stronger its impact on errors, which can be only justified in case of sticking to previously formed forecasting scenarios[16].

Table 4. Tests for individual data

				anchoring in:							
				past forecasts					past realizations		
variable	#forecasters	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Change in Nonfarm Payrolls	177	66,7%	4,4%	16,5%	0,07	0,58	12	46,4%	14,1%	0,08	0,55
Initial Jobless Claims	772	94,7%	34,0%	1,6%	0,20	0,36	1	18,1%	2,4%	0,07	0,15
GDP QoQ (Annualized)	179	33,8%	13,8%	11,5%	0,03	0,14	12	87,7%	11,5%	0,04	0,10
Consumer Confidence	176	76,7%	11,7%	30,7%	0,11	0,52	1	50,0%	25,7%	0,07	0,29
ISM Manufacturing	180	67,2%	14,9%	2,1%	0,25	0,55	1	11,9%	8,2%	0,07	0,44
Consumer Price Index (MoM)	180	60,6%	59,1%	9,6%	0,06	0,30	1	31,8%	4,5%	0,10	0,20
U. of Michigan Confidence	301	56,8%	25,9%	5,1%	0,17	0,46	1	9,9%	2,2%	0,06	0,32
Durable Goods Orders	168	23,8%	52,4%	22,0%	0,16	0,45	12	44,4%	26,8%	0,16	0,47
New Home Sales	169	22,0%	5,1%	6,5%	0,19	0,38	1	1,7%	7,9%	0,07	0,32
Advance Retail Sales	125	20,3%	65,6%	10,2%	0,20	0,31	12	54,7%	14,8%	0,10	0,37
Unemployment Rate	177	61,2%	83,6%	3,8%	0,26	0,49	1	4,5%	5,0%	0,10	0,24
Housing Starts	163	29,5%	16,4%	14,2%	0,11	0,55	1	8,2%	38,5%	0,05	0,29
Existing Home Sales	80	18,2%	2,3%	1,8%	0,30	0,33	1	6,8%	5,4%	0,08	0,30
Industrial Production	180	50,0%	37,5%	18,1%	0,12	0,33	12	60,9%	18,1%	0,11	0,42
Producer Price Index (MoM)	167	81,5%	26,2%	54,3%	0,12	0,60	12	93,8%	53,1%	0,10	0,58
Leading Indicators	176	36,2%	57,4%	24,3%	0,04	0,40	1	68,1%	22,6%	0,04	0,21
Personal Spending	177	43,8%	39,1%	28,5%	0,06	0,38	1	71,9%	31,9%	0,05	0,33
Personal Income	181	66,1%	59,7%	2,7%	0,12	0,28	12	45,2%	4,1%	0,12	0,23
Factory Orders	182	35,1%	19,3%	14,3%	0,05	0,25	12	1,8%	19,0%	0,05	0,25
Trade Balance	179	32,3%	38,7%	4,8%	0,05	0,20	1	16,1%	8,3%	0,06	0,20
Empire Manufacturing	106	71,4%	5,7%	27,8%	0,41	0,81	12	80,0%	22,2%	0,34	0,74
Wholesale Inventories	181	41,9%	64,5%	9,1%	0,17	0,26	12	22,6%	12,1%	0,19	0,32
Chicago Purchasing Manager	178	66,0%	44,0%	16,8%	0,12	0,38	1	10,0%	48,6%	0,07	0,30

Percentage values should be interpreted as a fraction of individual forecasters who meet column criteria:
(1) rejection of H0:a=0,b=1; (2) rejection of efficiency hypothesis;
(3) rejection in favor of H1:$\gamma>0$ in anchoring test with forecasting anchor;
(7) significance of anchoring models with recession dummy;
(8) rejection in favor of H1:$\gamma>0$ in anchoring test with past data anchor.
Other columns: (6) anchor length chosen with BIC criterion (in months);
(4) and (9) are min(1-λ) ; (5) and (10) are max(1-λ) - from anchor of past forecasts and past realizations respectively.

[16] Underconfidence would be supported by anchoring only in the latest available information.

Results in Table 2 show that not only information criteria choose longer anchor, but also significance of the anchoring effect and its magnitude in most cases increases with the length of anchor. Its weight [i.e. $1-\lambda$] can account for up to 55% of forecasted value. Adding recession dummy gives an interesting result of significantly smaller anchoring effect during recession time [in Table 3 for variables with significant anchor almost all $\gamma 2$ coefficients are negative].

Finally, aggregated results [for consensus forecasts] are compared with the individual forecasting data [see Table (4)]. First, notice that not all individuals who constitute aggregated consensus forecasts are included in analysis presented in Table (4) because of availability of the lagged data. Therefore, individual results should be interpreted as representing forecaster with experience of more than two years. In general tests for individual data give weaker support for irrationality and anchoring, which proves that more experienced forecasters provide better predictions.

Rationality and efficiency tests give similar results to aggregated data. Interestingly, significant anchoring effect ($\gamma>0$) occurs similarly for both anchors - in past forecasts and past realizations, but magnitude of anchor's influence on the forecast [i.e. $(1-\lambda)$] is on average much higher for forecasting anchor [not for realizations]. Similarly to the aggregated results, for variables with significant anchoring effect, information criteria choose significantly longer average [compare the 6^{th} column of Table (4)]. This corroborates presented hypothesis of overconfidence also on the disaggregated series.

