

ORGANISATION DE COOPÉRATION ET DE DÉVELOPPEMENT ÉCONOMIQUES



ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

OECD WORKSHOP ON BUSINESS AND CONSUMER TENDENCY SURVEYS

**Long-Run Biases in Consumer Sentiment.
Micro Evidence from European Surveys**

BY MAURIZIO BOVI

ISAE - Institute for Studies and Economic Analyses

Department of Macroeconomics

E-mail: m.bovi@isae.it

Rome, September 2006

Abstract

Using micro data from the Business Surveys Unit of the European Commission, I perform a long-run analysis of people's perception on the economic situation. Side by side with "logically expected" results, which I interpret as sustaining the reliability of the surveys, data also show "puzzling" outcomes. *E.g.*, Europeans tend to judge over-pessimistically and to forecast over-optimistically. Cognitive psychology suggests that these emotionally-driven asymmetries could be "expected" as well. Thus, on the one hand, they do not affect the trustworthiness of survey data and, on the other hand, they recommend enriching the analysis of consumer sentiment indexes with psychological considerations.

JEL Codes: C42, C82, D12, D84.

Keywords: Consumer attitudes and behavior; consumer confidence, survey research, cognitive economics.

1. INTRODUCTION

In many countries consumer sentiment indexes (CSI) play a relevant role in public discussions about the economic situation. A striking feature of CSI is that they are so commonly diffused and commented at their “face value”, that it seems that what is behind them be considered as known as not to require explanation. Possibly, this is so because mainstream literature (among others, Ludvigson, 2004) seems to support the existence of links between CSI and National Account data. Similarly, perhaps as a consequence of the everyday practice attitude, the economic literature often takes CSI as a given starting point paying little, if any, attention to its components¹ and to how they are assembled. An indirect confirmation of the issues surrounding this topic can be drawn from the fact that virtually identically targeted CSI are based on different approaches. In the US, some of the single questions pertaining to the (University of Michigan) overall CSI have three possible answers referring to "good times", "no change" and "bad times". The weights of the answers are respectively 1, 0 and -1. In Europe there are always five (similar) response options weighted, from “a lot better” to “a lot worse”, 1, 1/2, 0, -1/2, -1. In Japan the five-responses quantification method is: 1, 3/4, 2/4, 1/4 and 0. Thus, somewhat curiously, while in Japan pessimists are left out by CSI, elsewhere the zero weight is assigned to “no change” persons. Also, as mentioned, in the US some questions do not allow extreme positions. Then, the attitudinal research (Dominitz and Manski, 2003, 1999; Das, Dominitz, and van Soest; 1999; Manski, 2004) has made clear why the analysis of the micro foundations of CSI may be important. One obvious problem is that some of the events about which respondents are queried are remarkably vague, *e.g.* “general economic conditions”, and it may be that different respondents do not interpret the same question in the same way. Thus, responses could be not comparable across individuals. Dominitz and Manski (2004) conclude that indexes based on disparate and non-commensurate elements are not the best way to decipher information on consumer beliefs. Other, actually sparse, works (Oppenlander and Poser, 1986; European Commission 1997; Dominitz, 1998) have addressed the additional potential

¹ Among the few papers working with micro-level data are Souleles, 2004; Alessie and Lusardi, 1997.

criticism that there may be little incentive for respondents to reply truthfully². As a consequence, the study of what is behind CSI seems to be very important from both the theoretical and operational standpoints.

Against this backdrop my aim is to shed some light on the trustworthiness of respondents' replies and on some issues related to the quantification method underlying CSI. To this end, unlike previous works, I extract some structural (long run) information contained in people's response options, with an attempt to explain puzzling behaviors via cognitive psychology.

Data are from the Business Surveys Unit of the European Commission. In particular, I deal with the percentage of respondents replying "better", "same", etc., to the queries about economic conditions. From the empirical point of view, I analyze exclusively individual data – *i.e.* I do not associate survey data with "corresponding" National Account items. Moreover, I compute full-sample descriptive statistics for fifteen European countries. Examining i) long run facts (monthly data span twenty years) via ii) basic statistics about iii) the representative individual within iv) the survey framework should reduce some data issues (seasonality, heterogeneity, vagueness, the lack of re-interviews), allowing robust findings.

Results show both predictable (*e.g.*, people know the past better than the future) and puzzling (*e.g.*, people tend to judge over-pessimistically and to forecast over-optimistically) perceptions about the economic situation. I interpret the former as supporting the reliability of the survey data. I argue that the paradoxical outcomes do not perturb the picture. This is so because the detected biases may be validly addressed via cognitive psychology – why should economically astonishing but psychologically explainable outcomes affect the trustworthiness of survey data? Clearly, to the extent one believes in the long-lasting presence of emotionally-driven asymmetries, a psychological dimension should be added to the technical design of the CSI quantification scheme.

The paper is organized as follow. The next section deals with the data, section 3 focuses on statistical analyses and results. Section 4 offers a tentative arrangement of the most puzzling findings. Concluding remarks close the paper.

² According to Smith (1976), monetary incentives for respondents are essential in behavioral economics experiments.

2. DATA

The data are drawn from the Business Surveys Unit of the European Commission. They are based on monthly surveys carried out at a national level by public and private institutes in the framework of the Joint Harmonised European Union Programme of Business and Consumer Surveys³ (European Commission, 1997). In order to achieve representativeness, the bigger member countries use a larger sample. The starting date is January 1985 for nine out of fifteen European Union (EU) countries. Exemptions are Austria (starting date 1995:10), Finland (1987:11), Luxembourg (2002:01), Portugal (1986:06), Spain (1986:06), and Sweden (1995:10). The sample stops in July 2005 for all countries. Currently, almost 33,000 consumers are surveyed each month across the EU. Persons are usually selected by a random stratified sampling procedure or by simple random sampling. At the moment, the most widespread method is the telephone interview. Participants in the survey are asked the following questions, which are harmonized in all countries according to the EU guidelines:

Ex post questions:

Q1 How has the financial situation of your household changed over the last 12 months? It has ...

Q3 How do you think the general economic situation in the country has changed over the past 12 months? It has ...

- PP) got a lot better
- P) got a little better
- E) stayed the same
- M) got a little worse
- MM) got a lot worse
- N) don't know.

Ex ante questions:

Q2 How do you expect the financial position of your household to change over the next 12 months? It will ...

Q4 How do you expect the general economic situation in the country to develop over the next 12 months? It will ...

- PP) get a lot better
- P) get a little better

³ Detailed information on the Joint Harmonised EU Programme of Business and Consumer Surveys can be found in European Commission (1997, 2004).

- E) stay the same
- M) get a little worse
- MM) get a lot worse
- N) don't know.

In fact, national surveys contain other ex ante and ex post questions about the labor market, spending intentions on major purchases (furniture, electrical/electronic devices, etc.), savings, etc. While each question has a potential information content, I focus only on the mentioned four queries. Hopefully, they should constitute a sufficient information set in the present context. Let PP, P, E, etc. denote the percentages respondents having chosen the corresponding option, so that $PP+P+E+M+MM+N=100$. These percentages are the data I deal with while, as mentioned, the most common way of presenting/commenting/analyzing consumers survey data is the balance (Anderson, 1952), that is the difference between positive and negative percentages. Balances are calculated as $B = (PP + \frac{1}{2}P) - (\frac{1}{2}M + MM)$. The aggregate index is then calculated as a simple average of individual indicators. The data set suffer from some change throughout the sample. Since 1995, for instance, Italy substituted on-the-spot interviews with the telephone method. In Germany, apart from the issues stemming from the re-unification of 1991, there have been some modifications in the order as well as in the wording of some questions. I was not been able to find data issues for other countries, but a first impression of them can be drawn by the graphs reported in Appendix 2 (*e.g.*, for Portugal data from February 1997 to August 1997 are not available). Altogether it means that there are some problems in the time series comparability of the data. Then, it is easily understood that the queries are remarkably vague and there are not incentives/disincentives related to a particular answer. Finally, persons are usually selected randomly, in that somewhat preventing to perform analyses via re-interviews⁴. On the positive side, my aim is to highlight some basic, structural, facts via full-sample descriptive statistics referring to representative agents. This should reduce data issues such as seasonality, heterogeneity, vagueness and the lack of re-interviews. Otherwise stated, studying basic statistics within the framework of the

⁴ This kind of analysis was suggested by the so called Smithies Committee (1955), as well as by Tobin (1959) and Juster (1964). As mentioned in the main text, this recommendation turned out to be the minority view.

long-running continental-scale harmonized “experiment” managed by EU should allow achieving very robust results.

3. STATISTICAL ANALYSIS AND RESULTS

While agents may act differently from what they say, the message contained in surveys of people in households can be validly and usefully decoded to infer intriguing aspects of consumers’ way of thinking. In the present setting, some quick and simple experiments based on reply options allow verifying some commonsense-predictable scores. For instance, do consumers know their own situation better than the system wide one? If this is the case, the average share of individuals answering “don’t know” to questions about the general environment should be greater than the average share of individuals which do not know how their own situation is going on. A similar trial can be performed by comparing corresponding ex ante vs ex post replies in order to see whether consumers feel, as logically expected, more uncertain about the future. Thus, the simple comparison between usually neglected information can confirm/negate interesting behavioral conjectures.

Tab. 1. Consumers' uncertainty on Personal vs General and Past vs Future economic conditions

	Personal		General	
	(Q1) Past	(Q2) Future	(Q3) Past	(Q4) Future
AUSTRIA	0.98	2.97	2.43	4.04
BELGIUM	2.96	6.69	6.05	10.8
GERMANY	1.36	4.73	2.32	5.58
DENMARK	0.66	3.70	7.11	8.85
GREECE	0.16	4.46	1.95	7.68
SPAIN	1.09	9.68	5.37	14.8
FINLAND	0.60	3.59	3.34	4.89
FRANCE	0.48	4.23	1.66	8.76
IRELAND	0.94	4.88	1.97	6.97
ITALY	0.50	4.68	1.95	6.32
LUXEMBOURG	1.16	3.46	4.67	5.59
NETHERLANDS	1.08	4.50	7.00	11.0
SWEDEN	0.77	2.21	5.49	4.77
PORTUGAL	1.33	11.2	5.47	15.7
UK	1.57	5.93	4.62	9.84
EU11	1.02	5.08	3.93	8.58

EU11=Belgium, Germany, Denmark, Greece, Spain, Finland, France, Ireland, Italy, Netherlands, UK (sample 87:11-05:07)

Full sample average of responses "don't know" (in % of total) to the questions:

Q1=How has the financial situation of your household changed over the last 12 months?

Q2=How do you expect the financial position of your household to change over the next 12 months?

Q3=How do you think the general economic situation in the country has changed over the past 12 months?

Q4=How do you expect the general economic situation in the country to develop over the next 12 months?

The picture emerging from table 1 is plain. Europeans do report a greater uncertainty in addressing both system wide conditions, as opposed to familiar ones, and future developments as opposed to past situations. Another intriguing experiment deals with the "E" answer. Since the queries are about "developments/changes", individuals should respond, on average, "the same" the most part of times (Theil, 1961), because it is hard to think to ever improving/worsening economic conditions (whatever it means for common people⁵) over many years. In addition, the preference of being "E" may be partly due to the fact that this "neutral" option may be chosen by uninformed and/or uninterested respondents. In appendix 1, I report the empirical distributions of the full sample means of the six response options. The visual impact of the histograms clearly shows that individuals reply, on average, "the same" the most part of times. It is worth recalling that in the European quantification method, both "N" and "E" respondents do not affect the overall CSI.

⁵ To the extent i) GDP growth coincides with people's view of "development in economic condition", and ii) GDP growth follows a stationary process agents should, on average, accumulate towards the "stationary" item of the questionnaire. It is worth noticing that the average GDP (Consumption) growth has been positive for each and every country during the years under scrutiny (the same holds in per capita terms), whereas Europeans seem to be more pessimists than optimists.

