
Inflation Expectations, Labour Markets and EMU

UNIVERSITY OF OXFORD DPHIL THESIS

Fabien Curto Millet

(Balliol College)

INFLATION EXPECTATIONS, LABOUR MARKETS AND EMU

Fabien Curto Millet
Balliol College, University of Oxford

Thesis submitted for the degree of DPhil Economics, Trinity Term 2006
Granted leave to supplicate 1 May 2007

ABSTRACT

This thesis examines the measurement, applications and properties of consumer inflation expectations in the context of eight European Union countries: France, Germany, the UK, Spain, Italy, Belgium, the Netherlands and Sweden.

The data proceed mainly from the European Commission's Consumer Survey and are qualitative in nature, therefore requiring quantification prior to use. This study first seeks to determine the optimal quantification methodology among a set of approaches spanning three traditions, associated with Carlson-Parkin (1975), Pesaran (1984) and Seitz (1988). The success of a quantification methodology is assessed on the basis of its ability to match quantitative expectations data and on its behaviour in an important economic application, namely the modelling of wages for our sample countries. The wage equation developed here draws on the theoretical background of the staggered contracts and the wage bargaining literature, and controls carefully for inflation expectations and institutional variables. The Carlson-Parkin variation proposed in Curto Millet (2004) was found to be the most satisfactory. This being established, the wage equations are used to test the hypothesis that the advent of EMU generated an increase in labour market flexibility, which would be reflected in structural breaks. The hypothesis is essentially rejected. Finally, the properties of inflation expectations and perceptions themselves are examined, especially in the context of EMU. Both the rational expectations and rational perceptions hypotheses are rejected. Popular expectations mechanisms, such as the "rule-of-thumb" model or Akerlof et al.'s (2000) "near-rationality hypothesis" are similarly unsupported. On the other hand, evidence is found for the transmission of expert forecasts to consumer expectations in the case of the UK, as in Carroll's (2003) model. The *distribution* of consumer expectations and perceptions is also considered, showing a tendency for *gradual* (as in Mankiw and Reis, 2002) but non-rational adjustment. Expectations formation is further shown to have important qualitative features.

JEL Classifications:

D84, E31, E24, J30, J31, J32, J33, J38, J51, J58, J80, J88, E52, E58, E61, E65, E27, N14, N34, C22, C32, C42, C51, C52, C53, C82.

Keywords:

Inflation expectations, inflation perceptions, survey data, quantification, wage equations, labour markets, EMU, euro, flexibility, competitiveness, labour market institutions, rationality, Phillips curve, consumers, expert forecasts, expectations distribution, inflation targeting.

TABLE OF CONTENTS

INTRODUCTION	11
---------------------------	-----------

CHAPTER I:

Finding the Optimal Method of Quantifying Inflation Expectations on the Basis of Qualitative Survey Data

1 INTRODUCTION	14
2 QUALITATIVE DATA: THE EC CONSUMER SURVEY.....	16
3 OVERVIEW OF QUANTIFICATION METHODOLOGIES.....	19
3.1 THE CARLSON-PARKIN OR PROBABILITY APPROACH	19
3.2 THE REGRESSION APPROACH.....	30
3.3 MODELS WITH STOCHASTIC TIME-VARYING PARAMETERS.....	37
3.4 COMPARATIVE THEORETICAL DISCUSSION AND SUMMARY	49
4 COMPARING METHODS ON THE BASIS OF PREDICTIVE ABILITY	53
4.1 INTRODUCTION	53
4.2 RESULTS	53
5 QUANTITATIVE SURVEY DATA	60
5.1 UNITED KINGDOM: THE GALLUP SURVEY	60
5.2 SWEDEN: THE HIP SURVEY (SCB/GfK)	62
6 BENCHMARKING QUALITATIVE DATA ON QUANTITATIVE DATA.....	65
6.1 PREDICTIVE PERFORMANCE STATISTICS FOR THE QUANTITATIVE MEASURES	65
6.2 DESCRIPTIVE STATISTICS	65
6.3 NON-NESTED TESTING OF QUANTIFICATION METHODS	69
7 INFLATION EXPECTATIONS MEASURES IN WAGE EQUATIONS.....	73
7.1 MODEL SPECIFICATION AND SIGN PRIORS	73
7.2 NOTES ON THE ECONOMETRIC APPROACH.....	76
7.3 RESULTS	77
7.3.1 Sweden.....	80
7.3.2 United Kingdom	82
7.3.3 France	83
7.3.4 Spain	84
7.3.5 Belgium.....	86
7.3.6 The Netherlands	86
7.3.7 Italy.....	88
7.3.8 Germany	89
8 CONCLUSION	90

CHAPTER II: The Impact of European Monetary Unification on Labour Market Flexibility

1	INTRODUCTION	94
1.1	MOTIVATION AND CHAPTER PLAN	94
1.2	THE MEANING OF LABOUR MARKET FLEXIBILITY	97
1.3	SUGGESTIVE EVIDENCE OF LABOUR MARKET CHANGES LINKED TO EMU	98
2	THEORETICAL BACKGROUND AND LITERATURE REVIEW.....	100
2.1	THEORETICAL FOUNDATIONS AND EXPECTATIONS	100
2.1.1	<i>The sources of nominal wage rigidity and their implications for EMU.....</i>	<i>101</i>
2.1.2	<i>Real wage rigidity: sources and expectations</i>	<i>106</i>
2.2	GUIDANCE FROM THE EMS AND OTHER PRE-EMU EMPIRICAL STUDIES.....	107
2.2.1	<i>Country case studies.....</i>	<i>108</i>
2.2.2	<i>Multi-country econometric studies</i>	<i>109</i>
2.3	EMU AND LABOUR MARKET FLEXIBILITY: ECONOMETRIC STUDIES.....	114
2.4	INTERIM CONCLUSION	121
3	ECONOMETRIC SPECIFICATION.....	124
3.1	THEORETICAL FOUNDATIONS OF OUR WAGE EQUATION.....	124
3.1.1	<i>Long-run determinants of wages</i>	<i>124</i>
3.1.2	<i>Dynamics of the wage equation.....</i>	<i>126</i>
3.2	THE EMPIRICAL WAGE EQUATION	129
3.3	ECONOMETRIC APPROACH	132
3.4	THE TIME AGGREGATION PROBLEM	134
3.5	CORRESPONDENCE BETWEEN MEASURES AND CONCEPTS	135
3.6	DATA DESCRIPTION AND MOTIVATION	138
3.6.1	<i>General macroeconomic variables</i>	<i>138</i>
3.6.2	<i>Wedge variables</i>	<i>143</i>
3.6.3	<i>Institutional variables.....</i>	<i>146</i>
3.6.4	<i>General Remark</i>	<i>159</i>
4	EMPIRICAL RESULTS AND ANALYSIS.....	159
4.1	COUNTRY-SPECIFIC RESULTS	161
4.1.1	<i>Italy.....</i>	<i>161</i>
4.1.2	<i>France</i>	<i>176</i>
4.1.3	<i>Belgium.....</i>	<i>190</i>
4.1.4	<i>Spain.....</i>	<i>201</i>
4.1.5	<i>Germany</i>	<i>208</i>
4.1.6	<i>The Netherlands</i>	<i>219</i>
4.1.7	<i>United Kingdom</i>	<i>229</i>
4.1.8	<i>Sweden.....</i>	<i>235</i>
4.2	COMPARATIVE OVERVIEW OF THE RESULTS.....	241
4.3	POOLED MEAN GROUP (PMG) ESTIMATION	248
5	CONCLUSION	255
5.1	SUMMARY OF KEY POINTS.....	255
5.2	LOOKING AHEAD: IMPLICATIONS FOR EMU.....	257

CHAPTER III: Properties and Applications of Data on Inflation Perceptions and Expectations

1	INTRODUCTION	264
2	UNDERSTANDING INFLATION EXPECTATIONS AND PERCEPTIONS.....	268
2.1	DESCRIPTIVE ANALYSIS OF OUR INFLATION EXPECTATIONS AND PERCEPTIONS SERIES	268
2.1.1	<i>Inflation perceptions.....</i>	268
2.1.2	<i>An additional justification for our quantification method.....</i>	274
2.1.3	<i>Inflation expectations</i>	277
2.2	UNDERSTANDING THE LINK BETWEEN INFLATION EXPECTATIONS AND PERCEPTIONS	282
2.3	UNDERSTANDING THE FORMATION OF INFLATION PERCEPTIONS	287
2.3.1	<i>Modelling inflation perceptions</i>	287
2.3.2	<i>Impact on assessments of the rationality of expectations and quantification</i>	298
2.3.3	<i>Conclusion.....</i>	303
3	TESTING THE ‘NEAR-RATIONALITY’ HYPOTHESIS	306
3.1	INTRODUCTION	306
3.2	TESTING NEAR-RATIONALITY IN THE CONTEXT OF REGIME CHANGES.....	308
3.3	TESTING NEAR-RATIONALITY IN THE CONTEXT OF TRANSITORY CHANGES IN INFLATION.....	314
3.4	CONCLUSION.....	316
4	FOUNDATIONAL MODELS OF INFLATION EXPECTATIONS	318
4.1	INTRODUCTION	318
4.2	OVERVIEW OF KEY DEVELOPMENTS.....	319
4.2.1	<i>From adaptive to rational expectations</i>	319
4.2.2	<i>The New Keynesian Phillips Curve literature and further developments</i>	322
4.3	THE ADAPTIVE EXPECTATIONS MODEL	325
4.4	INVESTIGATING THE RULE-OF-THUMB APPROACH	327
4.4.1	<i>Presenting the Gerberding (2001) results.....</i>	327
4.4.2	<i>Replicating the Gerberding (2001) results.....</i>	330
4.4.3	<i>Sample sensitivity</i>	335
4.4.4	<i>Sensitivity to expectations measure.....</i>	337
4.4.5	<i>Conclusion and further implications of our ‘rule-of-thumb’ results.....</i>	341
4.5	CONCLUSION.....	344
5	THE EXPECTATIONS OF CONSUMERS & ECONOMIC EXPERTS.....	348
5.1	INTRODUCTION	348
5.2	INTRODUCING THE CONSENSUS FORECASTS SURVEY	349
5.3	FORECASTING POWER OF CONSUMERS’ AND EXPERTS’ EXPECTATIONS	353
5.4	THE RELATIONSHIP BETWEEN THE EXPECTATIONS OF CONSUMERS AND EXPERTS	358
5.5	INFLATION NEWS COVERAGE AND INFLATION EXPECTATIONS IN THE UK	367
5.6	CONCLUSION.....	378
6	THE ECONOMICS OF DISAGREEMENT	384
6.1	INTRODUCTION	384
6.2	THE “STICKY INFORMATION” MODEL OF MANKIW AND REIS (2002)	386
6.2.1	<i>Overview.....</i>	386
6.2.2	<i>Testing the model across the Atlantic: the Volcker and ERM disinflations</i>	388
6.3	DOUBLE-PEAKEDNESS THROUGH SURVEY DESIGN: THE VAN LELYVELD HYPOTHESIS.....	397
6.3.1	<i>The arguments and evidence presented in van Lelyveld (2000)</i>	398
6.3.2	<i>A critique of the van Lelyveld (2000) double-peakedness hypothesis.....</i>	401
6.3.3	<i>Conclusion.....</i>	409
6.4	ANALYSING DISAGREEMENT IN UK QUANTITATIVE GALLUP DATA	410
6.5	IS THE CARLSON-PARKIN INTERQUANTILE RANGE A GOOD MEASURE OF DISAGREEMENT?	415

6.6	QUALITATIVE ANALYSIS OF THE EXPECTATIONS DISTRIBUTION	419
6.6.1	<i>A note on the choice between the mean and median of the distribution</i>	420
6.6.2	<i>Introducing the Bank of England / NOP Survey</i>	423
6.6.3	<i>Studying the pattern of “don’t know” responses across surveys</i>	428
6.6.4	<i>“Zero barrier”, “digit preferences” and the impact of inflation targeting</i>	441
6.6.5	<i>On the comparative advantages of quantitative and qualitative surveys</i>	458
7	CONCLUSION	461

CONCLUSION	464
------------------	-----

REFERENCES	472
------------------	-----

APPENDIX I	490
-------------------	-----

PLOTS OF INFLATION EXPECTATIONS AND PERCEPTIONS DATA

APPENDIX II	502
--------------------	-----

ADDITIONAL DATA AND DIAGRAMS FOR CHAPER III

ACKNOWLEDGEMENTS	517
------------------------	-----

LIST OF TABLES

TABLE 1: EC CONSUMER SURVEY, QUESTIONS 5 AND 6	17
TABLE 2: EC CONSUMER SURVEY, SAMPLE COVERAGE BY COUNTRY	18
TABLE 3: PESARAN REGRESSION FOR THE UK, NLS, 1981m11-2005m7	35
TABLE 4: SMITH & McALEER REGRESSION, UK, NLS, 1981m11-2005m7	37
TABLE 5: SUMMARY OF SPECIFICATION SELECTION FOR TVP MODEL	46
TABLE 6: EXPECTATIONS MEASURES SELECTED FOR COMPARISON	49
TABLE 7: FREQUENCY OF ZEROES IN EC MONTHLY SURVEY DATA	50
TABLE 8: RMSE RESULTS FOR PREDICTING ACTUAL INFLATION	54
TABLE 9: AVERAGE RANK & RMSE FOR PREDICTIVE SUCCESS	55
TABLE 10: MSE DECOMPOSITION, DETAILS	57
TABLE 11: MSE DECOMPOSITION, AVERAGES BY MEASURE	57
TABLE 12: % CORRECTLY PREDICTED TURNING POINTS	58
TABLE 13: AVERAGE RANK, SUCCESS IN TURNING POINT PREDICTION	59
TABLE 14: HIP SURVEY, QUANTITATIVE QUESTIONS	62
TABLE 15: PREDICTIVE PERFORMANCE OF QUANTITATIVE MEASURES	65
TABLE 16: CORRELATIONS WITH QUANTITATIVE MEASURES	66
TABLE 17: ABILITY TO MATCH QUANTITATIVE DATA	67
TABLE 18: % CORRECTLY INDICATED TURNING POINTS	67
TABLE 19: NON-NESTED TESTING RESULTS, UK	70
TABLE 20: NON-NESTED TESTING RESULTS, SWEDEN	70
TABLE 21: NON-NESTED RESULTS WITH HOMOGENEITY RESTRICTION	72
TABLE 22: WAGE EQUATION STRUCTURE	74
TABLE 23: EMPIRICAL WAGE EQUATION, SIGN PRIORS	75
TABLE 24: WAGE EQUATIONS, RETAINED LAGS OF EXPECTATIONS	77
TABLE 25: STUDIES DEALING WITH THE EMS AND OTHER PRE-EMU EXPERIENCES	113
TABLE 26: ECONOMETRIC STUDIES ON EMU AND LABOUR MARKET FLEXIBILITY	120
TABLE 27: WAGE EQUATION STRUCTURE	130