5 - Conclusions

Authors find reasons behind traditionally understood irrationality of professional expectations in cognitive heuristics introduced by Kahneman and Tversky (1974) such as cognitive dissonance, avoidance of regret, anchoring in previously formed scenarios and overconfidence. They extend the test proposed by Campbell and Sharpe (2009) for a different type of anchor – previously formulated predictive scenarios, which finds justification in both theory and empirics. Although correlation between past data and past forecasts is high, they find empirical evidence that supports the overconfidence hypothesis and therefore changes the interpretation behind behavioral reasons of forecaster irrationality, i.e. inversely to underconfidence assumption made by CS, forecasters are proved to be overconfident. Authors find that anchoring effect accounts for up to 50% of the forecasted value and is highly significant for more than half of the most relevant variables for financial market participants. These results are consistent for both aggregated and individual data. Finally, authors find that impact of anchoring effect decreases during recessions, which proves that in financial institutions costs of macroeconomic forecast errors increase during difficult times.

References

Adams R., 2011 Nov. 14. Financial-Data Suppliers Brace for Ebb. *Wall Street Journal*. Retrieved 13 Apr. 2014 from http://online.wsj.com/news/articles/SB10001424052970203537304577031870359843842

Andersson M., L.J. Overby and S. Sebestyén, 2009. Which news moves the euro area bond market? *German economic review*, 10(1), 1-31.

Batchelor R., 2000. The IMF and OECD versus consensus forecasts. City University Business School, London, August.

Batchelor R., 2007. Bias in macroeconomic forecasts. *International Journal of Forecasting*, 23(2), 189–203.

Batchelor R. and P. Dua, 1990. Forecaster ideology, forecasting technique, and the accuracy of economic forecasts. *International Journal of Forecasting*, 6(1), 3-10.

Campbell S.D. and S.A. Sharpe, 2009. Anchoring bias in consensus forecasts and its effect on market prices. Journal of Financial and Quantitative Analysis 44.02: 369-390.

Carroll C.D., 2003. Macroeconomic expectations of households and professional forecasters. *The Quarterly Journal of economics*, 118(1), 269-298.

Croushore D., 1997. The Livingston Survey: still useful after all these years. *Business Review*, (Mar), 15-27.

Deschamps B. and C. Ioannidis, 2013. Can rational stubbornness explain forecast biases? *Journal of Economic Behavior & Organization*, 92, 141–151.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. Journal of Business & economic statistics, 20(1).

Elliott G., I. Komunjer and A. Timmermann, 2008. Biases in macroeconomic forecasts: irrationality or asymmetric loss? *Journal of the European Economic Association* 6.1: 122-157.

Elliott G. and A. Timmermann, 2008. Economic forecasting. *Journal of Economic Literature*, 3-56.

Fildes R. and H. Stekler, 2002. The state of macroeconomic forecasting. *Journal of Macroeconomics*, 24(4), 435–468.

Fujiwara I., H. Ichiue, Y. Nakazono and Y. Shigemi, 2013. Financial markets forecasts revisited: Are they rational, stubborn or jumpy? *Economics Letters*, 118(3), 526-530.

Fuster A., B. Hebert and D. Laibson, 2011. Natural expectations, macroeconomic dynamics, and asset pricing (No. w17301). National Bureau of Economic Research.

Heilemann U. and H. Stekler, 2010. Perspectives on Evaluating Macroeconomic Forecasts. *Working Papers 2010-002. The George Washington University, Department of Economics, Research Program on Forecasting*, (46).

Hess D. and S. Orbe, 2013. Irrationality or efficiency of macroeconomic survey forecasts? Implications from the anchoring bias test. Review of Finance 17.6: 2097-2131.

Kahneman D. and A. Tversky, 1974. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.

Keane M.P. and D.E. Runkle, 1990. Testing the rationality of price forecasts: New evidence from panel data. *The American Economic Review*, 714-735.

Lamont O.A., 2002. Macroeconomic forecasts and microeconomic forecasters. *Journal of economic behavior & organization*, 48(3), 265-280.

Laster D., P. Bennett and I. S. Geoum, 1999. Rational bias in macroeconomic forecasts. *The Quarterly Journal of Economics*, 114(1), 293-318.

Lawrence M., P. Goodwin, M.O. Connor and O. Dilek, 2006. Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting*, 22, 493–518.

Mincer J.A. and V. Zarnowitz, 1969. The evaluation of economic forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance* (pp. 1-46). NBER.

Nordhaus W., 1987. Forecasting efficiency: concepts and applications. *The Review of Economics and Statistics*, 69(4), 667–674. Ottaviani, M., & Sørensen, P. (2006). The strategy of professional forecasting. *Journal of Financial Economics*, 81(2), 441-466.

Ottaviani, M. and P.N. Sørensen, 2006. The strategy of professional forecasting. *Journal of Financial Economics*, 81(2), 441-466.

Patton A.J. and A. Timmermann, 2003. Properties of optimal forecasts. Centre for Economic Policy Research Discussion Papers 4037.

Patton A.J. and A. Timmermann, 2010. Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics*, 57(7), 803–820.

Shiller R.J., 1999. Human behaviour and the efficiency of the financial system. *Handbook of macroeconomics*, 1, 1305-1340.

Stekler H., 2007. The future of macroeconomic forecasting: Understanding the forecasting process. *International Journal of Forecasting*, 1–21.

Zarnowitz V., 1984. The accuracy of individual and group forecasts from business outlook surveys. *Journal of forecasting*, 3(1), 11-26.

Zarnowitz V., 1985. Rational expectations and macroeconomic forecasts. *Journal of Business & Economic Statistics*, 3(4), 293–311.