Given their “logically expected” outcomes, all the tests performed so far may be thought of as supporting the reliability of the survey “experiment”. In fact, the distributions proposed in Appendix 1 shed some light on other interesting long run features of consumers’ replies. The number of agents responding “the same” when elicited about personal as opposed to general economic developments is much higher. Perceptions about the personal context show a unimodal distribution in E, with a very high percentage of E. On the contrary, beliefs on the general environment display ten (out of thirty) M-peaks with an almost halved E. This calls for *ad hoc* experiments to contrast general vs personal and ex-ante vs ex-post response options. One simple way to address the former issue is computing mean values of $(Q1+Q2)-(Q3+Q4)$ for each single item (leaving aside the already studied E and N). The term $(Q1+Q2)$ refers to the two personal queries, the second to the pair of general questions (see section 2). Thus, negative (positive) values in columns “MM” and “M” (“PP” and “P”) of table 2 imply that the personal condition is perceived to be systematically better than the general one. Table 2 collects the results, where bold values highlight a personal condition perceived to be worse than the general one. Since such values amount to a low 28% (seventeen out of sixty experiments), there is a strong clue that the representative agents’ sentiment about their own economic condition is consistently better than the general one. Moreover, as already noticed, the response option “the same” shows a much lower share in queries eliciting general economic conditions. As a consequence, in passing from Q1,Q2 to Q3,Q4 questions, some “E” individual responds differently. Table 2 leads to conclude that most part of them becomes pessimist. By the comparison between the figures in column M and in column P it results that pessimists are more numerous than optimists in twelve out of fifteen experiments. All in all, bold values are a small minority and have smaller values than the others. It turns out that the personal condition is considered to be systematically better than the general one.

Tab. 2. Comparing Personal vs General Sentiment in fifteen European Countries

	PP	P	M	MM
AUSTRIA	1.0	-19.0	-25.0	-8.1
BELGIUM	1.0	-8.9	-26.0	-19.0
GERMANY	0.6	-5.6	-23.0	-12.0
DENMARK	10.0	-1.0	-17.0	-3.4
GREECE	-0.7	-6.2	-5.3	-5.5
SPAIN	-0.3	-9.6	-16.0	-9.3
FINLAND	3.8	-19.0	-14.0	-3.9
FRANCE	2.2	1.5	-31.0	-23.0
IRELAND	-3.8	-16.0	-10.0	-13.0
ITALY	-1.8	-15.0	-22.0	-26.0
LUXEMBOURG	2.8	1.7	-50.0	-8.0
NETHERLANDS	5.1	-12.0	-15.0	-9.9
SWEDEN	-0.4	-4.6	-22.0	-8.0
PORTUGAL	6.6	-11.0	-21.0	-2.9
UK	5.3	-2.7	-14.0	-14.0
EU11	2.1	-8.1	-18.6	-12.7

EU11=Belgium, Germany, Denmark, Greece, Spain, Finland, France, Ireland, Italy, Netherlands, UK (sample 87:11-05:07).

Country-rows report mean values of $(iQ1+iQ2)-(iQ3+iQ4)$ where $i=PP,P,M,MM$. PP=a lot better; P=a little better; M=a little worse; MM=a lot worse. Q1-Q4 see under table 1. E.g., when $i=MM$, $iQ1=$ % of agents responding "my financial situation changed a lot worse over the last 12 months". Clearly, negative (positive) values in columns "MM" and "M" ("PP" and "P") mean that the personal condition (Q1,Q2) is perceived to be systematically better than the general one (Q3,Q4). Bold values show the opposite.

Alike, table 3 gathers mean values of $(Q1+Q3)-(Q2+Q4)$ for each single response option (again excluding E and N). The first term, $(Q1+Q3)$, refers to the ex-post queries, the second to the ex-ante questions. In this case, positive (negative) values in columns "MM" and "M" ("PP" and "P") imply that judgments $(Q1,Q3)$ are systematically worse than forecasts $(Q2,Q4)$.

Tab. 3. Comparing Judgments vs Forecasts in fifteen European Countries

	PP	P	M	MM
AUSTRIA	-0.8	-13.0	14.0	12.0
BELGIUM	0.1	-7.4	14.0	11.0
GERMANY	1.2	1.0	12.0	9.4
DENMARK	1.3	-3.0	3.8	3.6
GREECE	-0.9	-12.0	27.0	5.8
SPAIN	0.8	-7.9	16.0	8.6
FINLAND	0.8	-8.1	6.1	7.3
FRANCE	-0.8	-11.0	12.0	15.0
IRELAND	3.7	-2.3	13.0	14.0
ITALY	-0.9	-15.0	23.0	17.0
LUXEMBOURG	-0.3	-15.0	18.0	4.2
NETHERLANDS	5.6	-1.5	-0.6	10.0
SWEDEN	0.3	-5.0	12.0	4.9
PORTUGAL	0.7	-9.0	5.9	5.2
UK	0.6	-9.9	12.0	15.0
EU11	1.0	-7.2	12.7	10.2

EU11=Belgium, Germany, Denmark, Greece, Spain, Finland, France, Ireland, Italy, Netherlands, UK (sample 87:11-05:07).

Country-rows report mean values of $(iQ1+iQ3)-(iQ2+iQ4)$. PP, P, M, MM and Q1-Q4 see under table 1. When $i=MM$, $iQ1=$ % of agents responding "my financial situation changed a lot worse over the last 12 months". Clearly, positive (negative) values in the columns "MM" and "M" ("PP" and "P") mean that judgments $(Q1,Q3)$ are systematically worse than forecasts $(Q2,Q4)$. Bold values show the opposite.

The picture arising from table 3 suggests that, when consulted about the economy, people's judgments are worse than people's forecasts even considering hundreds of tests performed across several countries. The detected difference between ex ante and ex post perceptions recommends refining the experiment. As a matter of fact, according to one of the basic axiom of the standard neoclassical models, agents should not consistently repeat the same mistakes. In the present framework, it may be addressed by looking at the gap between "contemporaneous" ex ante and ex post responses, to which I refer as the "forecast error". An example may help to clear the matter. Let the share of individuals forecasting that the system wide economic situation will be "a little worse" in the next year be, according to the survey performed in January 2000, 35%. After a year, interviewed are asked to say how the general economic situation in the country has changed over the past 12 months. If people's forecasts in January 2000 were corrected, then the share of individuals judging that the economic situation has got "a little worse" should be 35%. It is noteworthy that, in this setting, there is no need for agents to correctly address what a "general economic situation" really is. In fact, I just compare answers given to the same question. With the potential exemption of "don't know", which is a "non" response⁶ (*i.e.* it is not the outcome of an explicit elaboration but, rather, a declaration of no information), the equivalence should hold for each and every possible ex ante vs ex post same-period-referring pairs. Needless to say the bias should be zero only on average, allowing for short-living forecast errors (this is why it is usually called the "long-run bias"), perhaps partly due to the different individuals interviewed, too. In appendix 2, I plot the forecast errors and I report some descriptive statistics. Forecast errors sustain the previous conclusion. People show an evident asymmetry towards the future, keeping on thinking that things, as compared to what they themselves think it is happened, will improve. E-agents turn out to be the most "rational" (or, *rectius*, the less emotional) – the values of their forecast errors are the lowest. Anyway, as mentioned, this ex-ante-ex-post coherence could be partly due to the appealing of keeping neutral. Clearly, psychological neutrality is different from analytical rationality. Then, reflecting agents' difficulties in addressing more complex

⁶ On that, European Commission Users' Manual (1997, p. 18) claims that: "(...) there are six reply options: five "real" ones and a "do not know" option."

tasks (Thaler, 1999), opinions about the system wide situation should show a greater volatility. In order to save space I do not report single item standard errors⁷, although an indirect clue may be found by looking at the standard errors of personal vs general forecasts errors (Appendix 2).

Somewhat confirming earlier findings on consumers inflation expectations (Souleles 2004; Forsells and Kenny, 2004; Carrol, 2001, 2003), the picture emerging from the proposed empirical analysis leads to conclude that people behave sometimes as expected, sometimes not. Among predictable behaviors it results that, over time, across countries, and for the most part of the response options:

1. agents think to know their own situation better than the system wide one, and the past better than the future;
2. opinions about the system wide situation show a greater volatility as compared to personal situation replies;
3. responses accumulate towards a “long-run stationarity” of the economic stance.

Side by side with these conventional scores, somewhat supporting the reliability of survey experiments, some paradoxical outcomes emerge as well. They refer to:

1. people’s tendency to judge over-pessimistically and/or to forecast over-optimistically.

The ambiguity arises because of the lack of a “hard” benchmark (*e.g.*, GDP, Consumption, etc.). However, as suggested in the next section, there are reasons to believe that people systematically both judge over-pessimistically and forecast over-optimistically. Anyway, it implies that:

2. people’s forecasts show a long run bias. Last, but not least:
3. people think that their own situation can be systematically different from the general one.

To sum up the puzzling results, it seems that there is a “structural” mantra echoing across Europe. It sounds as follows:

⁷ Available on request from author. They support agents’ difficulties in addressing complex tasks.

AS USUAL, IT HAS GOT WORSE THAN I EXPECTED. ESPECIALLY FOR THE OTHERS.

NEVERTHELESS, I THINK THAT IT WILL GET BETTER. ESPECIALLY FOR ME.

4. EXPLAINING THE PARADOXES

Some undisputed findings in cognitive psychology are validly exploitable to address the over-pessimism in judgments and the over-optimism in people's forecasts (Kahneman and Tversky, 1973, 1974, 1982). Psychologists suggest that, compared to unfamiliar information, familiar information is more easily accessible from memory and is believed to be more real or relevant. Therefore, mere repetition of certain information in the media, regardless of its accuracy, makes it more easily available and therefore falsely perceived as more accurate. The explanation is completed by noticing that, according to Doms and Morin (2004), the media tend to overweight bad economic news⁸. In fact, that is just the very basic nature of the news media. So, there are reasons inducing individuals toward dispositional pessimism. What psychology may say about people's tendency to forecast over-optimistically? According to the psychological law of small numbers, as opposed to (*rectius*, side-by-side with) the statistical law of large numbers, people believe that the mean value from a small sample also has a distribution concentrated at the expected value of the random variable. This leads to a bias due to "overinference" from (too) short sequences of observations. In an overview of behavioral finance, Shleifer (2000) argues that the law of small numbers may explain the excess sensitivity of stock prices (Shiller, 1981) as a result of investors' overreacting to short strings of good news. Likewise, as suggested by Shiller (2000), another aspect of overconfidence (irrational exuberance, in the famous Mr. Greenspan's 1996 speech) is that people tend to make forecasts in uncertain situations by looking for familiar patterns and assuming that future favorable patterns will resemble past ones, often without sufficient consideration of the reasons for the pattern or the probability of the pattern repeating itself. Finally, we may resort to the so called hindsight bias (Shiller, 2000). Suppose there is an unexpected event. People tend to concoct explanations for it after the fact, which

⁸ Doms and Morin (2004) find that news affect consumers' sentiment about general situations more deeply than that on personal conditions. The approach of this paper may help in explaining their result.

makes them appear more predictable, and less random, than it is. More generally speaking, our minds are designed to retain, for efficient storage, past information that fits into a compressed narrative. This distortion prevents agents from adequately learning from the past. The point I want to stress is that these departures from mathematical/rational expectations may help in understanding the presence of the long run bias which, it is worth recalling, is computed “as if” people act/think in an unemotional Muthian way.