TABLE 28: COEFFICIENT SIGN PRIORS	130
TABLE 29: BARGAINING COVERAGE DATA – SAMPLES BY COUNTRY	152
TABLE 30: FULL-TIME EMPLOYEES EARNING THE MINIMUM WAGE (%)	179
TABLE 31: SIGNIFICANT INSTITUTIONAL VARIABLES, SUMMARY	245
TABLE 32: WAGE FLEXIBILITY MEASURES, ALL COUNTRIES	245
TABLE 33: COUNTRY RANKINGS, SPEED OF EQUILIBRIUM CORRECTION	246
TABLE 34: MAINTAINED HYPOTHESIS, H_U	251
TABLE 35: NESTED TEST RESULTS	252
TABLE 36: HICP COMPONENTS	290
TABLE 37: MODELLING INFLATION PERCEPTIONS (SWE), 1993m1-2005m7	291
TABLE 38: WEIGHTS OF HICP COMPONENTS FOR SWEDEN, 1990-2005	503
TABLE 39: SWEDISH INFLATION PERCEPTIONS, NESTED TEST RESULTS	293
TABLE 40: PARSIMONIOUS EqCM MODEL, INFLATION PERCEPTIONS	294
TABLE 41: HICP COMPONENTS	295
TABLE 42: FACTORS HYPOTHESISED TO AFFECT INFLATION PERCEPTIONS	296
TABLE 43: PREDICTIVE ABILITY, PERCEPTIONS-ADJUSTED INFLATION	300
TABLE 44: PREDICTIVE ABILITY, HICP INFLATION	301
TABLE 45: ABILITY TO MATCH QUANTITATIVE DATA	302
TABLE 46: RESULTS OF BAI-PERRON TESTS FOR INFLATION REGIMES	311
TABLE 47: SWEDEN, EXPECTATION ERRORS & INFLATION BY REGIME	312
TABLE 48: EC DATA, EXPECTATION ERRORS & INFLATION REGIMES	313
TABLE 49: TESTING ADP NEAR-RATIONALITY, ALL FLUCTUATIONS	315
TABLE 50: THE ADAPTIVE EXPECTATIONS MODEL	326
TABLE 51: GERBERDING (2001) RULE-OF-THUMB MODEL; 1991q1-1999q4	328
TABLE 52: REPLICATION OF GERBERDING'S EQUATIONS; 1991q1-1999q4	330
TABLE 53: DIAGNOSTICS FOR GERBERDING RULE-OF-THUMB MODEL	331
TABLE 54: RULE-OF-THUMB MODEL, ALL COUNTRIES, 1991q1-1999q4	333
TABLE 55: RULE-OF-THUMB MODEL, DIAGNOSTICS, 1991q1-1999q4	334
TABLE 56: RULE-OF-THUMB MODEL, FULL SAMPLE ESTIMATION	335
TABLE 57: RULE-OF-THUMB MODEL, DIAGNOSTICS, FULL SAMPLE	336
TABLE 58: RULE-OF-THUMB MODEL, EXPINFLPBIS, 1991q1-1999q4	338
TABLE 59: DIAGNOSTICS FOR EXPINFLPBIS RULE-OF-THUMB MODEL	338
TABLE 60: RULE-OF-THUMB MODEL, QUANTITATIVE DATA	340
TABLE 61: RULE-OF-THUMB, DIAGNOSTICS, QUANTITATIVE DATA	341
TABLE 62: CONSENSUS FORECASTS, SAMPLE USED IN THIS STUDY	350
TABLE 63: PREDICTIVE ABILITY OF CONSUMERS AND EXPERTS	353
TABLE 64: PREDICTIVE ABILITY CONSUMERS/EXPERTS, SUMMARY	354
TABLE 65: CONSUMERS & EXPERTS, INTRINSIC PREDICTIVE ABILITY	356
TABLE 66: GRANGER CAUSALITY, CONSUMER & EXPERT FORECASTS	359
TABLE 67: REGRESSION RESULTS FOR CARROLL'S (2003) MODEL	362
TABLE 68: DOES NEWS COVERAGE IMPROVE FORECAST ACCURACY?	373
TABLE 69: NEWS EFFECT ON THE CONSUMER/EXPERT FORECAST GAP	374
TABLE 70: NEWS COVERAGE AND SPEED OF UPDATING, MODEL 1	376
TABLE 71: NEWS COVERAGE AND SPEED OF UPDATING, MODEL 2	377
TABLE 72: LABOUR AND HOUSING MARKET INDICATORS	381
TABLE 73: UK INFLATION EXPECTATIONS, DISTRIBUTION DATA, 1991m1	391
TABLE 74: REGRESSIONS OF GALLUP IQR, 1991m1-1997m1	413
TABLE 75: REGRESSIONS OF GALLUP IQR, 1983m1-1997m1	414
TABLE 76: REGRESSIONS OF GALLUP IQR ON CP IQR, 1983m1-1997m1	418
TABLE 77: BANK OF ENGLAND / NOP SURVEY, QUESTIONS 1 AND 2	424
TABLE 78: 'DON'T KNOW' ANSWERS, BY SURVEY AND SAMPLE PERIOD	430
TABLE 79: GENERAL ELECTION EFFECTS, 'DON'T KNOW' RESPONSES	434
TABLE 80: US, UK, SWEDEN % OF "DON'T KNOW" RESPONSES	449
TABLE 81: KEY DATES IN UK MONETARY HISTORY	451

LIST OF FIGURES

FIGURE 1: AGGREGATE DENSITY FUNCTIONS: EXPECTATIONS, PERCEPTIONS	21
FIGURE 2: UK, CP EXPECTED INFLATION & INFLATION, 1981m11-2005m9	24
FIGURE 3: FRA, CP TRICHOTOMOUS QUANTIFICATION, 1981m11-2005m9	27
FIGURE 4: UK, CP PENTACHOMOUS QUANTIFICATION, 1981m11-2005m9	30
FIGURE 5: ALTERNATIVE EXPECTATIONS SERIES, PESARAN TRADITION	36
FIGURE 6: SEITZ METHOD; PERCEPTIONS & EXPECTATIONS DISTRIBUTIONS	38
FIGURE 7: MONTE CARLO SIMULATION RESULTS; COOLEY-PRESCOTT	43
FIGURE 8: KALMAN FILTER SIMULATION RESULTS, 1981m11-2005m6, UK	44
FIGURE 9: TIME-VARYING INDIFFERENCE THRESHOLDS, UK	48
FIGURE 10: UK INFLATION EXPECTATIONS, SEITZ METHODOLOGY	48
FIGURE 11: QUANTITATIVE GALLUP DATA & ACTUAL CPI INFLATION, UK	61
FIGURE 12: HIP PERCEPTIONS vs ACTUAL INFLATION, SWEDEN, 1979q1-2005q3	64
FIGURE 13: HIP EXPECTATIONS vs ACTUAL INFLATION, SWEDEN, 1979q1-2005q3	64
FIGURE 14: WORKER HOURLY COMPENSATION IN EUROS, EUROZONE	140
FIGURE 15: WORKER HOURLY COMPENSATION, SWEDEN (SEK) AND UK (£)	140
FIGURE 16: PRODUCTIVITY TREND INDEX, BY COUNTRY	142
FIGURE 17: STANDARDISED UNEMPLOYMENT RATES, BY COUNTRY	143
FIGURE 18: DIRECT/INCOME TAX RATE (%) 1970q1-2005q4	144
FIGURE 19: INDIRECT/CONSUMPTION TAX RATE (%) 1970q1-2005q4	145
FIGURE 20: RATIO OF IMPORT TO CONSUMER PRICES	146
FIGURE 21: BENEFIT REPLACEMENT RATES 1970q1-2005q4	148
FIGURE 22: BENEFIT DURATION INDEX 1970q1-2005q4	149
FIGURE 23: NET UNION DENSITY (%) (1970q1-2005q4)	151
FIGURE 24: ADJUSTED BARGAINING COVERAGE (%) (1970q1-2005q4)	153
FIGURE 25: BARGAINING COORDINATION INDEX, 1970q1-2005q4	157
FIGURE 26: EMPLOYMENT PROTECTION LEGISLATION [0-6], 1970q1-2005q4	158
FIGURE 27: ITALY, RECURSIVE COEFFICIENT ON DIRECT TAXATION (LCTAX)	164
FIGURE 28: ITALIAN MODEL, GRAPHICAL ANALYSIS, 1973q1-2005q2	166
FIGURE 29: ITALIAN MODEL, STABILITY TESTS	169
FIGURE 30: ITALIAN MODEL, RECURSIVE ECM COEFFICIENTS	170
FIGURE 31: DEPENDENT VARIABLE AND UNEMPLOYMENT, RECURSIVE	171
FIGURE 32: ITALY, INFLATION EXPECTATIONS, RECURSIVE COEFFICIENT	172
FIGURE 33: ITALIAN MODEL, FORECASTS FROM 1999:1	175
FIGURE 34: ITALIAN MODEL, RECURSIVE ECM COEFFICIENTS	175
FIGURE 35: BENEFIT DURATION AND REPLACEMENT RATE COEFFICIENTS	178
FIGURE 36: WAGE DISTRIBUTION EXPRESSED IN SMIC MULTIPLES	180
FIGURE 37: AVERAGE HOURS PER EMPLOYEE AND RECURSIVE COEFFICIENT	183
FIGURE 38: FRENCH MODEL, GRAPHICAL ANALYSIS, 1982q4-2005q2	187
FIGURE 39: FRENCH MODEL, FORECASTS FROM 1999:1	187
FIGURE 40: FRENCH MODEL, FORECASTS FROM 2001:1	188
FIGURE 41: FRENCH MODEL, STABILITY TESTS	188
FIGURE 42: FRENCH MODEL, RECURSIVE ECM COEFFICIENTS	189
FIGURE 43: RECURSIVE COEFFICIENT, DLPRODI	189
FIGURE 44: LOG HOURLY COMPENSATION OF WORKERS (LRCOMPTh)	190
FIGURE 45: BELGIUM, RECURSIVE COEFFICIENT, INFLATION EXPECTATIONS	195
FIGURE 46: BELGIAN MODEL, GRAPHICAL ANALYSIS, 1985q3-2005q2	199
FIGURE 47: BELGIAN MODEL, FORECASTS FROM 1999:1	200
FIGURE 48: BELGIAN MODEL, STABILITY TESTS	200
FIGURE 49: BELGIAN MODEL, RECURSIVE ECM COEFFICIENTS	201
FIGURE 50: SPANISH MODEL, GRAPHICAL ANALYSIS, 1987q2-2005q2	206
FIGURE 51: SPANISH MODEL, FORECASTS FROM 1999:1	207
FIGURE 52: SPANISH MODEL, STABILITY TESTS	207
FIGURE 53: SPANISH MODEL, RECURSIVE ECM COEFFICIENTS	208