Let us now turn the attention to the other stylized fact emphasized in the previous section. Why people believe that their own situation can be consistently better than the general one? To answer, the phenomenon of the illusion of control is validly exploitable (DeBondt and Thaler, 1995). It belongs to the more general class of egocentric biases, among overconfidence and unrealistic optimism (Msetfi *et al.*, 2006 for a survey). It is magnified by skill-related factors, in the present case the greater familiarity with the personal situation, and it is defined as “an expectancy of a personal success probability inappropriately higher than the objective probability would warrant” (Langer 1975, p. 313). Closely related to the illusion to control, there is the theory of depressive realism. In a seminal paper, Alloy and Abramson (1979) found that non-depressed people are more likely than depressed people to think that outcomes are contingent on their actions when they are not. They concluded that as opposed to depressed people, whose perceptions are apparently accurate, common people distort reality in an optimistic fashion. One interpretation of depressive realism is that non-depressed people possess a positive bias, which allows them to feel in control of their environment. Since, hopefully, the representative European is non-depressed, evidence supports agents’ tendency to think that it is systematically less likely that they themselves will suffer an adverse event than the average agent.

CONCLUDING REMARKS

Everyday practice presents survey data on consumer sentiment by means of aggregate measures. Likewise, usually, mainstream literature analyzes summary information from the surveys in the form of confidence indexes. Otherwise stated, CSI are so commonly diffused and commented at their face value, that it seems that what is behind them be considered as known as not to require explanation. While it is easy to

understand the factual need to report easily accessible indicators, aggregation/quantification procedures necessarily imply loss of information (not mentioning the issues/assumptions they involve).

This paper deals with micro data to shed some light on what is behind aggregate CSI, with some attempt to explain “puzzling” results. The focus is on the long-run information content of people’s perceptions on economic conditions. The aim is to shed some light on the trustworthiness of respondents’ replies and on some potential issue related to the quantification method underlying CSI. Examining i) long run facts (monthly data span twenty years) via ii) basic statistics about iii) the representative individual within iv) the survey framework should reduce some data issues (seasonality, heterogeneity, vagueness, the lack of re-interviews), allowing robust findings.

Results suggest that sometimes survey respondents reply, on average, as “logically” expected. Agents think to know their own situation better than the system wide one, and the past better than the future. Also, perceptions accumulate towards the “stationarity” of the economic stance. Altogether these commonsense outcomes may be thought of as supporting the hypothesis that survey data give a faithful representation of people’s opinions. This is not the only story told by the data, which show the presence of a structural and puzzling “mantra” echoing across EU. It is worth noticing that the experiments highlight the emergence of paradoxical responses even when dealing with familiar conditions. Thus, it is not only a problem of the amount and quality of available information and/or the difficulty of the exercise - there is something else preventing “rational” results. Offering an explanation of these biases via cognitive psychology this paper argues that i) the puzzling outcomes not necessarily reduce the reliability of survey data and that ii) adding psychological considerations to the technical design of CSI could be of some help. After all, everybody should agree that the sentiment stems from the psyche.

REFERENCES

- Alloy, L.B., and L.Y. Abramson (1979): "Judgment of contingency in depressed and nondepressed students. Sadder but wiser?" *Journal of Experimental Psychology*, 108, 441-485.
- Anderson, O. (1952), The Business Test of the IFO Institute for Economic Research, *Review of the International Statistical Institute*, Vol. 20, 1-17
- Blinder, A.S., and A.B. Krueger (2004): "What Does the Public Know about Economic Policy, and How Does It Know It?" *NBER WP*, 10787.
- Bovi, M. (2005): "Consumer Surveys and Reality", *Statistics in Transition*, Central Statistical Office, Poland. Special Issue, 7, 239-258.
- Carroll, C. (2003): "Macroeconomic Expectations of Households and Professional Forecasters," *Quarterly Journal of Economics*, 118, 269-298.
- Carroll, C. (2001): "The Epidemiology of Macroeconomic Expectations," *NBER Working Paper*, 8695.
- Cyert, R.M., H.A. Simon, and Trow, D.B. (1956): "Observation of a business decision," *Journal of Business*, 29, 237-248.
- Das, M., J. Dominitz, and A. van Soest (1999): "Comparing Predictions and Outcomes: Theory and Application to Income Changes," *Journal of the American Statistical Association*, 94, 75-85.
- DeBondt, W., and R. Thaler (1995): "Financial decision-making in markets and firms: a behavioral perspective," in *Handbooks in OR and MS*, ed. by Jarrow, R. V. Maksimovic and W. Ziemba.
- Dominitz, J. (1998): "Earnings expectations, revisions and realizations," *Review of Economics and Statistics*, 80, 374-388
- Dominitz, J., and C. Manski (1999): "The Several Cultures of Research on Subjective Expectations," in: *Wealth, Work, and Health: Essays in Honor of F. Thomas Juster*, ed. by Smith, J. and R. Willis University of Michigan Press, Ann Arbor.
- Dominitz, J., and C. Manski (2003): "How Should We Measure Consumer Confidence (Sentiment)? Evidence from the Michigan Survey of Consumers," *NBER Working Paper*, 9926.
- Dominitz, J., and C.F. Manski (2004): "How Should We Measure Consumer Confidence?," *Journal of Economic Perspectives*, 18, 51-66.
- Doms, M., and N. Morin (2004): "Consumer Sentiment, the Economy, and the News Media", *Finance and Economics Discussion Series*, Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, 51, Washington, D.C.
- Driver, C., and G. Urga (2004): "Transforming Qualitative Survey Data: Performance Comparisons for the UK," *Oxford Bulletin of Economics & Statistics*, 66, 71-89.
- EUROPEAN COMMISSION (1997): The Joint Harmonised EU Programme of Business and Consumer Surveys. User Guide, 6, D.G. ECONOMIC AND FINANCIAL AFFAIRS
- EUROPEAN COMMISSION (2004): The Joint Harmonised EU Programme of Business and Consumer Surveys. User Guide (updated 26/05/2004):, D.G. ECONOMIC AND FINANCIAL AFFAIRS

- FEDERAL RESERVE CONSULTANT COMMITTEES ON CONSUMER SURVEY STATISTICS (1955): "Smithies Committee Report in Reports of the Federal Reserve Consultant Committees on Economic Statistics," Hearings of the Subcommittee on Economic Statistics of the Joint Committee on the Economic Report, 84th US Congress.
- Forsells, M., and G. Kenny (2004): "Survey Expectations, Rationality and the Dynamics of Euro Area Inflation," *Journal of Business Cycle Measurement and Analysis*, 1, 13-41.
- Hayek, F.A. (1952), *The Sensory Order. An Inquiry into the Foundations of Theoretical Psychology*, London: Routledge & Kegan Paul.
- Juster, T. (1964): *Anticipations and Purchases. An Analysis of Consumer Behavior*. Princeton: Princeton University Press.
- Kahneman, D. and A. Tversky (1973): "Availability: A heuristic for judging frequency and probability", *Cognitive Psychology*, 5, 207-232.
- Kahneman, D. and A. Tversky (1974): "Judgment under uncertainty: Heuristics and biases", *Science* 185, 1124-1131.
- Kahneman, D. and A. Tversky (1982): "Judgment of and by representativeness", in *Judgment Under Uncertainty: Heuristics and Biases*, ed. by Kahneman, D., P. Slovic, and A. Tversky. Cambridge University Press, Cambridge.
- Katona, G. (1951): *Psychological Analysis of Economics Behaviour*, New York: McGraw-Hill.
- Katona, G. (1957): "Federal Reserve Board Committee Reports on Consumer Expectations and Savings Statistics," *Review of Economics and Statistics*, 39, 40-46.
- Langer, E. (1975): "The Illusion of Control," *Journal of Personality and Social Psychology*, 32, 311-328.
- Ludvigson, S. (2004): "Consumer Confidence and Consumer Spending," *Journal of Economic Perspectives*, 18, 29-50.
- Manski, C.F. (2004): "Measuring Expectations," *Econometrica*, 72, 1329-1376.
- Mitchell, J., and M. Weal (2005): "Quantitative inference from qualitative business survey panel data: a microeconomic approach," *National Institute of Economic and Social Research Discussion Paper*, 261.
- Mitchell, J., R.J. Smith and M. Weal (2005): "Forecasting Manufacturing Output Growth Using Firm-Level Survey Data," *National Institute of Economic and Social Research Discussion Paper*, 251.
- Msetfi, R.M., Murphy, R.A. and Simpson, J. (2006): "Depressive realism and contingency judgement: Further support for the contextual processing hypothesis," *Quarterly Journal of Experimental Psychology*, Forthcoming.
- Mueller, E. (1957): "Effects of Consumer Attitudes on Purchases," *American Economic Review*, 47, 946-965.
- North, D.C. (2000): "Economics and Cognitive Science" in *The MIT Encyclopaedia of Cognitive Science* Cambridge: MIT Press.
- Oppenlander, K. H. and G., Poser (1986): *Business Cycle Surveys in the Assessment of Economic Activity*. Aldershot:Gower.

Shiller, R. (1981): "Do stock prices move too much to be justified by subsequent changes in dividends?", *American Economic Review* 71, 421-436.

Shiller, R. (2000): *Irrational Exuberance*, Princeton: Princeton University Press.

Shleifer, A. (2000): *Inefficient Markets. An Introduction to Behavioral Finance*. Oxford: Oxford University Press.

Simon, H.A. (2001): *Models of my Life*, New York, Wiley.

Smith, V. L. (1976): Experimental economics: Induced value theory. *American Economic Review* 66: 274-279.

Souleles, N. (2004): "Expectations, Heterogeneous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys," *Journal of Money, Credit and Banking*, 36, 39-72.

Stephens Jr., M. (2003): Job Loss Expectations, Realizations, and Household Consumption Behavior, *NBER WP*, 9508.

Theil, H. (1961): *Economic Forecasts and policy*. North-Holland, Amsterdam.

Thaler, R.H. (1999): "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12, 183-206

Tobin, J. (1959): "On the Predictive Value of Consumer Intentions and Attitudes," *Review of Economics and Statistics*, 41, 1-11.