FIGURE 54: INFLATION EXPECTATIONS IN GERMANY, 1980Q1-2005Q2	211
FIGURE 55: RECURSIVE COEFFICIENT, GERMAN INFLATION EXPECTATIONS	213
FIGURE 56: GERMAN MODEL, GRAPHICAL ANALYSIS, 1991Q3-2005Q2	216
FIGURE 57: GERMAN MODEL, FORECASTS FROM 1999:1	216
FIGURE 58: GERMAN MODEL, FORECASTS FROM 2000:1	217
FIGURE 59: GERMAN MODEL, STABILITY TESTS	217
FIGURE 60: GERMAN MODEL, RECURSIVE ECM COEFFICIENTS	217
FIGURE 61: LABOUR SHARE OF GROSS VALUE ADDED (BASIC PRICES), SA	218
FIGURE 62: DUTCH INFLATION EXPECTATIONS, RECURSIVE COEFFICIENT	223
FIGURE 63: DUTCH MODEL, GRAPHICAL ANALYSIS, 1974Q3-2005Q2	227
FIGURE 64: DUTCH MODEL, FORECASTS FROM 1999Q1	227
FIGURE 65: DUTCH MODEL, STABILITY TESTS	228
FIGURE 66: DUTCH MODEL, RECURSIVE ECM TERMS	228
FIGURE 67: UK MODEL, GRAPHICAL ANALYSIS, 1983Q1-2005Q1	233
FIGURE 68: UK MODEL, FORECASTS FROM 1999:1	233
FIGURE 69: UK MODEL, STABILITY TESTS	234
FIGURE 70: UK MODEL, RECURSIVE ECM COEFFICIENTS	234
FIGURE 71: SWEDISH MODEL, GRAPHICAL ANALYSIS, 1983Q1-2005Q1	239
FIGURE 72: SWEDISH MODEL, FORECASTS FROM 1999:1	239
FIGURE 73: SWEDISH MODEL, STABILITY TESTS	240
FIGURE 74: SWEDISH MODEL, RECURSIVE ECM COEFFICIENTS	240
FIGURE 75: MACROECONOMIC VOLATILITY INDEX, ALL COUNTRIES	260
FIGURE 76: ACTUAL INFLATION & PERCEPTIONS BY COUNTRY, POST-1996	268
FIGURE 77: ACTUAL INFLATION & PERCEPTIONS, GERMANY, POST-1996	269
FIGURE 78: GERMANY, RESPONSE SHARES (PERCEIVED INFLATION QUESTION)	270
FIGURE 79: EUROZONE COUNTRIES, RESPONSE SHARES (PERCEPTIONS)	272
FIGURE 80: UK AND SWEDEN, RESPONSE SHARES (INFLATION PERCEPTIONS)	274
FIGURE 81: GERMAN INFLATION PERCEPTIONS POST-1996, CP METHOD	275
FIGURE 82: ACTUAL INFLATION & EXPECTATIONS BY COUNTRY, POST-1996	278
FIGURE 83: ITALY, INFLATION EXPECTATIONS VS ACTUAL OUTTURNS	279
FIGURE 84: EXPECTED INFLATION, RESPONSE SHARES BY EUZ COUNTRY	281
FIGURE 85: EXPECTED INFLATION, RESPONSE SHARES, SWEDEN AND UK	282
FIGURE 86: ABS. DIFF. (EXPECTATIONS – PERCEPTIONS) OF INFLATION	283
FIGURE 87: ABS. DIFF. (EXPECT. – PERCEPTIONS) OF CONSTANT PRICES	284
FIGURE 88: ABS. DIFF. (EXPECTATIONS – PERCEPTIONS) OF DEFLATION	284
FIGURE 89: UK, ITA, ABS. DIFF. (EXPECTATIONS – PERCEPTIONS) OF INFLATION	285
FIGURE 90: INFLATION PERCEPTIONS EqCM, GRAPHICAL ANALYSIS	504
FIGURE 91: INFLATION PERCEPTIONS MODEL, STABILITY TESTS	504
FIGURE 92: PERCEPTIONS-ADJUSTED INFLATION vs HICP INFLATION	299
FIGURE 93: OLD / NEW EXPINFLPBIS AGAINST QUANTITATIVE EXPECTATIONS	303
FIGURE 94: ADP HYPOTHETICAL LONG-RUN PHILLIPS CURVE	306
FIGURE 95: AGGREGATE EXPECTATIONS ERRORS AND INFLATION	505
FIGURE 96: GERBERDING MODEL, FIT AND STANDARDISED RESIDUALS	332
FIGURE 97: FIT FOR GERBERDING’S RULE-OF-THUMB MODEL, 1991Q1-1999Q4	507
FIGURE 98: STANDARDISED RESIDUALS FOR THE MODEL, 1991Q1-1999Q4	507
FIGURE 99: RULE-OF-THUMB MODEL, FITTED & ACTUAL VALUES, FS	508
FIGURE 100: RULE-OF-THUMB MODEL, STANDARDISED RESIDUALS, FS	508
FIGURE 101: HEINEMANN AND ULLRICH (2004) RECURSIVE COEFFICIENTS	342
FIGURE 102: CONSENSUS FORECASTS (AT TIME REFERRED) AND INFLATION	351
FIGURE 103: INDEX OF NEWS COVERAGE INTENSITY, 1988M10-2005M12	369
FIGURE 104: VOLCKER DISINFLATION, “STICKY INFORMATION” PREDICTIONS	389
FIGURE 105: UK CPI INFLATION, ALL ITEMS; RECENT HISTORY	391
FIGURE 106: UK EXPECTATIONS, ERM DISINFLATION, ENDPOINTS	392
FIGURE 107: UK EXPECTATIONS DISTRIBUTION & THE ERM DISINFLATION, DETAIL	393
FIGURE 108: US INFLATION EXPECTATIONS, VOLCKER DISINFLATION	395
FIGURE 109: VAN LELYVELD’S DISTRIBUTION OF EXPECTATIONS (1995M1)	399

FIGURE 110: VAN LELYVELD INDEX OF DOUBLE-PEAKEDNESS THROUGH TIME.....	400
FIGURE 111: EXAMPLE OF NORMAL EXPECTATIONS DISTRIBUTION.....	402
FIGURE 112: GALLUP EXPECTATIONS DISTRIBUTION, 1983M12	406
FIGURE 113: GALLUP EXPECTATIONS DISTRIBUTION, 1984M10 AND 1985M6	407
FIGURE 114: GALLUP (1984M10 & 1985M6), NORMAL AND KERNEL DENSITY	408
FIGURE 115: GALLUP EXPECTATIONS DATA, INTERQUANTILE RANGE	410
FIGURE 116: US, MICHIGAN SURVEY, INTERQUANTILE RANGE	411
FIGURE 117: GALLUP INTERQUANTILE RANGE AND CPI INFLATION	509
FIGURE 118: GALLUP INTERQUANTILE RANGE AND CPI INFLATION, SCATTER.....	510
FIGURE 119: GALLUP INTERQUANTILE RANGE AND CHANGES IN INFLATION	511
FIGURE 120: GALLUP INTERQUANTILE RANGE AND THE OUTPUT GAP	512
FIGURE 121: DISTRIBUTION OF PRICE CHANGES ACROSS HICP COMPONENTS	513
FIGURE 122: DISPERSION OF GALLUP EXPECTATIONS & HICP INFLATION.....	514
FIGURE 123: CP AND GALLUP INTERQUANTILE RANGES, UK.....	417
FIGURE 124: GALLUP EXPECTATIONS DISTRIBUTION, 1992M12	421
FIGURE 125: TIME SERIES OF ‘LARGE’ GALLUP EXPECTATIONS RANGES	422
FIGURE 126: BOE/NOP EXPECTATIONS AT TIME SURVEYED VS PRICE INDICES.....	425
FIGURE 127: BOE/NOP MEDIAN PERCEPTIONS VS PRICE INDICES	426
FIGURE 128: AVG. ABSOLUTE DIFFERENCE, EXPECTATIONS - PERCEPTIONS.....	427
FIGURE 129: ‘DON’T KNOW’ ANSWERS, GALLUP AND BOE/NOP SURVEYS.....	428
FIGURE 130: ‘DON’T KNOW’ ANSWERS, EC CONSUMER SURVEY	429
FIGURE 131: EC SURVEY ‘DON’T KNOW’ ANSWERS & UK GENERAL ELECTIONS.....	433
FIGURE 132: QUANTITATIVE ‘DON’T KNOW’ ANSWERS & UK ELECTIONS	436
FIGURE 133: US SRC “DON’T KNOW” ANSWERS & US PRESIDENTIAL ELECTIONS	437
FIGURE 134: EC CONSUMER SURVEY, “DON’T KNOW” ANSWERS & ELECTIONS	439
FIGURE 135: US, MICHIGAN SURVEY; EXPECTATIONS DISTRIBUTION, 1978-2004	441
FIGURE 136: DISTRIBUTION OF INFLATION EXPECTATIONS, US & SWEDEN	443
FIGURE 137: SWEDISH INFLATION EXPECTATIONS ACROSS REGIMES.....	444
FIGURE 138: AVERAGED GALLUP EXPECTATIONS DISTRIBUTION, 1980s & 1990s.....	446
FIGURE 139: AVERAGE BOE/NOP EXPECTATIONS DISTRIBUTION, 1999Q4-2005Q4.....	447
FIGURE 140: UK INFLATION EXPECTATIONS, GALLUP DATA.....	515
FIGURE 141: UK INFLATION EXPECTATIONS, BOE/NOP DATA.....	516
FIGURE 142: BOE/NOP PERCEPTIONS/EXPECTATION BY INFLATION RANGE.....	454
FIGURE 143: US, SWEDEN AND UK, 1991-1999 CPI INFLATION	456

INTRODUCTION

Inflation expectations are one of the fundamental building blocks of modern macroeconomics. Yet their empirical properties remain under-studied, and their use in support of other empirical applications is almost inexistent, despite strong theoretical arguments attributing them a critical role. The present study aims to make a contribution towards changing this unsatisfactory situation. Our principal focus will be on consumer inflation expectations in a set of European Union countries (Belgium, France, Spain, Italy, Germany, the United Kingdom, the Netherlands and Sweden).

The most serious obstacle to a direct study and use of this concept is of course measurement-related. Indeed, how does one *measure* inflation expectations? This is the central theme of **Chapter 1**. The data source considered is a set of *qualitative* surveys managed by the European Commission. The chapter carries out an empirical contrast between most of the quantification methodologies used in the literature for this type of survey data in the past, and introduces two variants thereof. The analysis concludes that the two quantification methodologies proposed here are the most successful ones, with one of the two being a clear leader.

Chapter 2 turns to an empirical application in which the availability of the inflation expectations measure developed in Chapter 1 proves crucial. It has been suggested that since membership of a monetary union implies the relinquishing of national monetary sovereignty, it may be the case that the labour market responds *endogenously* to this new

situation by increasing its flexibility. This is the null hypothesis considered in this chapter. A new specification for wage equations which carefully controls for both inflation expectations and institutional change is developed. The possibility of EMS and EMU effects on wage flexibility is then checked by a search for appropriate structural breaks for each country in our sample. The wage equations prove to be quite stable overall and provide no evidence of an increase in flexibility, although the close linkage found between wages and productivity may arguably be related to EMU to some extent in certain countries.

Chapter 3 considers the empirical properties of inflation expectations and perceptions *themselves*. Their behaviour in the context of the euro's introduction is analysed in detail, and a case is made against the rationality of *both* inflation perceptions and expectations. The chapter shows that commonly used models of expectations conceptualising these as a simple combination between "rational" and "non-rational" components are misleading. On the other hand, some evidence is found to support a transmission of expert forecasts into consumer expectations through the news media in the case of the UK. Finally, the distributional properties of inflation expectations are considered. It is argued that consumers update their expectations gradually in a way that is mostly adaptive in the face of changing inflation. Accordingly, it is inflation *regimes* that affect the distribution of expectations rather than the technical details of monetary policy implementation. Evidence is presented to support the idea that consumers form expectations qualitatively, to some extent.

CHAPTER I:

Finding the Optimal Method of Quantifying Inflation Expectations on the Basis of Qualitative Survey Data

1 INTRODUCTION

The importance of expectations is one of the central tenets of modern theoretical economics¹. For instance, expectations played a key role in the economic case underpinning the move towards central bank independence (Forder, 2000); they are recognised as essential in the transmission mechanism of monetary policy and are kept under watch as part of modern inflation targeting frameworks². For empirical economists, this has generated the fundamental challenge of seeking ways to measure these “animal spirits”, to borrow the words of Keynes. This chapter will focus in particular on the measurement of inflation expectations.

It has become clear that simply assuming rational expectations is an indefensible proposition in empirical work. Indeed, the hypothesis has been rejected by the data both directly (see *inter alia* Roberts, 1997; Bakhshi and Yates, 1998; Lyziak, 2003) and indirectly (Ericsson and Irons, 1995)³. This should have come as no surprise given the fundamentally non-stationary nature of the economy (Clements and Hendry, 1999). Simple ‘rule-of-thumb’ mechanisms for expectations are similarly unsupported by the evidence (e.g. Roberts, 1997; Chapter 3, Section 4.4 of this study).

¹ For a historical overview, see Evans and Honkapohja (2001).

² For instance, the Bank of England considers both survey and market-based measures of inflation expectations in its Inflation Reports, the former being derived from the Bank of England/NOP survey. See Bank of England (2005), for instance, and Chapter 3 below, Section 6.6.2.

³ Further discussion and evidence is provided in Chapter 3.

The failure of these and other formulations of the mechanism of expectations generation⁴ has made the development of direct measures of the concept particularly valuable. This can be done in several ways. A first approach consists in inferring expectations from the prices of financial instruments (e.g. Bank of Canada, 1998). Indeed, comparing the prices of nominal and index-linked bonds with identical characteristics (risk, maturity and liquidity) can yield a measure of expected inflation – expectations which have been acted upon by economic agents. Unfortunately, the difficulty of finding directly comparable assets means that usually additional assumptions are required (Berk, 1999). Furthermore, the expectations of actors in financial markets may only be representative of the agents operating in those sectors⁵, and may involve additional complications (e.g. herding behaviour, time-varying risk premia, regulatory requirements⁶).

A second approach will be followed here, and makes use of survey data. We will in particular be interested in the inflation perceptions and expectations of consumers as a whole in a set of countries belonging to the European Union. Such measures can be thought to be particularly useful as part of the toolkit of economic “compasses” deployed in inflation targeting frameworks or in the investigation of wage dynamics for instance. Given the qualitative nature of these data, the task of the present chapter will be to find the optimal method of quantification to make these data useable.

⁴ We note that there is a vast empirical literature making use of IV or GMM methods to construct some form of model-based expectations (e.g. Kremers, 1990).

⁵ This is likely given the findings in Bryan and Venkatu (2001) and Carroll (2003).

⁶ For instance, in the UK, regulatory requirements force institutional investors such as pension funds to invest in index-linked bonds to some extent, thereby artificially depressing their yield. Expectations computed on this basis would exaggerate the ‘true’ underlying inflation expectations.

Section 2 presents the qualitative data under study here, collected as part of the European Commission's Consumer Survey. Section 3 then presents the competing quantification techniques for this type of data. A first comparison is carried out in Section 4 by using descriptive statistics typically considered in the literature to make an assessment of predictive ability. We argue however that a more thorough comparison is required. This will be done in two steps. First, we use quantitative data for the UK and Sweden (presented in Section 5) as benchmarks for the output of quantification techniques (Section 6). These techniques are then contrasted in an important application (Section 7), namely the estimation of wage equations for eight European countries: France, Spain, Belgium, Italy, Germany, the UK, the Netherlands and Sweden. Section 8 concludes.

2 QUALITATIVE DATA: THE EC CONSUMER SURVEY

We use the European Commission's Consumer Survey to derive a numerical measure of inflation expectations. It is part of the Joint Harmonised EU Programme of Business and Consumer Surveys, launched in 1962. The survey samples about 33,000 consumers every month at present (DG ECFIN, 2004). Participants in the survey are asked the following questions:

TABLE 1: EC CONSUMER SURVEY, QUESTIONS 5 AND 6

Q5. How do you think that consumer prices have developed over the last 12 months? They have...	Q6. By comparison with the past 12 months, how do you expect that consumer prices will develop in the next 12 months? They will ...
1 risen a lot	1 increase more rapidly
2 risen moderately	2 increase at the same rate
3 risen slightly	3 increase at a slower rate
4 stayed about the same	4 stay about the same
5 fallen	5 fall
9 don't know.	9 don't know.

Unfortunately, the survey has not always been fully harmonised to the wording presented in Table 1 in the case of two countries: France and Spain (Gerberding, 2001). In France, prior to harmonisation in January 2004, question 5 referred to the period of the previous six months, while question 6 referred to the “coming months” (INSEE, 2004). This latter indeterminacy in the horizon to be considered by the respondents can potentially be damaging in terms of measurement error. As in Gerberding (2001), we assume that the horizon considered is twelve months and quantify the French data in the same ways used for the other countries, although this caveat should be borne in mind when interpreting the results.

In Spain, options 1-3 of question 6 respectively read [more sharply / rather sharply / more slowly]⁷. Given that options 1 and 3 are essentially equivalent under both wording variations, we will assume that the difference in wording for option 2 is insignificant, being appropriately “sandwiched” in such a way that it can be regarded as immaterial. Of

⁷ Options 1-3 of question 5 are also different and read [very sharply / rather sharply/ slightly].

course, to the extent that this assumption is flawed Spanish inflation expectations could suffer from measurement error⁸.

The EC Consumer Survey itself dates back to 1972. Data availability and frequency differs over time and across countries. Alas, much of the early data collected by this programme has been lost (Papadia and Basano, 1981; Reckwerth, 1997). The European Commission can now only make available data from January 1985 onwards at the earliest. After a major data recovery effort – contacting national statistical institutes, private research firms and individual researchers – we have managed to achieve the sample coverage detailed in Table 2. This has left us with the most complete dataset to be used to date for consumer inflation expectations in Europe.

TABLE 2: EC CONSUMER SURVEY, SAMPLE COVERAGE BY COUNTRY

Country	Thrice yearly from...	Quarterly from...	Monthly data available from...
Belgium	N/A	January 1985	October 1985
Germany⁹	N/A	N/A	January 1980
Spain	N/A	N/A	June 1986
France	N/A	January 1985	April 1986
Italy	January 1973	N/A	January 1982
Netherlands	October 1973	January 1984	April 1986
Sweden¹⁰	N/A	N/A	January 1996
United Kingdom	N/A	N/A	November 1981

⁸ Gerberding (2001) proposes an alternative solution, the cost of which is the assumption that inflation perceptions are ‘correct’, in that they correspond to actual CPI inflation. This may be rather demanding assumption – see Chapter 3, Section 2.3.

⁹ Data on inflation perceptions (‘question 5’) is only available from October 1980. Expectations only refer to Western Germany prior to 1991. This should however not affect our results materially, as the German wage equation in Section 7.3.8 is modelled post-1991 and the CPI data used in the tests for predictive success in Section 6.1 also refers to West Germany alone prior to 1991.

¹⁰ The European Commission offers data from October 1995. However, a number of inconsistencies in the initial observations with the dataset sourced from the Konjunkturinstitutet made us prefer the latter for now after a private communication with Klas-Göran Warginger on the issue.

3 OVERVIEW OF QUANTIFICATION METHODOLOGIES

We will now present the different quantification methodologies considered in this chapter. Their output will be plotted for illustrative purposes in the case of the United Kingdom.

3.1 The Carlson-Parkin or Probability Approach

The probability approach has a long track record of literature¹¹ (Theil, 1952; Carlson and Parkin, 1975). It is also known as the Carlson-Parkin method after the seminal contribution of the two authors (henceforth CP). Batchelor and Orr (1988) extended the originally trichotomous approach to the case of pentachomous data, a methodology also adopted by Reckwerth (1997) and Berk (1999, 2000). Given the pentachomous structure of the EC Consumer Survey's question 6, it is this CP extension that we implement here.

The CP method assumes that respondents at time $t-12$ (months) have formed an expectation Π_{it}^{exp} about inflation in the coming 12 months when answering the survey. This is based on a subjective probability distribution for each individual i over future inflation $f_i(\Pi_{it} | I_{i,t-12})$, conditional on the information set $I_{i,t-12}$ available at time $t-12$ ¹².

The aggregation of the subjective probability distributions is denoted $g(\Pi_t | \Omega_{t-12})$,

where $\Omega_{t-12} = \bigcup_{i=1}^N I_{i,t-12}$ - the union of the individual information sets. Quantification is

an exercise in recovering the mean of the aggregate distribution, Π_t^{exp} .

¹¹ A good overview is provided by Nardo (2003); we borrow some of her notation in what follows.

¹² It seems sensible to assume heterogeneity in subjective probability distributions – notably given the surprisingly different experiences of inflation found in the population (Crawford and Smith, 2002).

The shares of respondents in each survey category are then interpreted as maximum likelihood estimates of areas under the aggregate density function of inflation expectations, that is, as probabilities¹³.

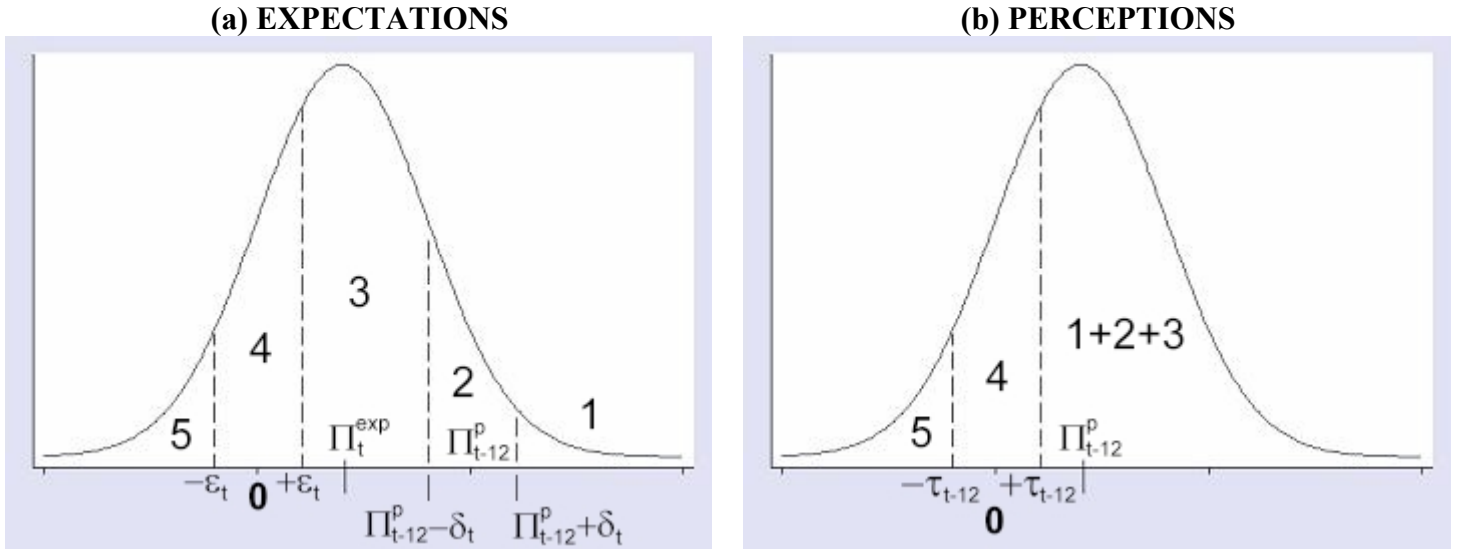
We then make three key assumptions, reflected in Figure 1(a):

1. There is a symmetric range of inflation values around zero (denoted ϵ_t) and around the perceived rate of inflation over the past twelve months Π_{t-12}^p (denoted δ_t) which the respondent cannot distinguish. This will lead the respondent to provide answers 4 and 2 respectively in their immediate vicinity¹⁴.
2. Second, as implied by the above notation, we assume that the indifference intervals are equal for all agents, but allow for time-variation.
3. The shape of the aggregate distribution function is known and corresponds to the normal distribution.

¹³ We will focus on response options 1-5 of question 6 (cf. Table 1). The share of respondents choosing the ‘don’t know’ option is allocated proportionally to the other categories (Visco, 1984).

¹⁴ Such indifference intervals can be given psychological groundings, and are termed “difference limens” in the specialised literature (Osgood, 1953).

FIGURE 1: AGGREGATE DENSITY FUNCTIONS: EXPECTATIONS, PERCEPTIONS



The shape of the assumed distribution has been a particularly contentious point in the literature. The normal distribution has a good justification from statistical theory, through some central limit theorem result. For instance, assuming independence of the expectations variables generated by the individual subjective distributions, the application of the Lindeberg-Lyapunov Central Limit Theorem would ensure that the aggregate distribution is normal (without the need to assume normality of the individual expectations).

It is nevertheless debatable whether the normal distribution is the most appropriate assumption to be made, and indeed there are good theoretical and empirical reasons to think otherwise from a descriptive standpoint. First, kurtosis has been found to be more marked than that of a normal distribution. For instance, the distribution may be more peaked at times when well-publicised events are generating inflation (Batchelor, 1982). Second, the distribution of inflation expectations is assumed to be symmetric, when one

would expect right-skewness in periods of rising prices and left-skewness in the reverse situation. There is some evidence that this is indeed the case (Foster and Gregory, 1977; Batchelor, 1982). However, the skewness is likely to depend on the magnitudes of inflation and its acceleration, which in recent European experience would be relatively moderate. Indeed, taking this into account appears to make little difference to the derived expectations series, as shown by Nielsen (2003a) on the same EC dataset, using the logistic, chi-square, central and noncentral t-distributions. This is consistent with earlier results in Berk (1999). We will therefore maintain the assumption of normality.

The CP method then proceeds by standardising key points under the distribution function:

$$Z_t^1 = \frac{\Pi_{t-12}^p + \delta_t - \Pi_t^{\text{exp}}}{\sigma_t} \quad (1)$$

$$Z_t^2 = \frac{\Pi_{t-12}^p - \delta_t - \Pi_t^{\text{exp}}}{\sigma_t} \quad (2)$$

$$Z_t^3 = \frac{\varepsilon_t - \Pi_t^{\text{exp}}}{\sigma_t} \quad (3)$$

$$Z_t^4 = \frac{-\varepsilon_t - \Pi_t^{\text{exp}}}{\sigma_t} \quad (4)$$

Having assumed the shape of the cumulative density function of $g(\Pi_t \mid \Omega_{t-12})$, which we will denote $G_{t-12}(\cdot)$, it is possible to derive estimates of the Zs from the response proportions of the survey. Indeed, labelling areas 1-5 in Figure 1(a) by RR, ER, WR, C and F respectively, we can write the following:

$$\begin{aligned} RR_t^{\text{exp}} &= \text{prob}(Z_t \geq Z_t^1 \mid \Omega_{t-12}) \\ &= 1 - G_{t-12}(Z_t^1) \end{aligned} \quad (5)$$

$$\begin{aligned} ER_t^{\text{exp}} &= \text{prob}(Z_t^2 \leq Z_t \leq Z_t^1 \mid \Omega_{t-12}) \\ &= G_{t-12}(Z_t^1) - G_{t-12}(Z_t^2) \end{aligned} \quad (6)$$

$$\begin{aligned} WR_t^{\text{exp}} &= \text{prob}(Z_t^3 \leq Z_t \leq Z_t^2 \mid \Omega_{t-12}) \\ &= G_{t-12}(Z_t^2) - G_{t-12}(Z_t^3) \end{aligned} \quad (7)$$

$$\begin{aligned} C_t^{\text{exp}} &= \text{prob}(Z_t^4 \leq Z_t \leq Z_t^3 \mid \Omega_{t-12}) \\ &= G_{t-12}(Z_t^3) - G_{t-12}(Z_t^4) \end{aligned} \quad (8)$$

$$\begin{aligned} F_t^{\text{exp}} &= \text{prob}(Z_t \leq Z_t^4 \mid \Omega_{t-12}) \\ &= G_{t-12}(Z_t^4) \end{aligned} \quad (9)$$

It is then straightforward to derive estimates of the Z s by substituting the left-hand side objects by their observed quantities – the sample response proportions. Finally, we rearrange equations (1) to (4) and obtain an expression for the expected inflation in terms of observables:

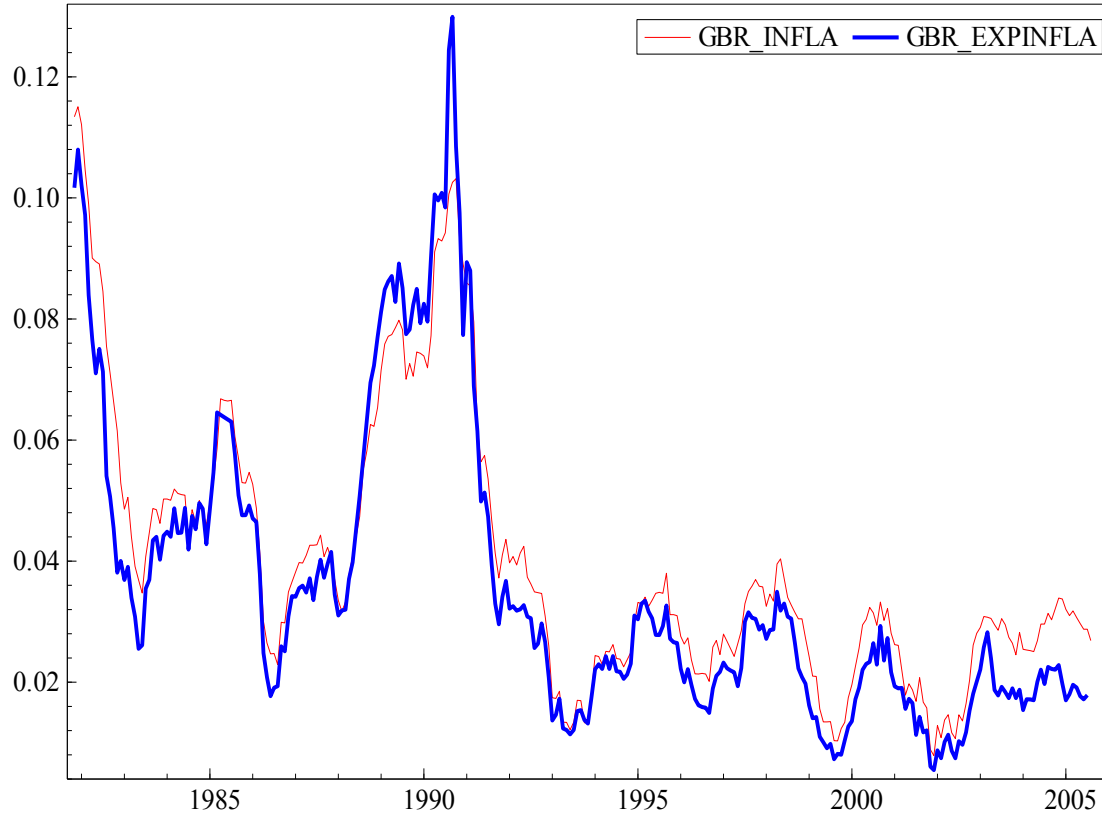
$$\Pi_t^{\text{exp}} = \Pi_{t-12}^p * \left[-\frac{Z_t^3 + Z_t^4}{Z_t^1 + Z_t^2 - Z_t^3 - Z_t^4} \right] \quad (10)$$

The only missing element in this expression is the perceived inflation rate Π_{t-12}^p . The following possibilities will be considered in this respect, thus generating alternative CP expectations series:

- 1. Assume that inflation is correctly perceived and use the officially published inflation rate to scale the expectations series (i.e. $\Pi_{t-12}^p = \Pi_{t-12}$).**

This results in Figure 2. Inflation expectations are plotted at the time they are *surveyed* rather than at the time they *refer* to (12 months later).