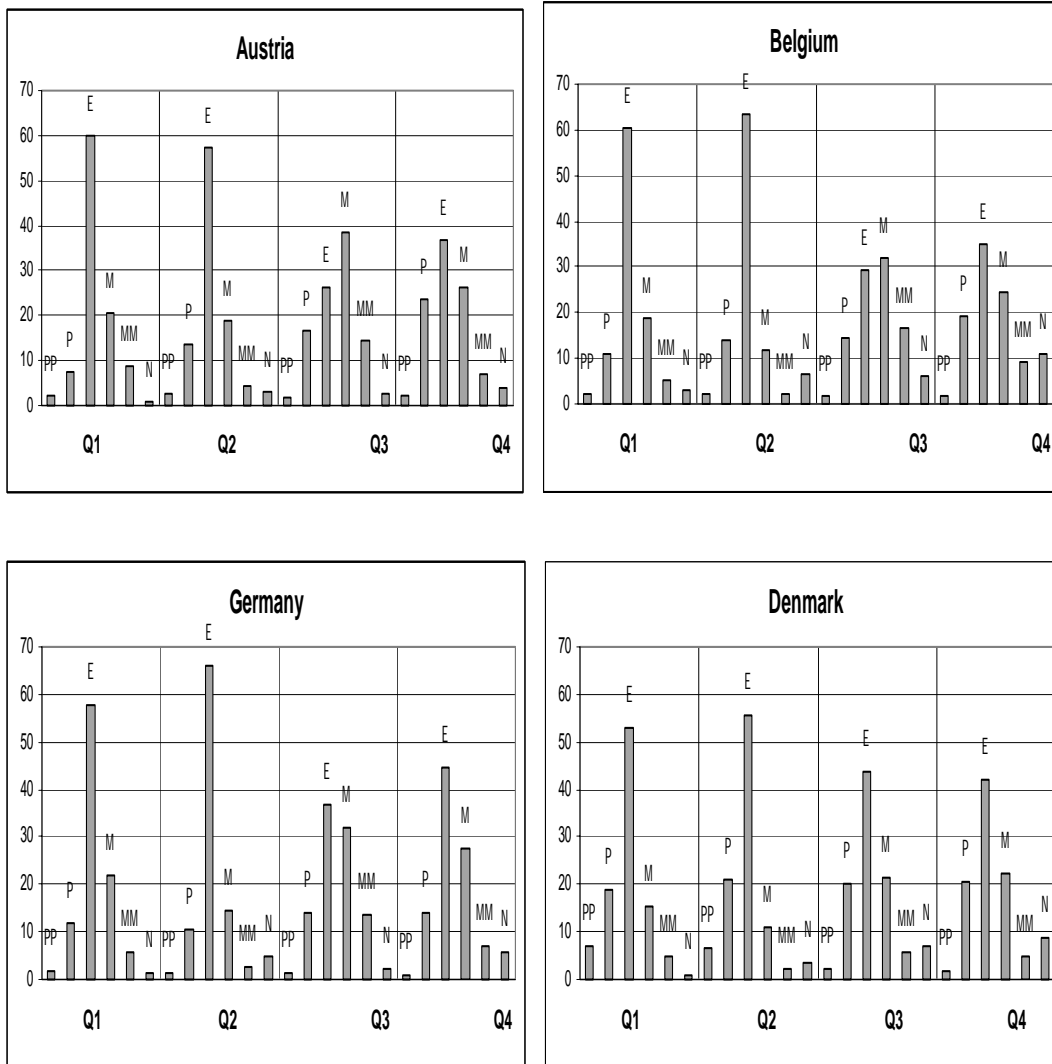
APPENDIX 1. THE DISTRIBUTION OF EUROPEANS' RESPONSES ON ECONOMIC CONDITIONS

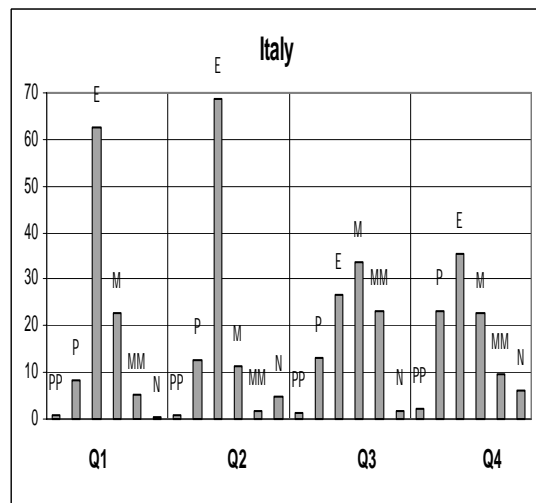
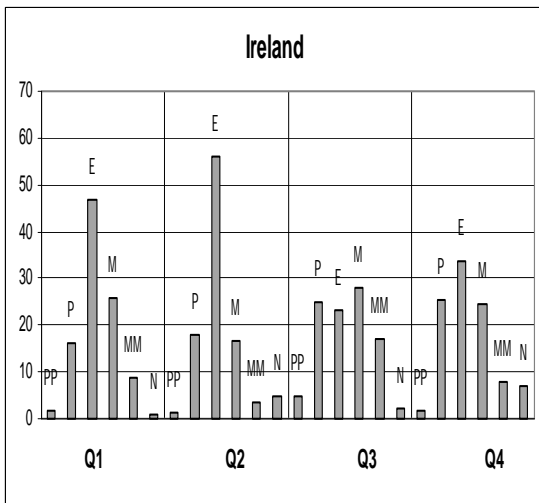
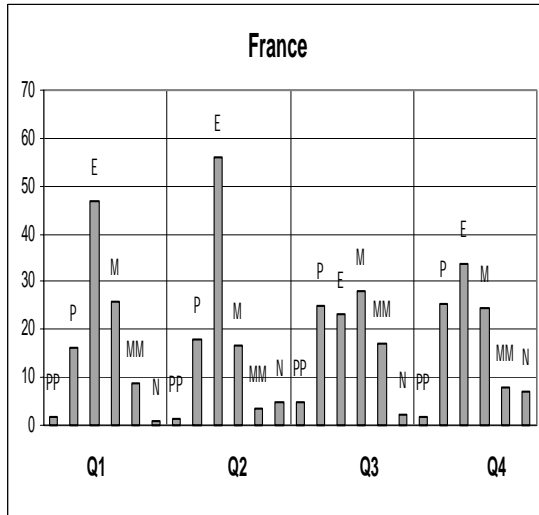
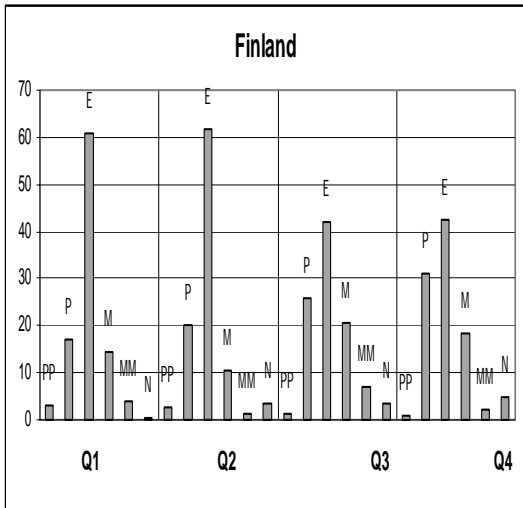
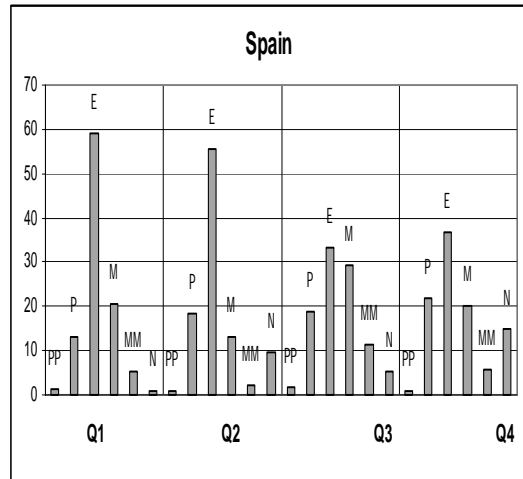
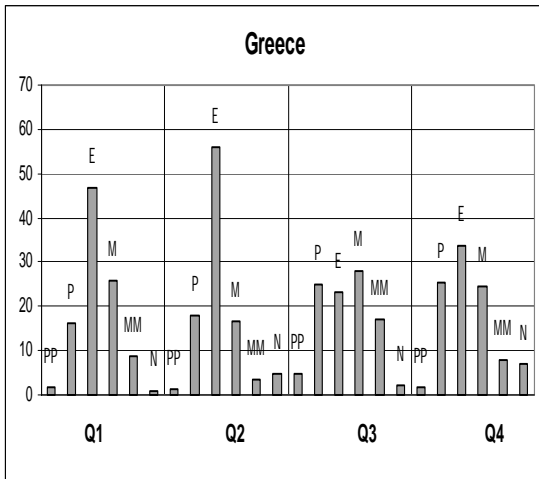
Sample 1985:01–2005:07 for all countries but for Austria (starting date 1995:10), Finland (1987:11), Luxembourg (2002:01), Portugal (1986:06), Spain (1986:06), and Sweden (1995:10).

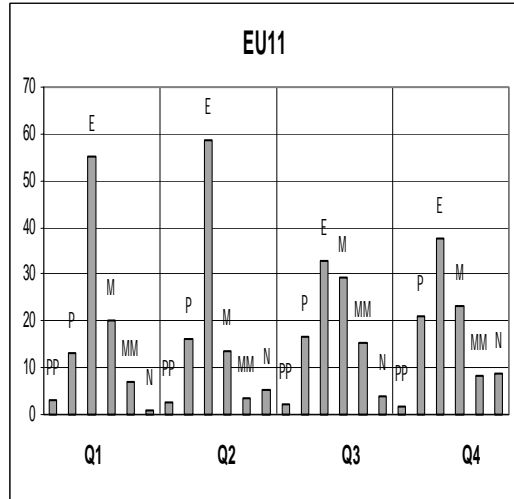
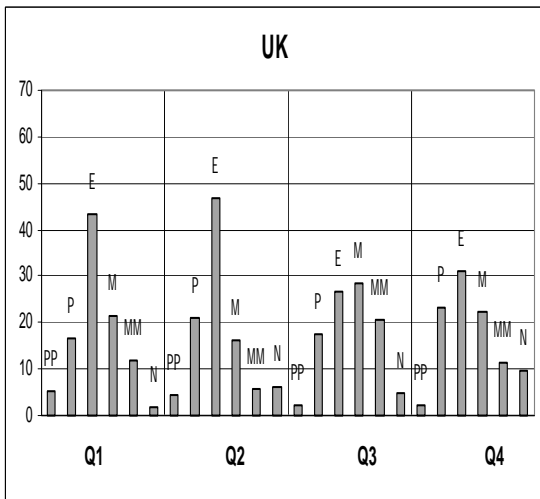
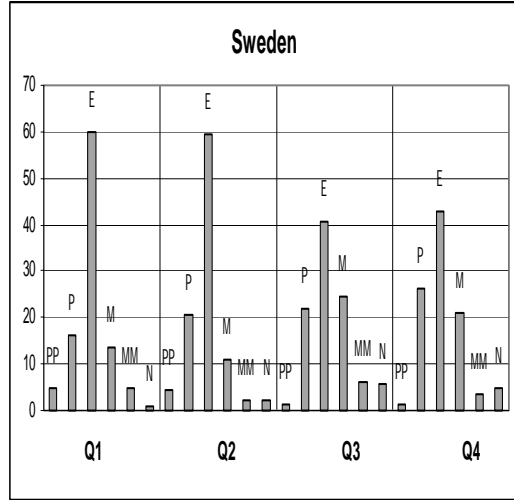
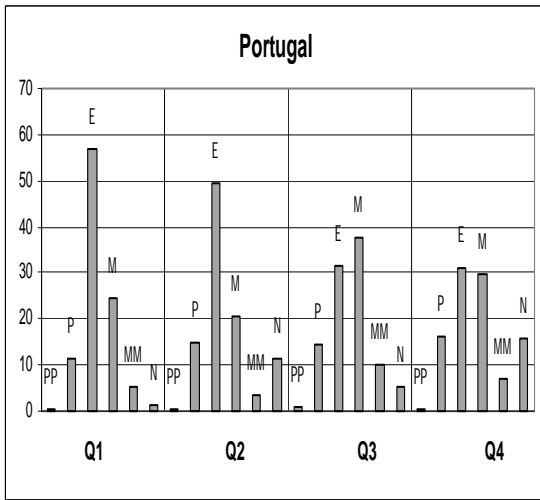
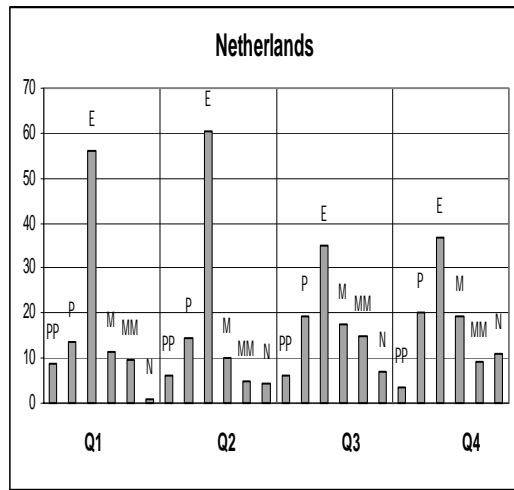
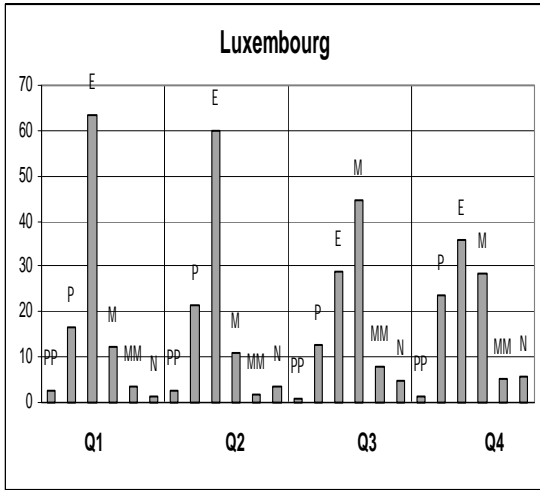
- Q1=How has the financial situation of your household changed over the last 12 months? It has ...
- Q2 How do you expect the financial position of your household to change over the next 12 months? It will ...
- Q3=How do you think the general economic situation in the country has changed over the past 12 months? It has ...
- Q4 How do you expect the general economic situation in the country to develop over the next 12 months? It will ...