FIGURE 2: UK, CP EXPECTED INFLATION & INFLATION, 1981M11-2005M9



2. Trichotomous quantification of question 5 (on perceptions) to obtain Π_{t-12}^p

Whereas question 6 of the survey offered information relative to two “clear” quantitative points (zero inflation and perceived inflation), question 5 (on perceptions) offers only one (zero inflation, option 4), the other being ambiguous (“moderate inflation”; Table 1).

This approach considers that no additional information can be gained from categories 1-3 and pools these together, thereby coding the survey answers in a trichotomous manner¹⁵.

With reference to Figure 1(b) define:

$$W_{t-12}^1 = \frac{+\tau_{t-12} - \Pi_{t-12}^p}{\omega_{t-12}} \quad (11)$$

$$W_{t-12}^2 = \frac{-\tau_{t-12} - \Pi_{t-12}^p}{\omega_{t-12}} \quad (12)$$

The indifference thresholds are denoted by “ τ ”. Rearranging (11) and (12) we obtain:

$$\Pi_{t-12}^p = \left(\frac{W_{t-12}^2 + W_{t-12}^1}{W_{t-12}^2 - W_{t-12}^1} \right) * \tau_{t-12} \quad (13)$$

The distributional assumption and the response proportion data then allow us to derive the W s in the expression in a way analogous to equations (5) to (9).

The thresholds require additional assumptions to be obtained. Carlson and Parkin (1975) propose thresholds that are symmetric and constant across individuals and time. Further, assume that consumers’ inflation perceptions are unbiased. We then obtain:

$$\hat{\tau} = \frac{\sum_{t=1}^T \Pi_t}{\sum_{t=1}^T \left(\frac{W_t^2 + W_t^1}{W_t^2 - W_t^1} \right)} \quad (14)$$

This assumption of constant indifference thresholds may be problematic, especially in the face of large structural changes such as EMU. Indeed, it is shown in Section 2.1 of Chapter 3 that the introduction of the new currency was generally associated with a

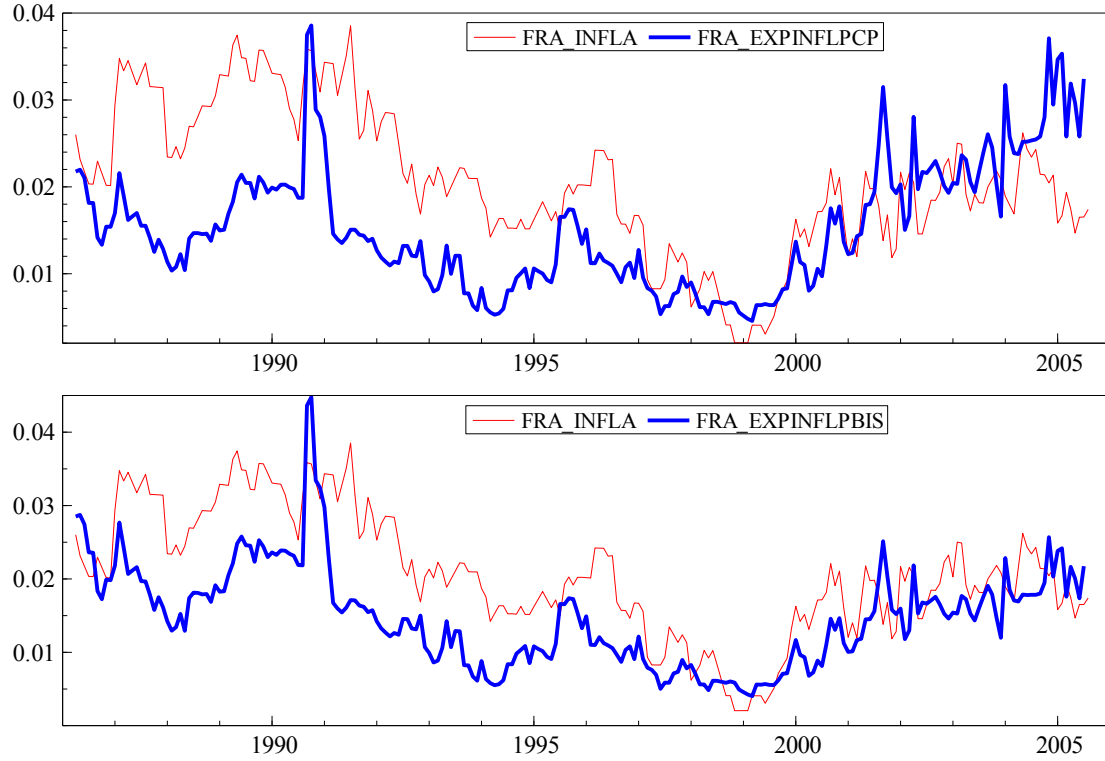
¹⁵ The ‘don’t know’ answers are, as before, proportionately assigned to the other response categories (Visco, 1984).

collapse in the proportion of respondents perceiving zero inflation in favour of that perceiving positive inflation. With constant thresholds, this forces a ‘flattening’ of the distribution to fit the data, in such a way that the mean perceptions (and consequently, expectations) derived are rather large. One interpretation of this is that the response thresholds have shifted inwards over the EMU period, so that it takes a smaller absolute value of perceived inflation or deflation for the respondents to call it as such¹⁶. To allow for this and other such changes in the threshold parameters without risking the imposition of excessive structure, we apply a criterion of unbiasedness for each half of the sample separately. The threshold values derived for the first half of the sample are assigned to the first observation, while the threshold values derived for the second half of the sample are assigned to the last observation; values in between are filled in by linear interpolation. This smooth transition aims to capture any trending behaviour of the thresholds, whilst preserving the main data features (with some attenuation). Further discussion will be provided in Section 2.1.2 of Chapter 3.

The results for the previous two methods (labelled ‘EXPINFLPCP’ and ‘EXPINFLPBIS’, respectively) are plotted in Figure 3 for France against actual inflation, as the differences between them are not readily visible on UK data. Here again, expectations are plotted at the time they were *surveyed*.

¹⁶ Further discussion is provided in Section 2.1.2 of Chapter 3.

FIGURE 3: FRA, CP TRICHOTOMOUS QUANTIFICATION, 1986M1-2005M9



3. Pentachomous quantification of question 5 (perceptions) to obtain Π_{t-12}^p

This consists in regarding option 3 in Table 1 (“risen moderately”) as a quantitative indication – there is at any time a given inflation rate which people consider to be “moderate”, Π_t^m . Assuming we know this rate, we can then proceed to quantifying perceptions in exactly the same manner as we quantified expectations.

Thus, we assume that there is a range of values around zero and around moderate inflation Π_t^m which are indistinguishable to respondents, and are enclosed by indifference thresholds. The relevant picture is then equivalent to that in Figure 1(a), applied to the case of perceptions, and we define points $A_t^1 \rightarrow A_t^4$ analogously to points $Z_t^1 \rightarrow Z_t^4$ earlier (see Eqs. (1)-(4)). We obtain the expression:

$$\Pi_{t-12}^p = \Pi_t^m * \left[-\frac{A_t^3 + A_t^4}{A_t^1 + A_t^2 - A_t^3 - A_t^4} \right] \quad (15)$$

The only missing item is then the moderate rate of inflation Π_t^m . We consider three proxies for this quantity¹⁷; options (a) and (c) follow the literature; option (b) is newly introduced here:

- a. The average value of the actual inflation rate over the whole sample¹⁸.**
- b. A linear interpolation between the average value of inflation over the first half of the sample and that over the second half of the sample after those values are assigned to the first and last months in the sample, respectively.**
- c. The running mean of inflation from the beginning of the sample to the point where expectations are surveyed¹⁹.**

¹⁷ Batchelor and Orr (1988) – having the benefit of a subsample of numeric survey answers as well as qualitative data – solve this problem through the minimisation of a loss function that penalises both deviations from unbiasedness and certain forecasting errors over the quantified data. One may argue as to the generality of the approach and the assumptions behind the specified loss function, but in any case it is an approach that cannot be followed here due to data limitations.

¹⁸ This approach is followed by Reckwerth (1997).

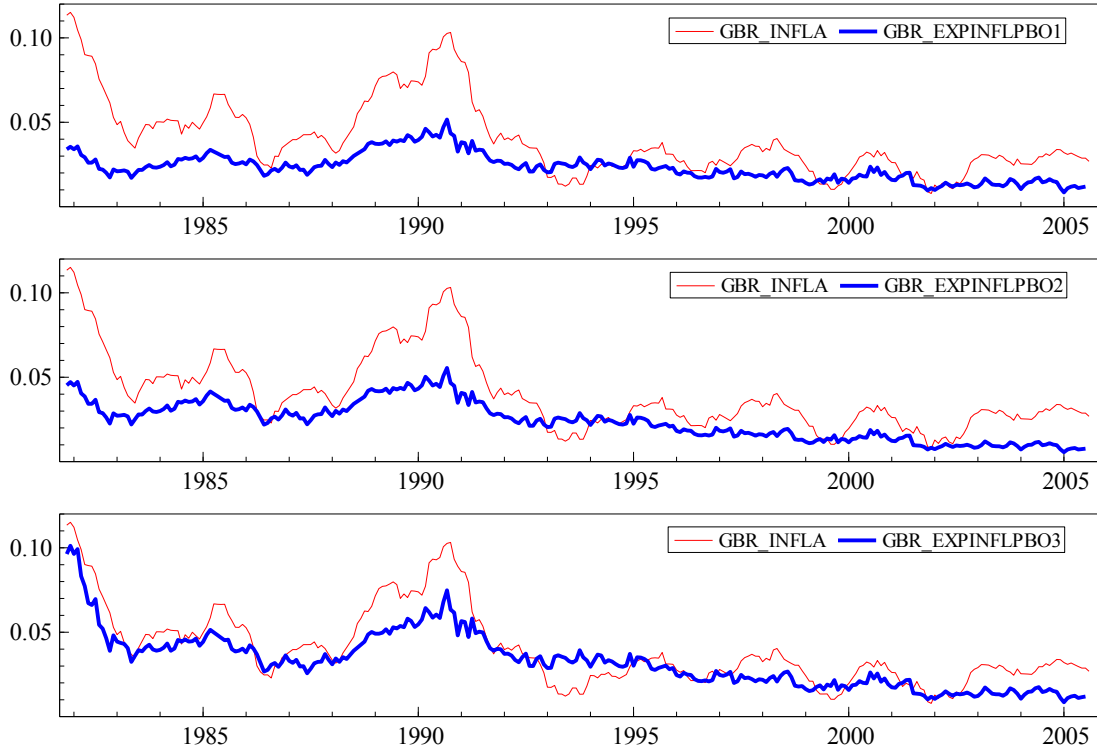
Although option (a) may be sensible in periods where mean inflation exhibits no trend, the definition of what is “moderate” has very likely changed in many European countries that saw large declines in inflation over the past two decades. This is indeed the case for the UK, and as can be seen from Figure 4 (top), the result is an expected inflation series that appears to be suspiciously ‘flat’ relative to the evolution of inflation.

To take this observation into account without imposing excessive structure by assumption, we propose option (b) to capture the potentially trending behaviour of the moderate rate without being any more ad hoc than the alternatives.

Option (c) solves the problem in a different way, although there are conceptual doubts as to whether the moderate rate can be equated with the running mean of inflation: not only is the variance of the former likely to be lower, but it might be considered that – except in situations of low and stable inflation – agents would have a concept of the moderate rate lower than actual and past inflation. The conceptual advantage of option (c) relative to the other methods is that it only makes use of information that can be known to the agents at any time t . However, the decision to initialise the running mean at the start of the sample is ad hoc and open to criticism.

¹⁹ This is suggested in Nielsen (2003a).

FIGURE 4: UK, CP PENTACHOMOUS QUANTIFICATION, 1981M11-2005M9



NOTE: Expectations are plotted *at the time they are surveyed*, but refer to the coming year.

3.2 The Regression Approach

The ‘*regression approach*’ originates in Anderson (1952) and was greatly developed by Pesaran (1984, 1985), so that it often takes the name of the latter author²⁰.