PP=got/get a lot better; P=got/get a little better; E=stayed/stay the same; M=got/get a little worse; MM=got/get a lot worse; N=don't know.

Histograms report full (individual) sample means of each item.



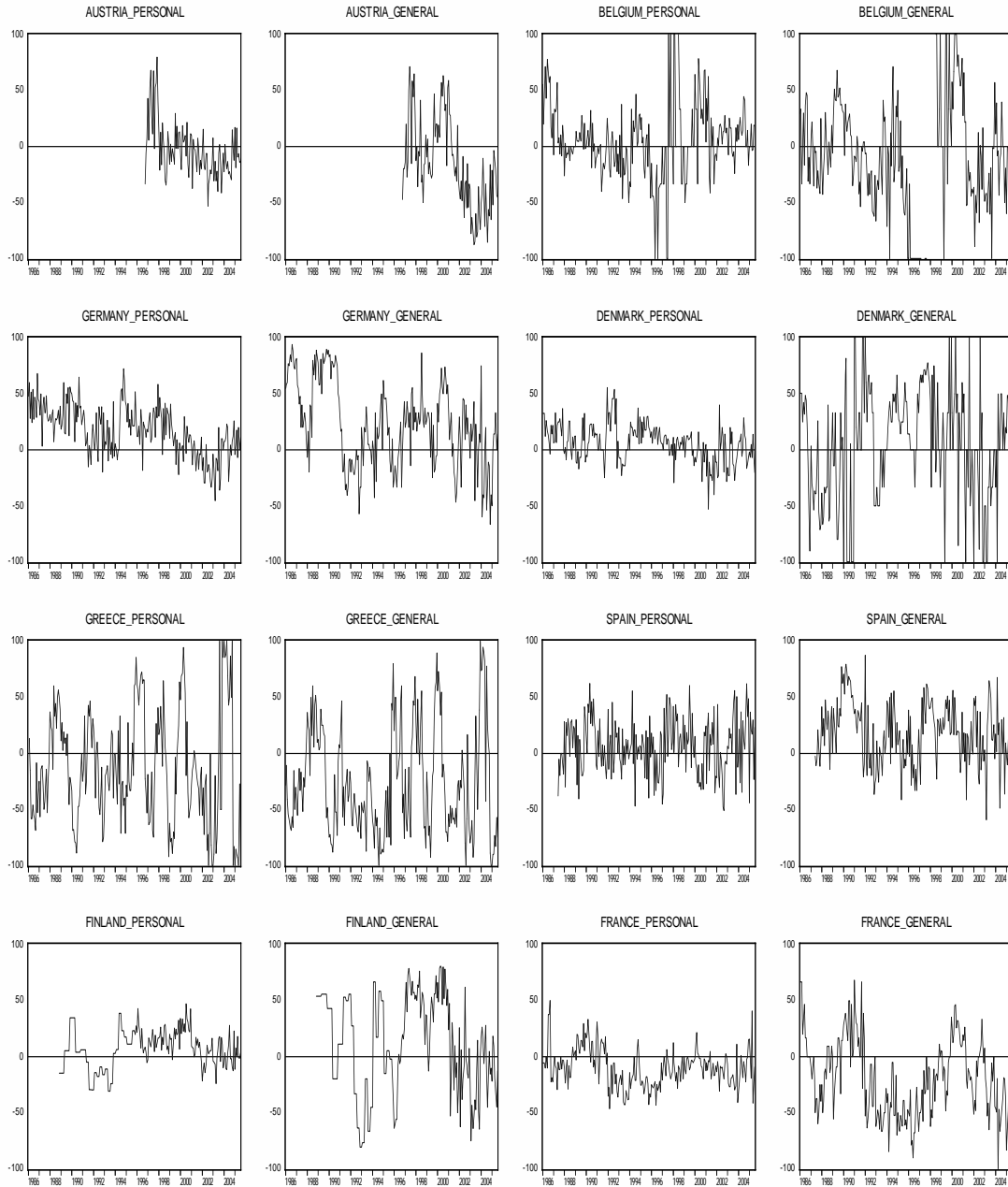




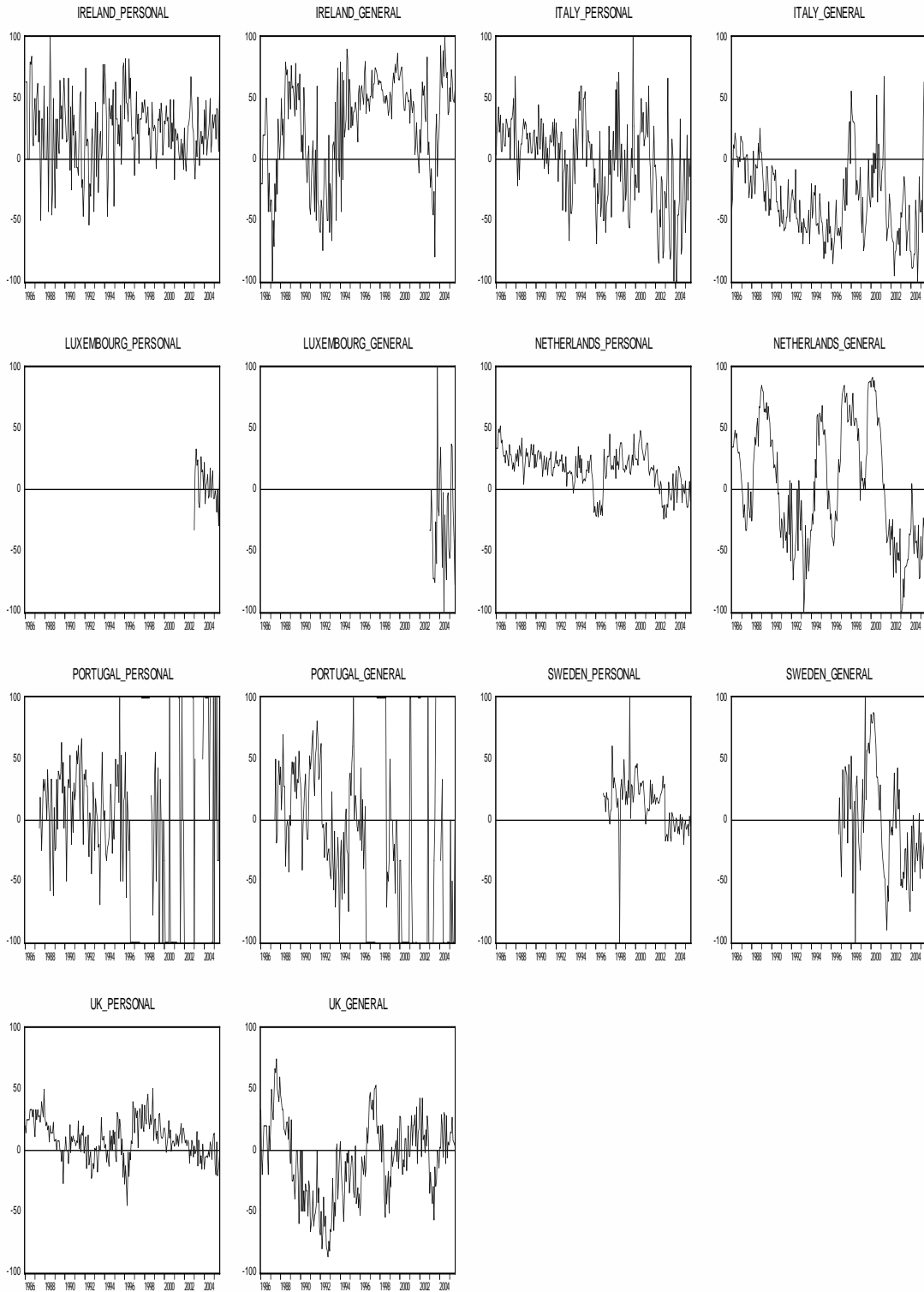
APPENDIX 2. EUROPEANS' FORECASTS ERRORS.

Graphs plot the “forecast error” computed as $100 \cdot [Q1_i - Q2_i(-12)] / [Q1_i + Q2_i(-12)]$ and $100 \cdot [Q3_i - Q4_i(-12)] / [Q3_i + Q4_i(-12)]$ where $i = PP, P, E, M, MM$. The headers (PP, P, etc.) refer to the six possible responses. Other details in Appendix 1. The statistics reported at the end of each page follow the order of graphs. Thus, e.g., the second column refer to Austria_Personal, the last to France_General. The last row (% in $\pm 5\%$ band) report the number of values (as % of total) within the $\pm 5\%$ band, i.e. the share of months with (quasi) no bias. The forecast error of a perfectly rational agent is: mean=0; ± 5 band=100.