Relate inflation π_t over the past 12 months to a weighted combination of inflation perceptions from respondents having experienced increasing (superscript +) or falling (superscript –) prices:

$$\pi_t = \sum w_i^+ \pi_{t,i}^+ + \sum w_i^- \pi_{t,i}^- \quad (16)$$

²⁰ The object of interest in Pesaran’s work is British wholesale manufactured goods inflation. We adapt the presentation of his model to the present context.

Pesaran (1984) assumes that during inflationary periods, there exists an asymmetrical relationship between the rate of change of individually experienced prices and overall inflation, depending on the direction of change reported:

$$\begin{cases} \pi_{i,t}^+ = \alpha + \lambda\pi_t + v_{i,t}^+ & \alpha \geq 0, 0 \leq \lambda \leq 1 \\ \pi_{i,t}^- = -\beta + v_{i,t}^- & \beta \geq 0 \end{cases} \quad (17)$$

Weighing all respondents equally and substituting the relations in (17) into (16) suggests the following nonlinear regression:

$$\pi_t = \frac{\alpha R_t - \beta F_t}{1 - \lambda R_t} + u_t \quad (18)$$

Where R_t and F_t denote the percentages of agents reporting price rises or falls in their answer to the *perceptions* question, respectively. Once the coefficients from (18) have been recovered, it is possible to apply them to the survey proportions relating to *expectations* this time, thereby deriving a measure of inflation expectations:

$$\hat{\pi}_{t,t+12}^{\text{exp}} = \frac{\hat{\alpha} R_t^{\text{exp}} - \hat{\beta} F_t^{\text{exp}}}{1 - \hat{\lambda} R_t^{\text{exp}}} \quad (19)$$

The key assumption is thus that the estimated relationship between survey data and inflation holds not only for realisations, but also for expectations. This can be thought to be a strong assumption to make a priori. Batchelor and Orr (1988) note that Pesaran

effectively assumes that the response thresholds relating to past and future inflation are the same, which contradicts signal-extraction theory.

It is important to understand the spirit in which this procedure is carried out, and therefore what its ambitions are. Pesaran (1987) emphasises that a regression like (18) “*is not a causal explanation of price changes but simply identifies the relationship between two different sources of information (namely official statistics and survey results), and serves as a ‘yardstick’ by means of which categorical responses concerning the direction of future changes in prices can be converted into quantitative measures*”²¹.

We estimated (18) for our UK data by nonlinear least squares using the TSP module of GiveWin2 and obtained the results in column(1) of Table 3²². The derived expectations series is plotted in Figure 5 (top).

As can be verified straightforwardly, the equation fails the basic congruence tests provided by exhibiting significant heteroscedasticity and residual autocorrelation. One possible attitude to this finding is to argue that since the goal of this exercise was not one of building a statistical model, but rather that of exploiting a ‘yardstick’, we are less concerned than normally about such failures. Given that the aim of this study is the empirical comparison of methodologies, we will accept this argument and move forward.

²¹ See Pesaran (1987), page 211.

²² The initial values for the procedure were taken from Smith and McAleer (1995)

Pesaran (1984) further suggested correcting for the residual autocorrelation by imposing an AR structure on the error term. However as Hendry and Doornik (2001) note, dozens of mis-specifications in econometrics can generate that result, making the assumption that autocorrelated residuals entail autoregressive errors an enormous non-sequitur. We nevertheless estimated for interest a model with an AR(2) correction, as in Pesaran (1984) and report the results in Table 3, column(3):

$$\pi_t = \frac{\alpha R_t - \beta F_t + \rho_1[(1 - \lambda R_{t-1})\pi_{t-1} - \alpha R_{t-1} + \beta F_{t-1}] + \rho_2[(1 - \lambda R_{t-2})\pi_{t-2} - \alpha R_{t-2} + \beta F_{t-2}]}{1 - \lambda R_t} + \varepsilon_t \quad (20)$$

Judging from the value of R^2 and the plot of the derived expectations series (Figure 5, bottom), this might look suspiciously like a case of spurious regression (note also that the coefficients on the autoregressive structure for inflation add up to one). However, there is no apparent autocorrelation problem according to the DW statistic. In any event, it is transparent that the survey data in this regression are uninformative relative to lags of inflation, which undermines the use of this modification of the approach.

This prompts the question of whether inflation should be treated as an I(1) variable. A DFGLS test for the UK over the period 1981m11-2005m7 would a priori seem to suggest so, as the null of nonstationarity cannot be rejected - even at the 10% level²³. However, it is well known that such tests are invalidated by the presence of breaks, and that series affected by them often “look” like I(2) series in levels. We will follow Hendry (2001) –

²³ This is a modified Dickey-Fuller t-test for a unit root in which the series has been transformed by a generalised least-squares regression. Detailed test results are available from the author.

who models UK inflation for the considerably more challenging period 1875-1991 – in assuming that (a) the price data are I(1) with superimposed breaks and (b) measurements thereof have I(1) deviations from their theoretical counterparts. This implies that inflation can be treated as an I(0) series with breaks, but measured with an I(0) error²⁴. This seems theoretically sensible and is further justified here by the application of the unit root tests of Clemente et al. (1998) – which adjust for the presence of structural breaks – and are able to reject the null of non-stationarity²⁵.

Smith and McAleer (1995) propose to modify Pesaran's approach by specifying the following symmetric response model instead of (17):

$$\begin{cases} \pi_{t,i}^+ = \alpha + \lambda \pi_t + v_{i,t}^+ & \alpha \geq 0, 0 \leq \lambda \leq 1 \\ \pi_{t,i}^- = -\beta + \lambda_2 \pi_t + v_{i,t}^- & \beta \geq 0 \end{cases} \quad (21)$$

Which leads to the regression:

$$\pi_t = \frac{\alpha R_t - \beta F_t}{1 - \lambda R_t + \lambda_2 F_t} + u_t \quad (22)$$

²⁴ This approach is also successfully adopted by Bowdler and Jansen (2004) for the Eurozone over 1981-2000 and implicitly by Sekine (2001) for Japan in the period 1971-1998.

²⁵ As noted by Baum (2004) such a finding suggests misspecification for ADF-type tests. We notably implemented the version of the Clemente-Montañés-Reyes test that allows for the control of two innovational outliers. The optimal breakpoints for the UK on this basis would be 1988m6 and 1991m1, yielding a t-statistic on the (rho-1) coefficient of -5.896, which exceeds the 5% critical value of -5.490.

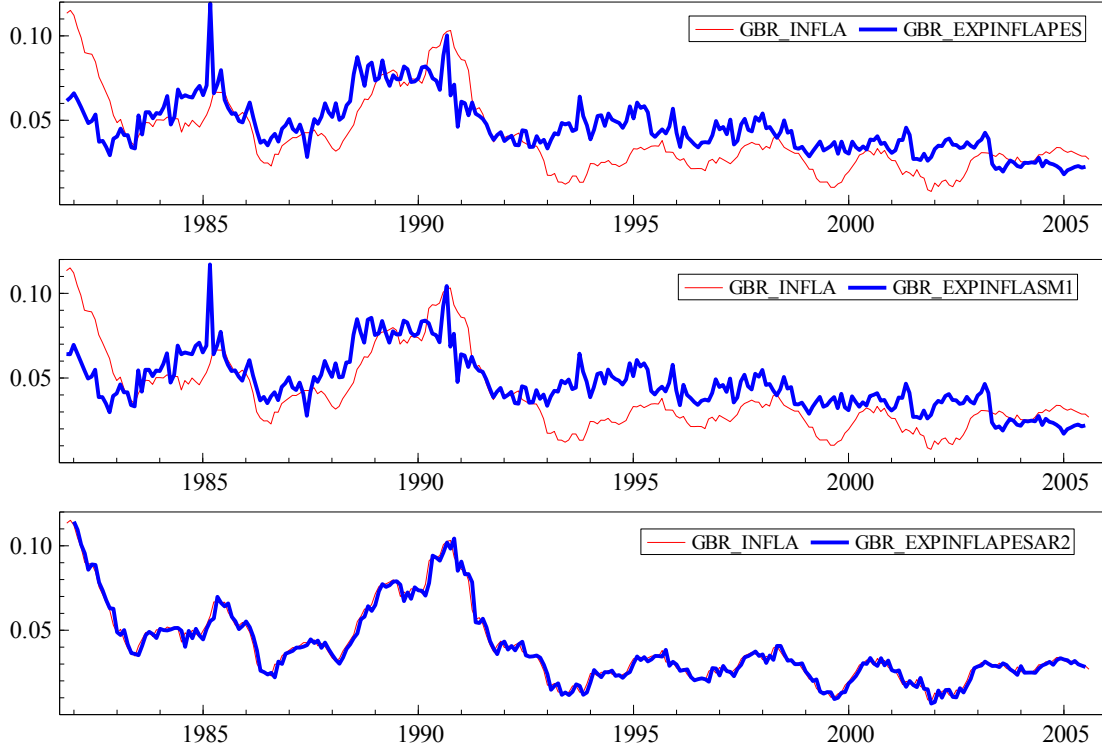
Column(2) of Table 3 makes clear that this is a very minor modification empirically, the additional coefficient being insignificant, and the others staying close to their earlier values (β excepted). The resulting expectations are plotted in Figure 5 (middle).

TABLE 3: PESARAN REGRESSION FOR THE UK, NLS, 1981m11-2005m7

Coefficient	Standard Pesaran Model [Eq. (18)]	Smith & McAleer Extension [Eq.(22)]	Pesaran AR(2) Correction [Eq.(20)]
α	0.0101 *** (0.0005)	0.0098 *** (0.0005)	-0.0035 (0.0097)
β	-0.0297 *** (0.0079)	-0.0124 (0.0152)	-0.0001 (0.0151)
λ	0.9504 *** (0.0071)	0.9501 *** (0.0069)	0.4345 ** (0.1776)
λ_2	-	0.8073 (0.5575)	-
ρ_1	-	-	1.3052 *** (0.0580)
ρ_2	-	-	-0.3148 *** (0.0577)
R^2	0.787	0.788	0.978
Eq. Std. Error	0.0104	0.010385	0.3282E-02
LM Het. Test	28.53 [.000]	27.55 [.000]	12.42 [.000]
Durbin-Watson	0.2782 [.000,.000]	0.2830 [.000,.000]	2.050 [.570,.746]

Notes: *, ** and *** denote significance at the 10%, 5% and 1% level respectively. Coefficient standard errors in parenthesis. The P-values for the test statistics are given in brackets. For DW, these relate to Durbin's h and Durbin's alternative statistic.

FIGURE 5: ALTERNATIVE EXPECTATIONS SERIES, PESARAN TRADITION



Smith and McAleer further propose to write parameters α and β as weighted averages of past π_t values, turning the equivalent of equation (18) into a dynamic nonlinear regression model.

$$\begin{cases} \alpha_t = \alpha + \lambda_1 \pi_t + \gamma_{11} \pi_{t-1} + \dots + \gamma_{1m} \pi_{t-m} \\ \beta_t = \beta + \lambda_2 \pi_t + \gamma_{21} \pi_{t-1} + \dots + \gamma_{2n} \pi_{t-n} \end{cases} \quad (23)$$

$$\pi_t = \frac{\alpha R_t - \beta F_t + R_t \sum_{j=1}^m \gamma_{1j} \pi_{t-j} - F_t \sum_{j=1}^n \gamma_{2j} \pi_{t-j}}{1 - \lambda_1 R_t + \lambda_2 F_t} + v_t \quad (24)$$

The results of this approach are presented in Table 4²⁶. The implausible coefficient values, very high R^2 and a ratio of γ_{21} and λ_2 which is close to one indicate that we have (predictably) collided with the same problem that occurred above with the AR correction. This is clearly a very marked case of overfitting; “rescuing” the model would likely have to involve further assumptions, in the form of constraints on the γ coefficients.

TABLE 4: SMITH & McALEER REGRESSION, UK, NLS, 1981m11-2005m7

	α	β	λ_1	λ_2	γ_{11}	γ_{12}	γ_{13}	γ_{14}	γ_{15}	γ_{21}	γ_{22}	γ_{23}	γ_{24}	γ_{25}
Value	0.01 ***	-0.09	0.67 ***	-186 ***	0.61	-0.55 **	0.12	0.59 **	-0.53 ***	219 ***	-2.53	-17.8	-31.7	15.7

$R^2 = .981871$ LM het. test = 8.26151 [.004] DW = 2.01578 [.249,.828] Eq. Std. Err = .309248E-02

Notes: *, ** and *** denote significance at the 10%, 5% and 1% level respectively. The P-values for the test statistics are given in brackets. For DW, these relate to Durbin’s h and Durbin’s alternative statistic.

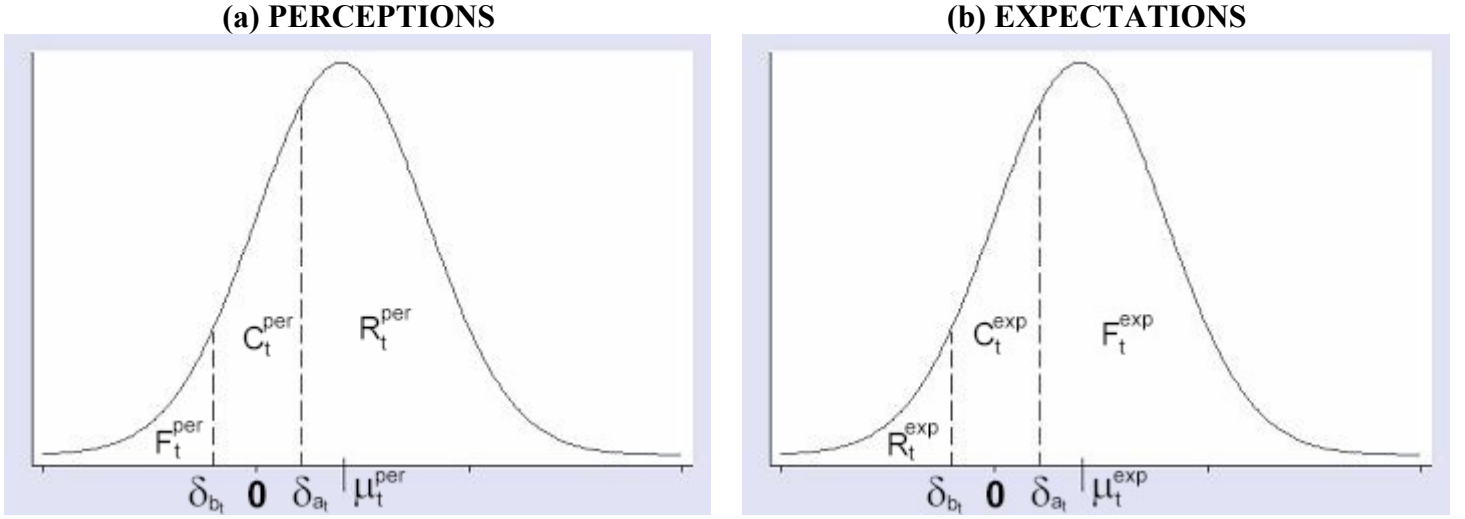
Given the results presented in this section, we choose to retain only the original Pesaran method for the comparisons that are carried out later in this study.