PP

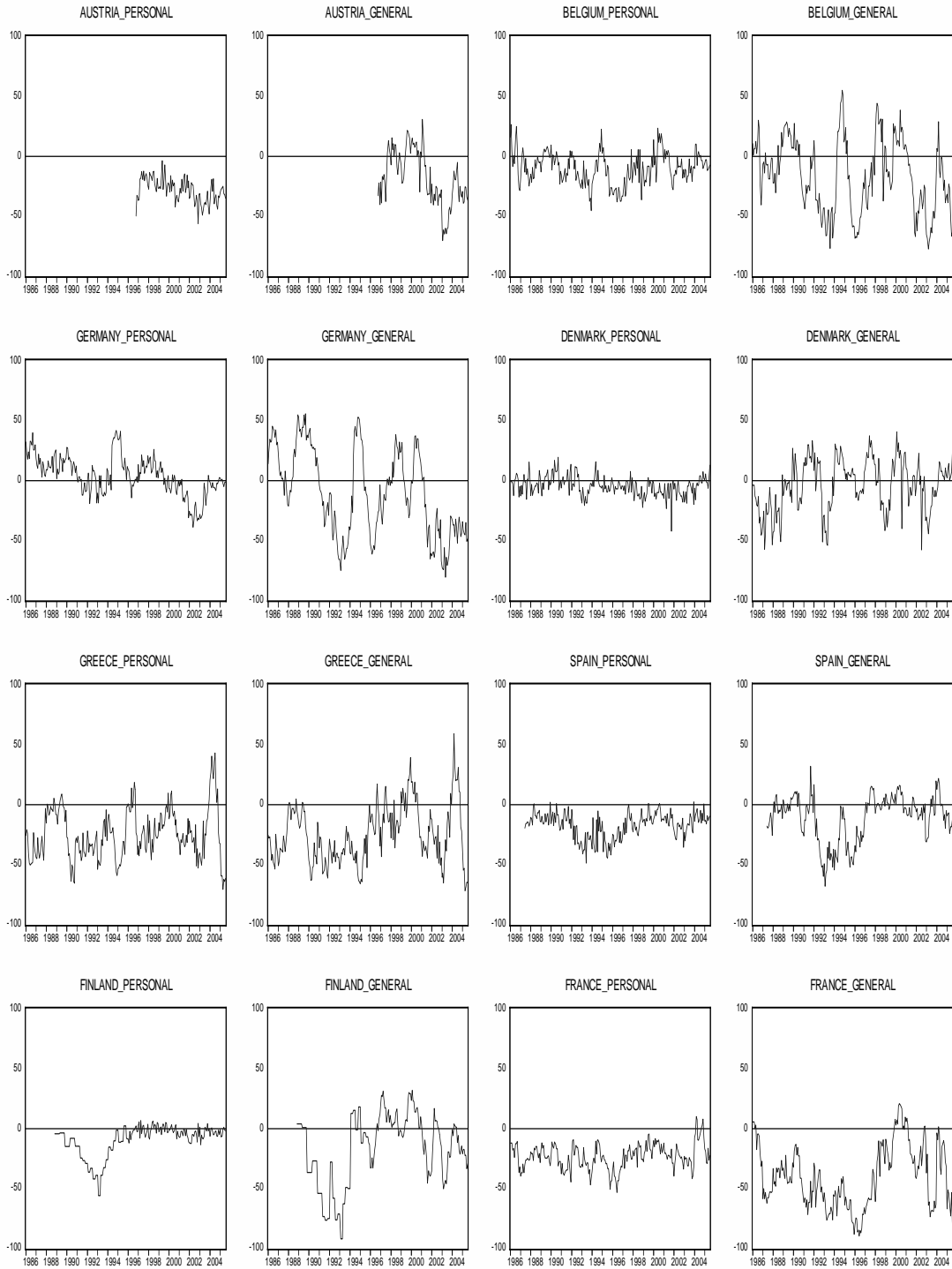


Mean	-3.1	-15.1	4.5	-10.1	17.0	22.1	6.1	4.2	-12.3	-25.2	6.2	17.4	6.4	11.9	-9.8	-19.3
Std. Dev.	24.1	39.8	32.1	49.0	22.3	37.5	16.8	50.3	49.8	47.8	24.7	28.2	17.0	43.8	16.9	34.7
J-B Prob.	0.00	0.17	0.00	0.73	0.42	0.06	0.26	0.05	0.02	0.00	0.24	0.47	0.41	0.00	0.00	0.06
% in $\pm 5\%$ band	14.2	4.0	25.1	16.6	14.0	6.8	17.9	21.3	7.7	6.4	15.1	12.8	17.9	4.5	20.0	8.9

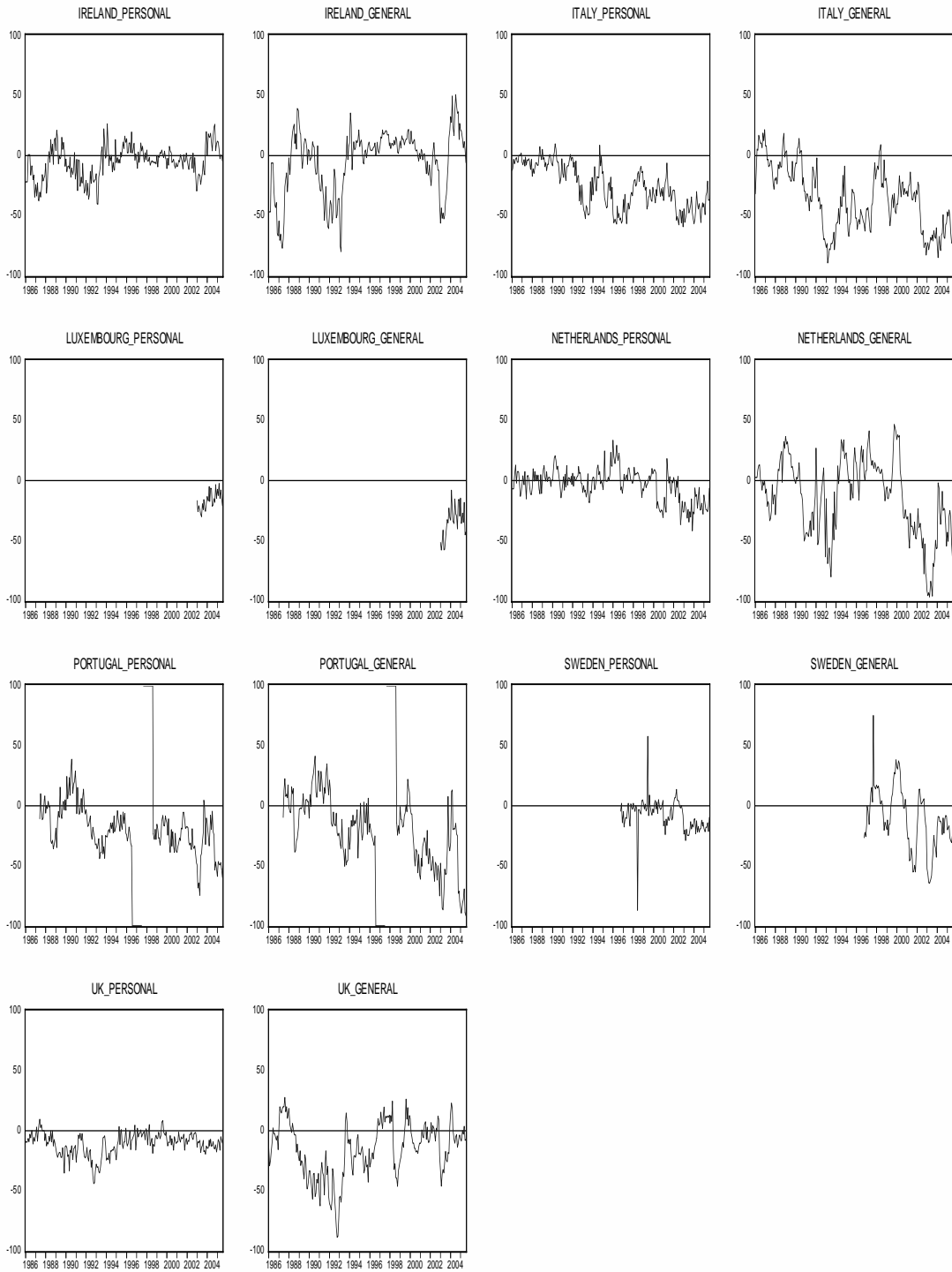


Mean	23.6	28.3	-1.7	-32.2	2.9	-27.5	16.8	4.0	2.1	9.2	12.8	-1.4	9.0	-7.3
Std. Dev.	28.1	40.9	35.5	31.2	16.1	42.6	16.0	49.6	63.6	9.3	21.6	42.6	15.5	33.2
J-B Prob.	0.31	0.00	0.02	0.00	0.77	0.10	0.00	0.00	0.03	0.67	0.00	0.40	0.76	0.18
% in $\pm 5\%$ band	11.1	5.5	13.6	7.7	19.4	9.7	8.5	6.4	18.3	17.9	17.9	5.7	23.4	14.9

P

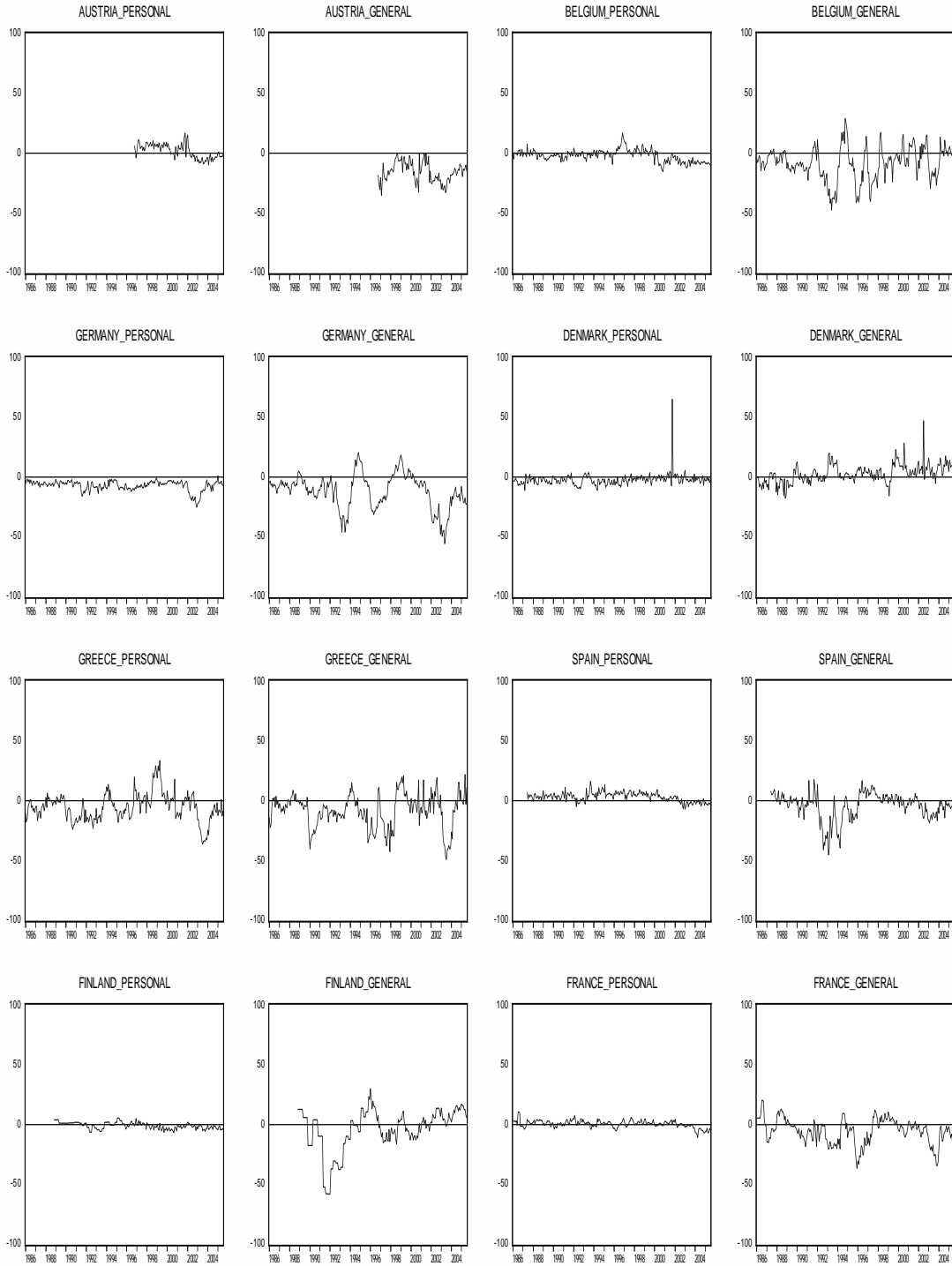


Mean	-28.4	-17.9	-10.0	-16.4	3.8	-10.4	-4.6	-4.4	-24.7	-24.6	-17.3	-10.9	-9.7	-16.7	-23.1	-39.9
Std. Dev.	9.9	23.0	13.2	30.4	16.1	36.0	8.1	21.4	20.9	24.1	10.5	19.2	12.9	29.6	10.5	26.1
J-B Prob.	0.74	0.29	0.88	0.02	0.75	0.00	0.00	0.02	0.01	0.02	0.00	0.00	0.00	0.00	0.42	0.01
in \pm 5% band	0.9	12.3	23.8	8.9	26.8	10.6	44.7	17.0	10.6	8.1	11.0	28.0	44.3	20.9	2.1	6.8