3.3 Models with stochastic time-varying parameters

Seitz (1988) presented a quantification methodology that permitted the relaxation of a crucial assumption in the original CP paper, namely the time-invariance of the threshold parameters. Consider the distribution of perceptions in Figure 6(a), split by whether responses have indicated (a rise/constancy/a fall) in prices:

²⁶ Note that even in the Smith and McAleer (1995) paper this regression is not congruent, exhibiting significant autocorrelation (in fact, almost significant at the 1% level).

FIGURE 6: SEITZ METHOD; PERCEPTIONS & EXPECTATIONS DISTRIBUTIONS



Let $W_1 = \frac{\delta_{a,t} - \mu_{per,t}}{\sigma_t}$ and $W_2 = \frac{\delta_{b,t} - \mu_{per,t}}{\sigma_t}$; from these two identities we obtain:

$$\mu_{per,t} = \frac{\delta_{a,t}W_{2,t} - \delta_{b,t}W_{1,t}}{W_{2,t} - W_{1,t}} \quad (25)$$

$W_{2,t}$ and $W_{1,t}$ can be found from the assumption of normality and the knowledge of the response proportions. Then, define:

$$x_{1,t} = \frac{W_{2,t}}{W_{2,t} - W_{1,t}} \text{ and } x_{2,t} = -\frac{W_{1,t}}{W_{2,t} - W_{1,t}}$$

This allows us to rewrite (25) as:

$$\begin{aligned} \mu_{per,t} &= \delta_{a,t}x_{1,t} + \delta_{b,t}x_{2,t} \\ &= x'_t\beta_t \end{aligned} \quad (26)$$

Where β_t denotes the vector of threshold coefficients at time t: $(\delta_{a,t}, \delta_{b,t})$.

Seitz proposed to deploy the Stochastic Parameter Variation model of Cooley and Prescott (1976) to estimate the indifference parameters (vector β_t). The model has the following structure:

$$\begin{cases} \beta_t = \beta_t^p + u_t \\ \beta_t^p = \beta_{t-1}^p + v_t \end{cases} \quad (27)$$

Thus, parameters are assumed to be subject to transitory and permanent changes, the latter denoted by a superscript p. The u_t and v_t are assumed to be identically and independently distributed normal variates with mean vectors zero and the following covariance structure:

$$\begin{cases} \text{cov}(u_t) = (1 - \gamma)\sigma^2 \Sigma_u \\ \text{cov}(v_t) = \gamma\sigma^2 \Sigma_v \end{cases} \quad (28)$$

The parameter $\gamma \in [0,1]$ shows the relative variance of permanent and transitory changes in the β , with $\gamma=0$ corresponding to a purely random coefficients model. Matrices Σ_u and Σ_v need to be specified a priori. Seitz considers two specifications:

- 1) $\Sigma_u = \Sigma_v = I$ where I is the Identity matrix
- 2) $\Sigma_u = \Sigma_v = \Sigma$ where Σ is derived from the variance-covariance matrix of the OLS estimation of (26) with the actual inflation rate as dependent variable.

He however finds that the estimates of both specifications yield insignificantly different parameters, a result confirmed by Smith and McAleer (1995). We will therefore consider option (1) in what follows.

Now, substitute equation (27) into (26) to obtain a state-space representation with the following measurement and transition equations:

$$\begin{cases} \mu_{per,t} = x_t' \beta_t^p + x_t' u_t \\ \beta_t^p = \beta_{t-1}^p + v_t \end{cases} \quad (29)$$

Finally, the method replaces $\mu_{per,t}$ in (29) by the actual inflation rate at t , y_t . This procedure requires additional discussion.

The natural step to take at this point would be to assume unbiasedness for inflation perceptions and replace the unknown mean perceptions by actual inflation y_t *and* an i.i.d. error term w_t . An additional error term would in any event be needed since we only have *sample* rather than population response proportions. This would give:

$$\begin{cases} y_t = x_t' \beta_t^p + x_t' u_t + w_t \\ \beta_t^p = \beta_{t-1}^p + v_t \end{cases} \quad (30)$$

The problem with this step of course is that the system is not identified any more. This is obviously unsatisfactory yet is ignored in the literature, which seems to simply substitute $\mu_{per,t}$ for y_t . This supposes both that we are dealing with population quantities *and* that mean perceptions are identically equal to actual inflation.

The system considered is thus:

$$\begin{cases} y_t = x_t' \beta_t^p + \varepsilon_t \\ \beta_t^p = \beta_{t-1}^p + v_t \end{cases} \quad (31)$$

With $Var(\varepsilon_t) \equiv h_t = \sigma^2(1 - \gamma)(x_t' \Sigma_u x_t)$. The strategy for estimation of the indifference thresholds β_t then comprises two steps:

1. Follow Cooley-Prescott (1973, 1976) to obtain the parameter values β_0^p , γ and σ^2 .
2. Implement the Kalman filter on (31) given the results from the first step.

STEP 1: COOLEY-PRESCOTT ESTIMATION OF INITIAL VALUES

The initial parameter vector β_0^p , γ and the variance can be estimated by adapting the method proposed by Cooley and Prescott to the present problem.

Let $\rho_t = x_t' u_t + x_t' \sum_{i=1}^t v_i$ and $\beta = \beta_0^p$. Then, $y_t = x_t' \beta + \rho_t$, and the full model of equation

(31) can be rewritten:

$$Y = X\beta + \rho$$

The next step consists in maximising the concentrated likelihood function:

$$L(Y; \gamma) = -\frac{T}{2}(\ln 2\pi + 1) - \frac{T}{2} \ln s^2(\gamma) - \frac{1}{2} \ln |\Omega(\gamma)|$$

Where $\Omega(\gamma)$ arises from $\text{cov}(\rho) \equiv \sigma^2 \Omega(\gamma)$. Having estimated γ , estimates for β_0^p and σ can be obtained from:

$$\begin{cases} B(\gamma) = [X' \Omega(\gamma)^{-1} X]^{-1} X' \Omega(\gamma)^{-1} Y \\ s^2(\gamma) = \frac{1}{T} [(Y - XB(\gamma))' \Omega(\gamma)^{-1} (Y - XB(\gamma))] \end{cases}$$

STEP 2: KALMAN FILTERING

Given the initial values obtained in step 1, we use Kalman filtering techniques to obtain the full time series for the indifference thresholds.

Let $\beta_0^p : N(\beta, P_0)$ and $Q_t = \text{Var}(v_t)$. The Kalman filter is then given by the following recursive relationships (see Harvey, 1989):

$$\begin{cases} \hat{\beta}_{t|t-1}^p = \hat{\beta}_{t-1}^p \\ P_{t|t-1} \equiv \text{Var}(\hat{\beta}_{t|t-1}^p) = P_{t-1} + Q_t \\ \hat{\beta}_t^p = \hat{\beta}_{t|t-1}^p + P_{t|t-1} x_t f_t^{-1} (y_t - x_t' \hat{\beta}_{t|t-1}^p) \\ P_t = P_{t|t-1} - P_{t|t-1} x_t f_t^{-1} x_t' P_{t|t-1} \end{cases} \quad (32)$$

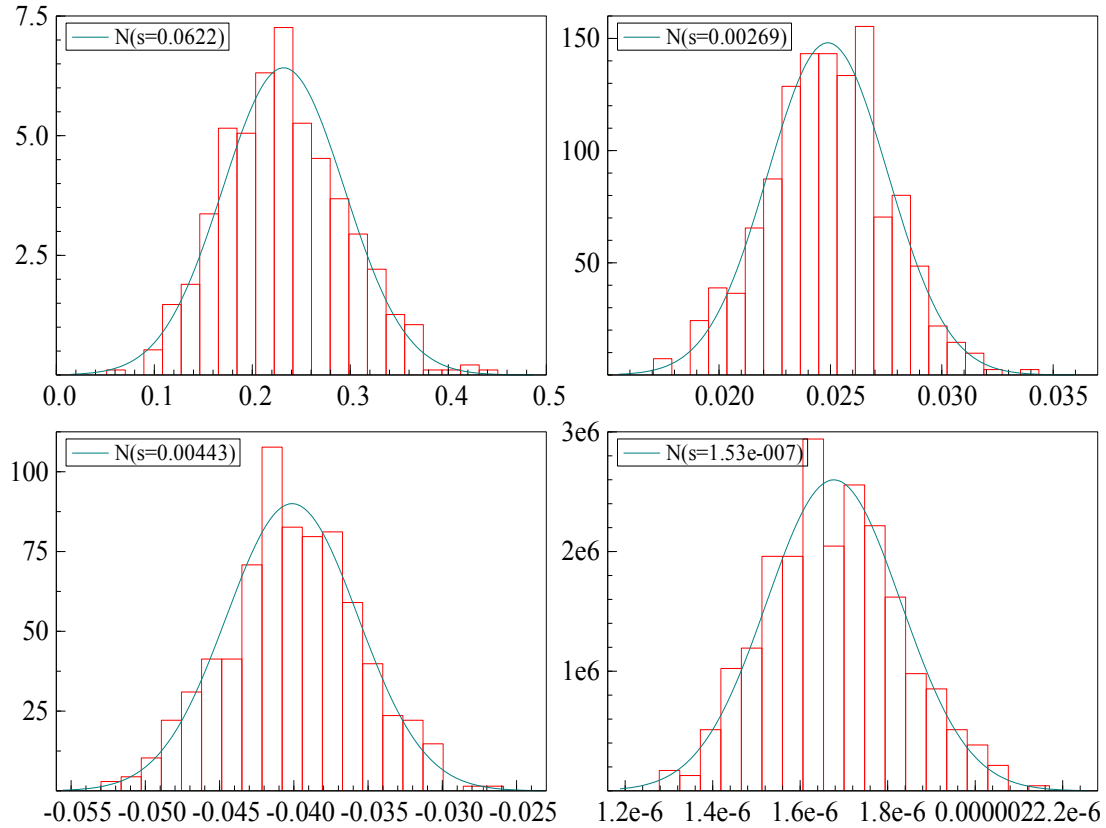
With $f_t = x_t' P_{t|t-1} x_t + h_t$.

We carried out a simulation exercise on the basis of the UK survey results, generating artificial ‘inflation/perceptions’ data for the following ‘true’ values:

$$TRUTH : \quad \beta_t \begin{cases} \delta_{a,0} = 0.025 \\ \delta_{b,0} = -0.040 \end{cases} \quad \gamma = 0.24, \quad \sigma^2 = 0.0013, \quad \Sigma_u = \Sigma_v = I \quad (33)$$

We implemented the Cooley-Prescott maximum likelihood approach in an Ox program. The validity of the code was checked by a Monte Carlo simulation. With 500 iterations, we obtained the following results, which are fully satisfactory:

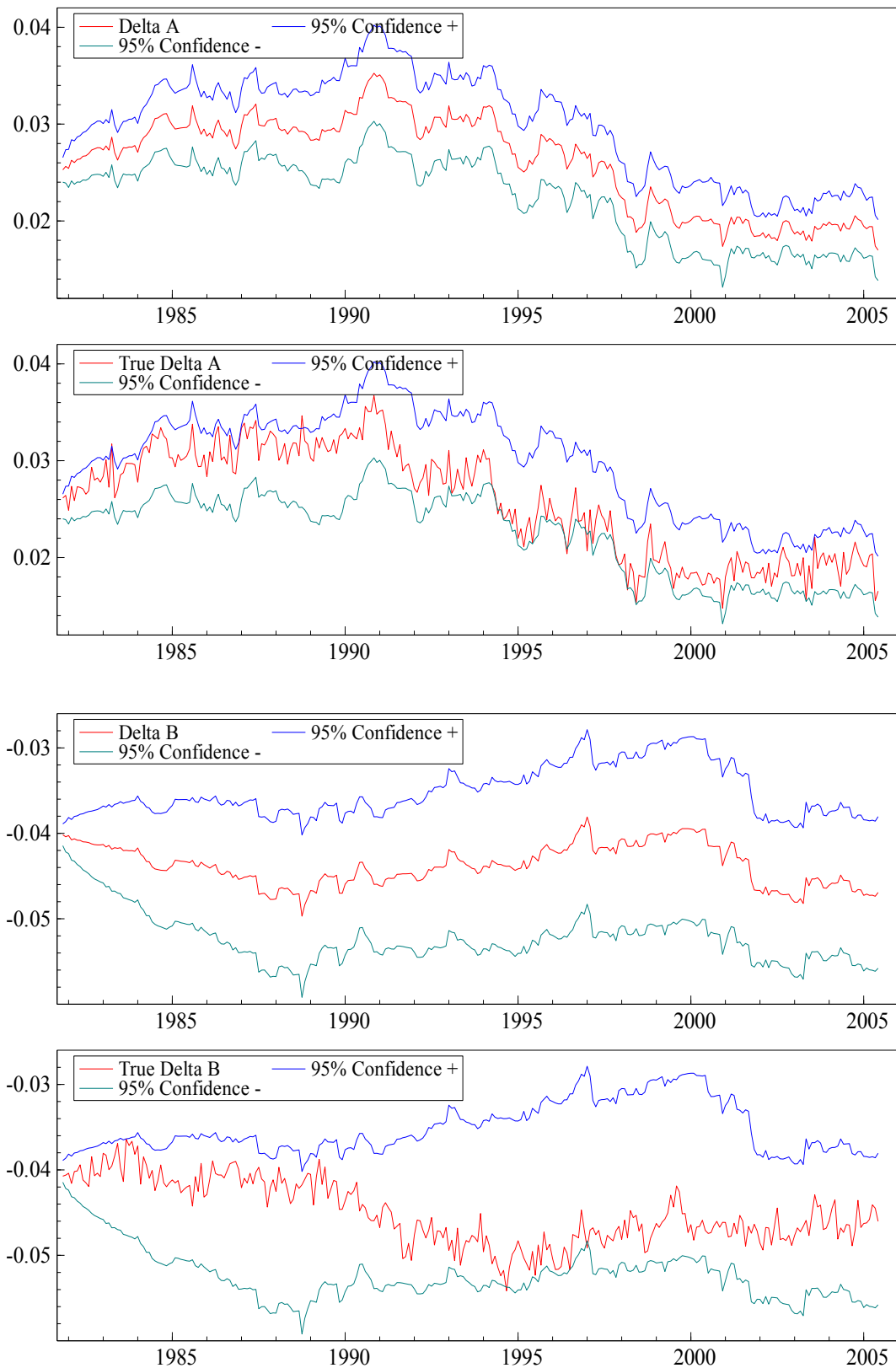
FIGURE 7: MONTE CARLO SIMULATION RESULTS; COOLEY-PRESCOTT



NOTE: Row 1 shows results for γ and δ_a ; Row 2 for δ_b and σ . Figures for 500 iterations.