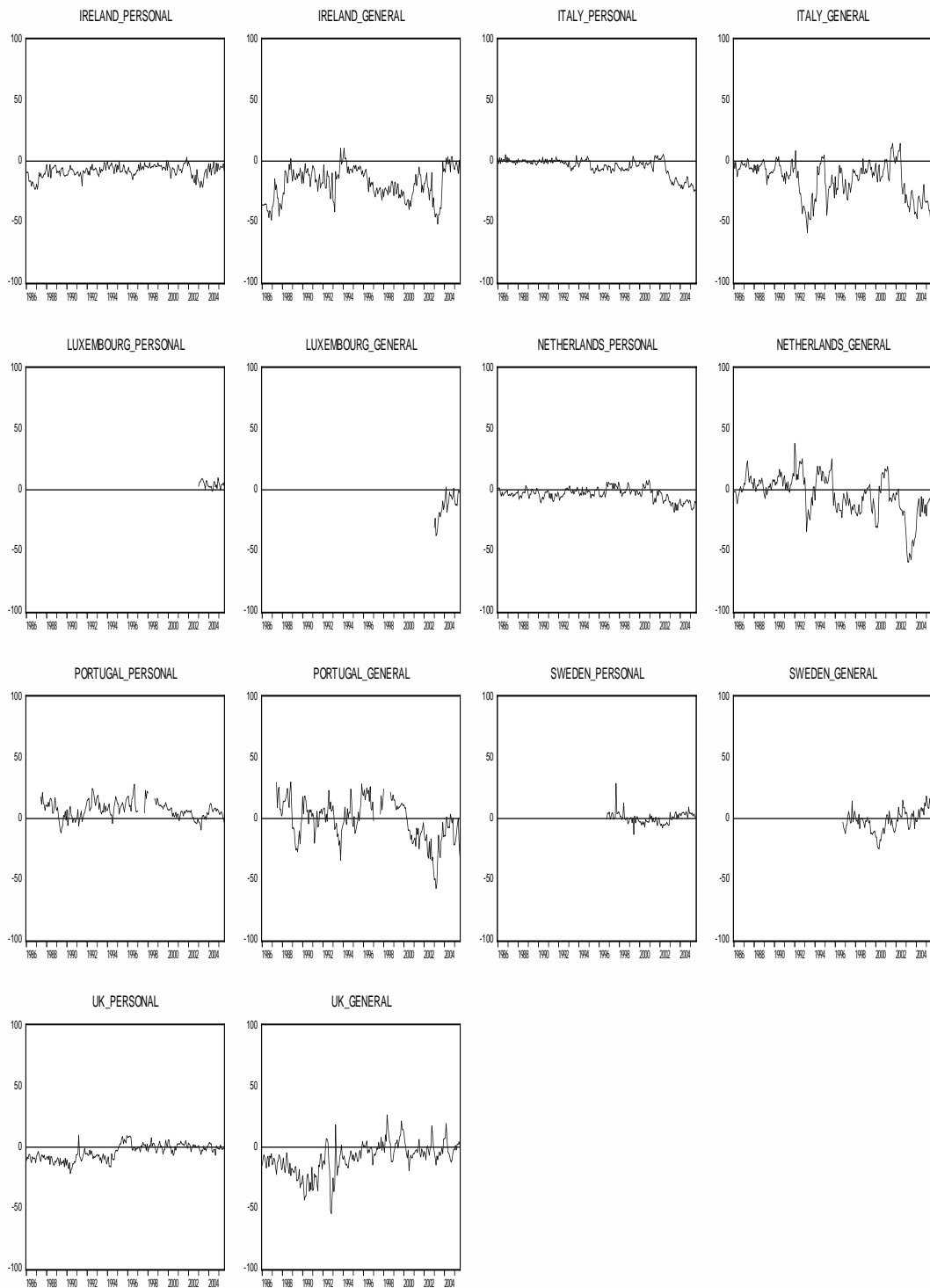


Mean	-5.5	-5.3	-25.5	-35.6	-16.5	-34.0	-3.2	-13.7	-16.9	-18.2	-8.1	-10.7	-10.8	-15.5
Std. Dev.	13.4	26.6	18.0	26.3	7.4	13.5	13.1	31.9	38.3	43.4	13.7	25.8	9.2	22.8
J-B Prob.	0.21	0.00	0.00	0.02	0.60	0.71	0.08	0.01	0.00	0.00	0.00	0.77	0.00	0.00
% in $\pm 5\%$ band	35.7	17.4	14.0	6.8	9.7	0.0	37.0	12.8	9.6	9.2	41.5	14.2	24.3	16.2

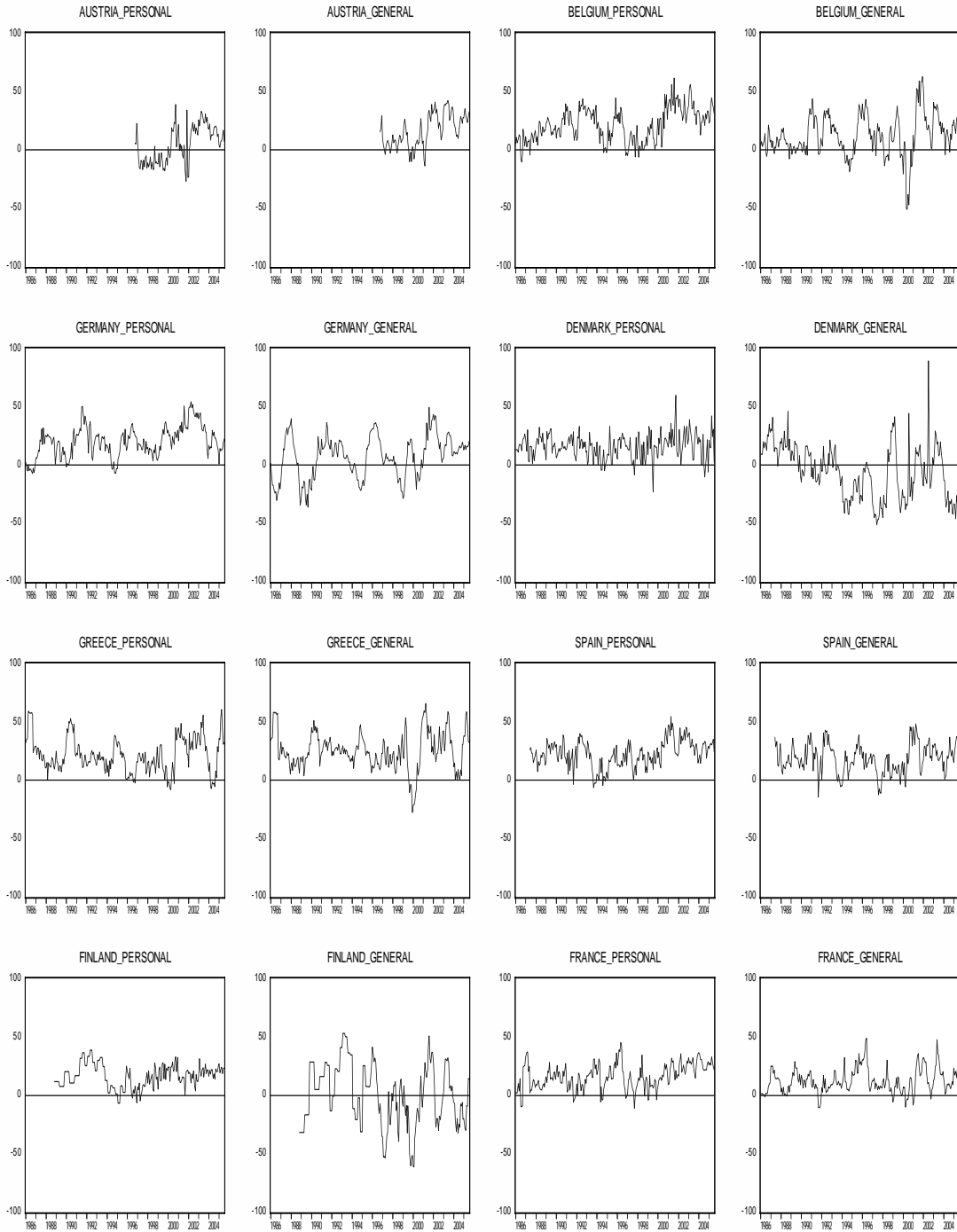
E



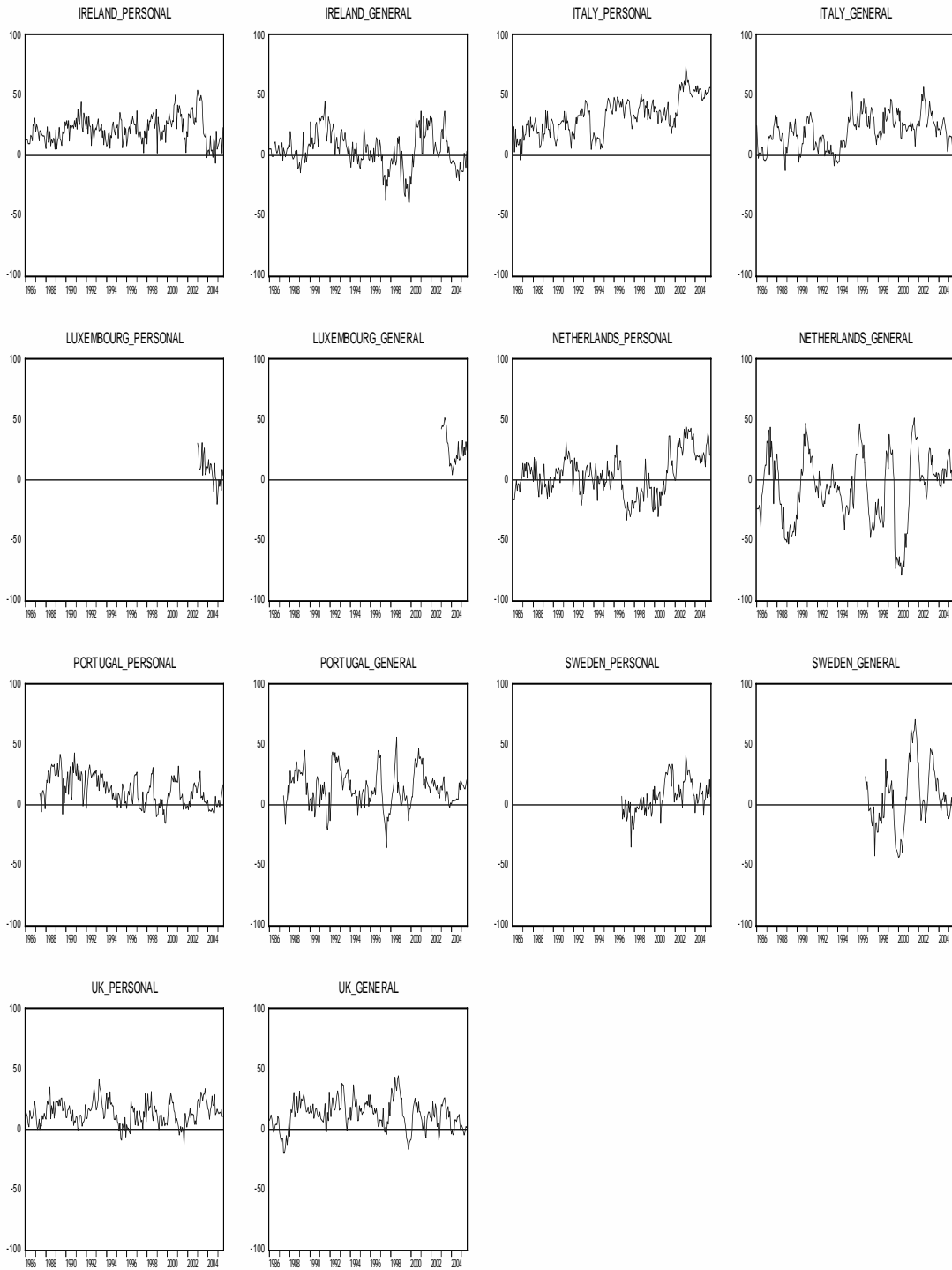
Mean	1.7	-15.7	-2.1	-8.5	-7.2	-12.0	-2.3	2.8	-5.4	-8.4	3.3	-4.4	-0.9	-4.0	0.6	-5.5
Std. Dev.	5.9	8.5	4.9	13.7	4.2	14.5	5.5	7.8	11.6	14.2	3.9	11.2	2.8	17.6	3.1	10.5
J-B Prob.	0.2	0.4	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.0	0.2	0.0	0.0	0.0
% in $\pm 5\%$ band	50.9	13.2	67.2	26.0	34.5	23.0	75.3	50.6	34.9	31.5	66.5	47.7	90.0	26.4	89.8	40.4



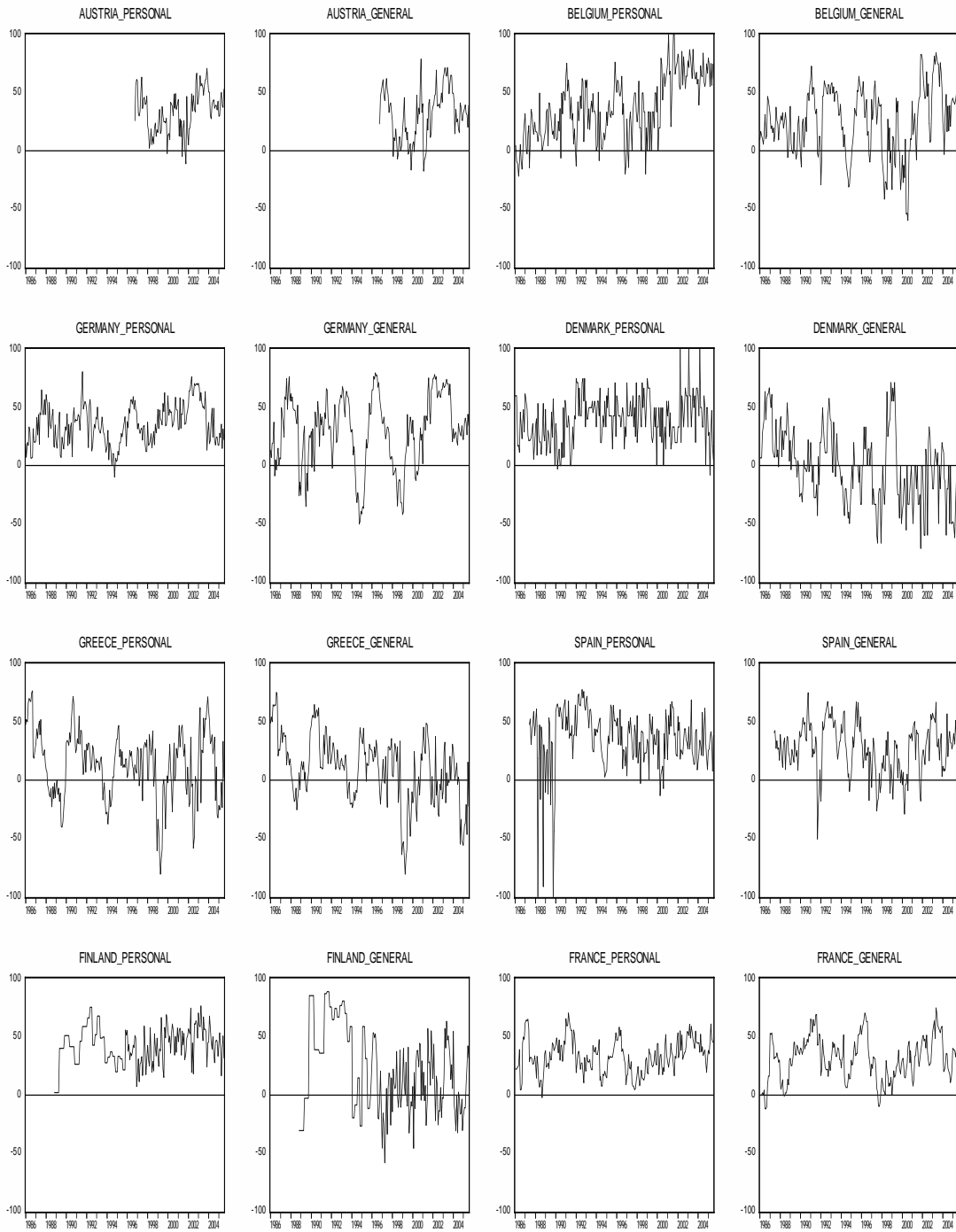
Mean	-8.3	-18.6	-4.5	-15.1	4.5	-13.4	-3.5	-4.0	7.2	-1.7	0.4	-1.7	-3.9	-8.6
Std. Dev.	5.0	12.9	6.7	14.8	3.0	10.7	5.1	16.6	7.0	16.9	4.9	8.7	6.3	12.8
J-B Prob.	0.00	0.03	0.00	0.00	0.66	0.30	0.05	0.00	0.61	0.00	0.00	0.59	0.15	0.00
% in $\pm 5\%$ band	26.4	12.3	64.7	23.0	58.1	25.8	64.3	24.3	36.2	23.9	80.2	46.2	47.7	29.8



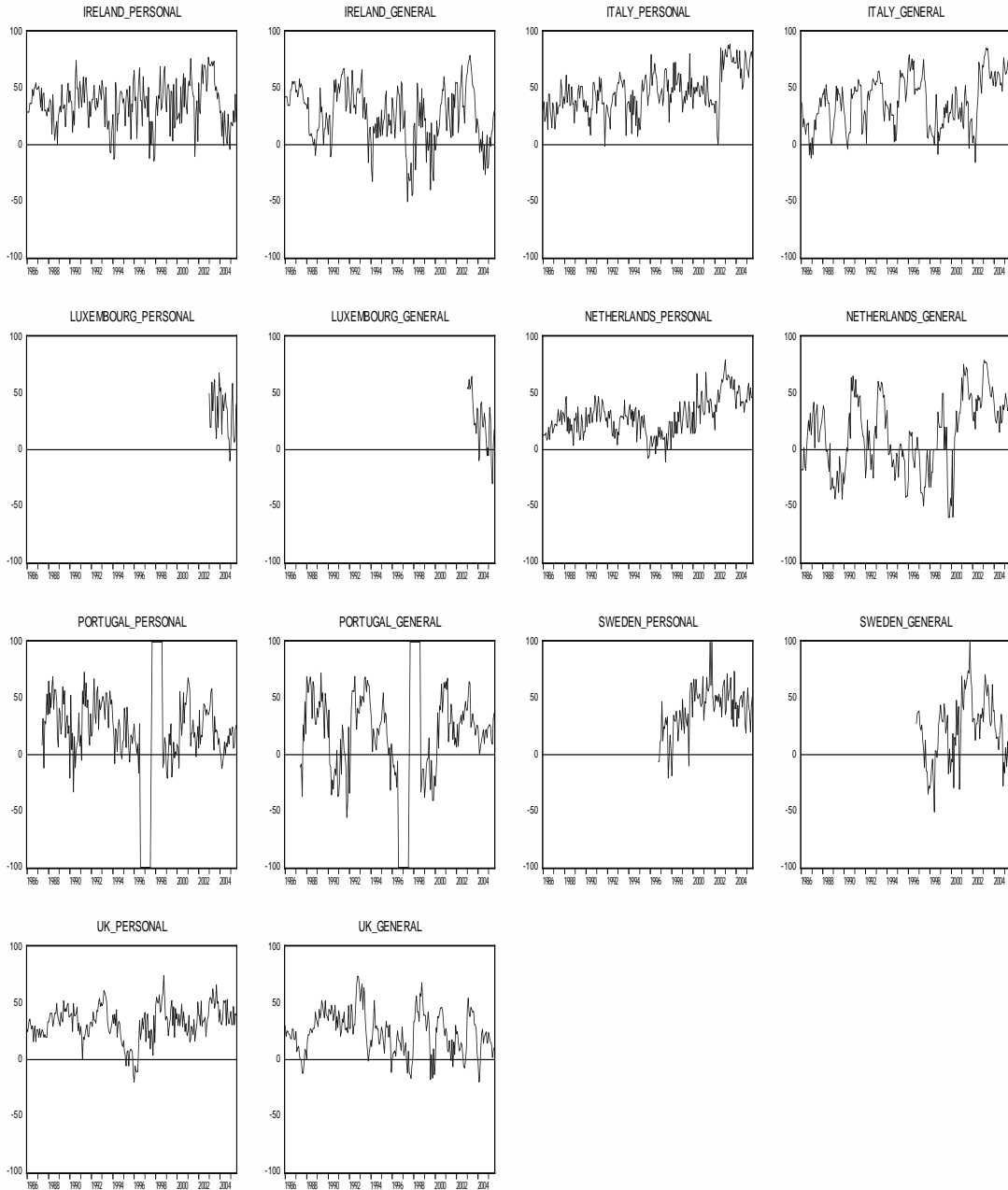
Mean	4.6	16.6	21.5	12.7	20.2	7.1	15.9	-5.3	21.8	25.3	22.9	18.7	16.7	0.6	16.4	12.6
Std. Dev.	15.8	14.0	14.2	17.8	13.4	18.1	10.6	22.8	14.9	16.0	11.4	12.6	9.9	26.3	10.9	10.6
J-B Prob.	0.1	0.1	0.2	0.0	0.5	0.1	0.0	0.3	0.0	0.1	0.7	0.7	0.5	0.1	0.6	0.0
% in $\pm 5\%$ band	25.5	26.4	12.8	21.7	11.1	15.7	15.7	13.6	8.9	3.8	6.9	11.0	10.4	9.0	12.3	15.7



Mean	21.7	4.6	32.8	20.3	7.6	26.1	4.4	-6.3	12.1	14.1	7.1	6.2	13.8	12.7
Std. Dev.	10.8	15.2	14.9	13.5	12.1	13.3	16.9	28.5	12.4	15.1	13.5	26.2	9.4	12.1
J-B Prob.	0.1	0.7	0.2	0.4	1.0	0.3	0.2	0.1	0.0	0.8	0.8	0.3	0.7	1.0
% in $\pm 5\%$ band	5.1	29.8	1.7	13.6	25.8	3.2	21.7	16.6	25.2	18.3	29.2	17.0	16.2	17.4



Mean	34.1	31.8	36.9	25.4	35.6	29.3	39.5	-0.7	13.3	9.6	35.6	27.1	42.3	21.3	33.5	31.2
Std. Dev.	17.7	22.2	27.2	28.6	17.1	30.5	19.9	31.1	29.1	28.8	25.9	21.7	16.0	36.0	14.6	18.7
J-B Prob.	0.4	0.3	0.1	0.1	0.2	0.0	0.9	0.4	0.1	0.0	0.0	0.0	0.2	0.0	0.3	0.3
% in $\pm 5\%$ band	3.8	7.5	8.1	9.8	1.3	5.1	4.3	18.7	11.5	11.5	1.8	5.0	3.0	10.4	1.7	7.2



Mean	36.5	25.6	46.7	38.2	33.0	24.4	29.5	13.7	21.8	18.4	40.5	22.0	32.8	24.4
Std. Dev.	19.3	25.7	18.4	22.7	20.8	24.7	16.6	32.4	39.8	43.7	21.3	28.7	15.2	19.0
J-B Prob.	0.1	0.0	0.7	0.1	0.4	0.6	0.1	0.0	0.0	0.0	0.0	0.9	0.0	0.9
% in $\pm 5\%$ band	3.8	5.5	0.9	5.1	3.2	6.5	3.0	14.0	8.3	6.0	1.9	6.6	1.3	8.5