The next step concerns the Kalman filtering, also implemented in Ox code. We plot in Figure 8 the δ values predicted by this procedure with their confidence bounds (graphs 1 and 3), and compare the 95% confidence bounds with the ‘true’ simulation data (graphs 2 and 4). It can be seen that the ‘true values’ are contained within our confidence bands, thereby validating our implementation of the Kalman Filter.

FIGURE 8: KALMAN FILTER SIMULATION RESULTS, 1981M11-2005M6, UK



Consider now the actual empirical implementation of this model. As noted in Nardo (2003), a situation in which the lower threshold δ_b becomes positive cannot be excluded in periods exhibiting an upward trend in inflation, for instance. If this were to be the case, the interpretation of the phenomenon would not be obvious as it would imply respondents reporting deflation for low but *positive* values of perceived inflation. We consider that such a finding would contradict the model and impose the following conditions both for the initial values of β_t^p and their evolution through time²⁷, as derived by the filter:

$$\delta_{a,t} > 0 \quad \text{and} \quad \delta_{b,t} < 0 \quad \forall t \in \{1, \dots, T\} \quad (34)$$

In practice, this involves carrying out a grid-search over possible values of γ and finding the one associated with the highest concentrated likelihood, subject to the constraint. The initial conditions β_0^p and σ^2 associated with this value of γ are then inputted with the latter to initialise the Kalman filter algorithm, and it is checked whether the constraint holds dynamically. Should this not be the case, the next-preferred value of γ is chosen to initialise the Kalman filter, until a satisfactory specification is found.

Table 5 summarises the results of this procedure for our data.

²⁷ There is therefore a constraint on the vector β – the claim to the contrary in Smith and McAleer (1995) is puzzling.

TABLE 5: SUMMARY OF SPECIFICATION SELECTION FOR TVP MODEL

Country	γ maximising unconstrained LogL	γ yielding acceptable β_0^p	γ yielding acceptable $\beta_t^p \forall t$
ESP	1	0.008	0.001
FRA	1	0.01	0.009
BEL	1	0.01	0.009
SWE	1	0.26	0.05
DEU	1	1	0
GBR	1	0.09	0.02
ITA	1	0	0
NLD	0.9	0.9	0.003

As can be seen, all countries are found to display very low acceptable values of γ . This is the first time the methodology has been implemented in the context of consumer prices to our knowledge. Seitz (1988) had found a value of 0.24 on the basis of survey data on product prices from the IFO-Institute's survey of manufacturing firms in West Germany between January 1975 and October 1985. Smith and McAleer (1995) estimated γ to be 0.25 using data from the Confederation of Australian Industries (CAI)/Westpac survey. Note that the latter authors fail to impose the model-consistent constraints of (34) and indeed seem to think it justifiable to do so²⁸, while the former did not encounter such a difficulty, it seems. Clavería González (2003) applied the methodology in the context of the EU Business Survey but does not publish the resulting estimates. Obviously, the prices and agents considered in these studies are very different from those that are the focus of this study²⁹, so comparisons are of limited value – an observation that is compounded by the differences between survey methodologies.

²⁸ Figure 2 on p.172 shows that their results violate the conditions listed in (34). They assert that “*in this model, no restrictions are imposed on a_t or b_t* ”.

²⁹ See Seitz (1988) p.428.

With the full time series of the indifference thresholds δ_a and δ_b having been derived, these are applied to an equation analogous to (25) for the case of *expectations* (see panel b of Figure 6):

$$\mu_{\text{exp},t} = \frac{\delta_{a,t}Y_{2,t} - \delta_{b,t}Y_{1,t}}{Y_{2,t} - Y_{1,t}} \quad (35)$$

where $Y_1 = \frac{\delta_{a,t} - \mu_{\text{exp},t}}{\omega_t}$ and $Y_2 = \frac{\delta_{b,t} - \mu_{\text{exp},t}}{\omega_t}$, and the Ys can be derived from the survey response proportions and the distributional assumption.

The selected indifference thresholds for the UK are plotted in Figure 9, and the derived expectations are presented against actual inflation in Figure 10. Please note that expectations are plotted at the time they are *surveyed* rather than at the time they are *referring to* (+12 months).

Although the trend decrease in absolute magnitude for the derived thresholds corresponds to the intuition of what ought to be observed in a context of falling inflation, the derived series appear to be implausibly volatile. This already raises the suspicion that alternative modelling approaches may be more appropriate³⁰.

³⁰ Alternatives generating smoother results, such as the constant intercept/stochastic slope specification of the STAMP package (Koopman et al., 1999) could be profitably explored in future research.

FIGURE 9: TIME-VARYING INDIFFERENCE THRESHOLDS, UK

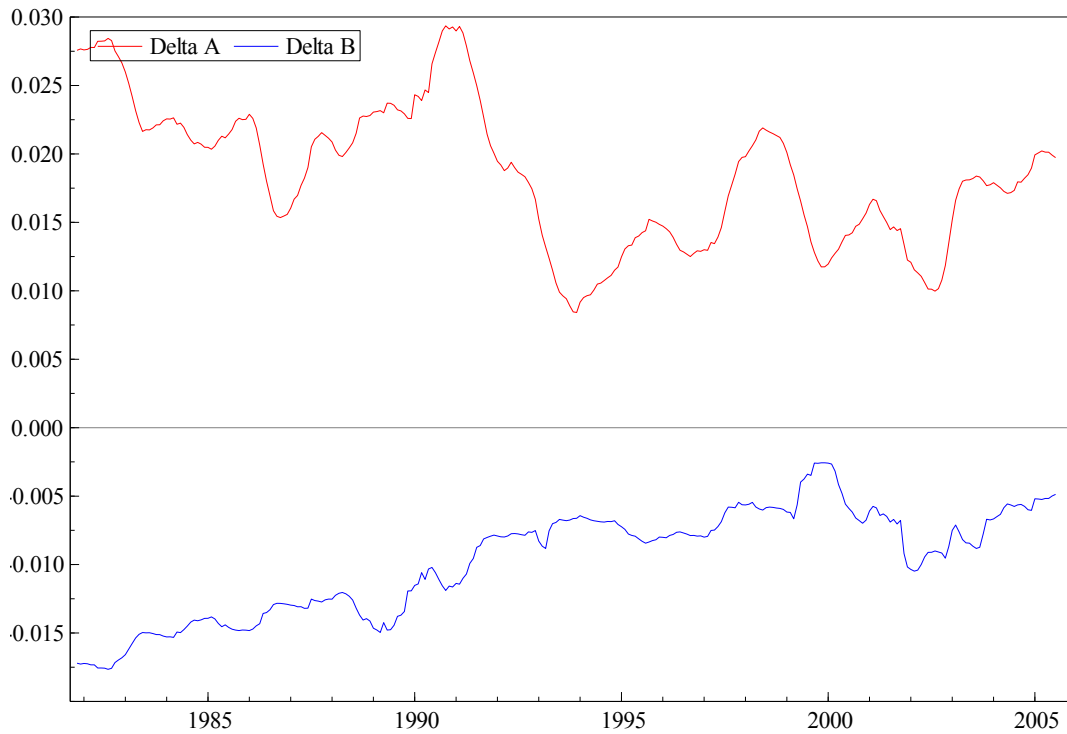
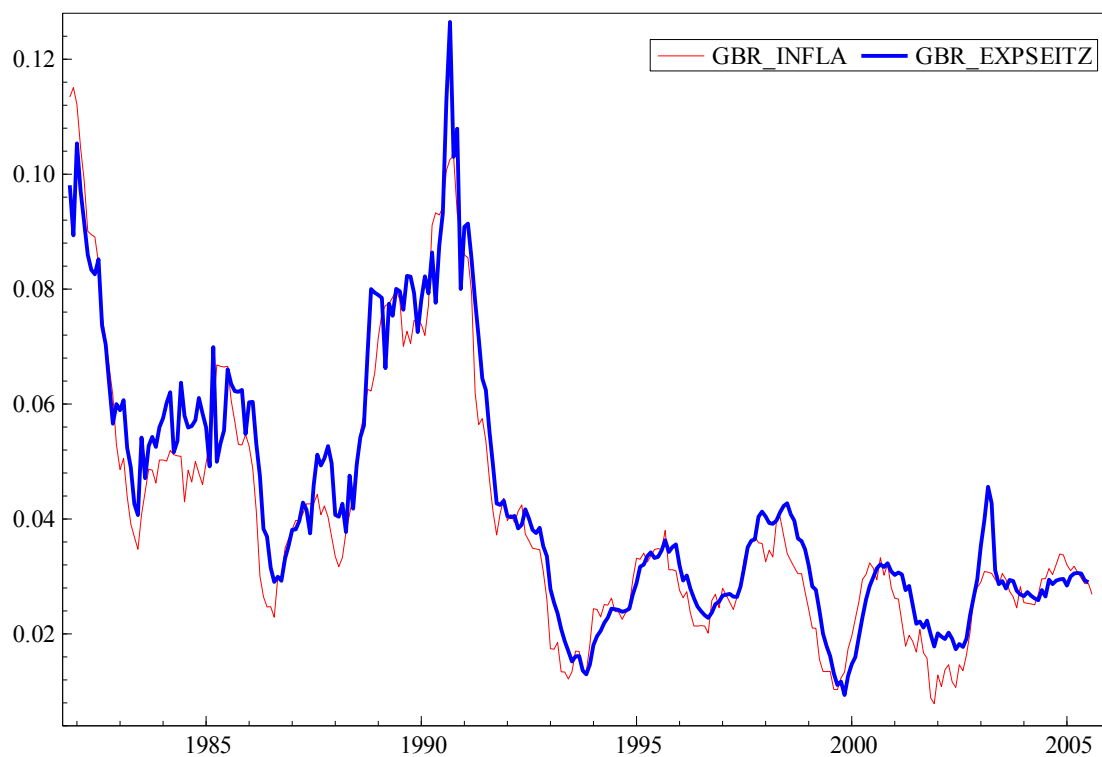


FIGURE 10: UK INFLATION EXPECTATIONS, SEITZ METHODOLOGY



3.4 Comparative theoretical discussion and summary

Following the discussion above, we are left with the following quantification methods for survey data to compare:

TABLE 6: EXPECTATIONS MEASURES SELECTED FOR COMPARISON

	Measure	Description
CP Tradition	EXPINFLA	Past inflation perceived correctly (p.23)
	EXPINFLPCP	Trichotomous quantification of perceptions (p.24)
	EXPINFLPBIS	Trichotomous perceptions, with threshold adjustment (p.24)
	EXPINFLPBO1	Pentachomous perceptions. Moderate rate is full sample inflation mean (p.27)
	EXPINFLPBO2	Pentachomous perceptions. Moderate rate from linear interpolation (p.27)
	EXPINFLPBO3	Pentachomous perceptions. Moderate rate is running average of inflation (p.27)
	EXPINFLAPES	Pesaran's regression approach (p.30)
	EXPINFLASEITZ	Time-varying stochastic parameter model for thresholds (p.37)

The relative performance of these measures is chiefly an empirical issue. However, it is possible to make a number of comments regarding their theoretical plausibility.

The most important shortcoming of the CP method is undoubtedly its requirement to maintain a firm assumption about the form of the aggregate distribution function, as has already been discussed when we justified the use of the normal distribution. Linked to this is the fact that the approach can break down when one of the response categories attracts a zero share of the ‘vote’, as this may contradict the normality assumption made for the aggregate distribution³¹.

³¹ See Section 6.3 of Chapter 3 for an exception to this statement.

TABLE 7: FREQUENCY OF ZEROES IN EC MONTHLY SURVEY DATA

Country	Perceptions data				Expectations data		
	Months Affected	as % Total	q5C	q5F	Months Affected	as % Total	q5F
Belgium	13	5.49	0	13	0	0.00	0
Germany	0	0.00	0	0	1	0.34	1
Spain	3	1.30	0	3	0	0.00	0
France	2	0.86	1	1	0	0.00	0
Italy ³²	28	10.00	0	28	14	5.00	14
Netherlands	7	3.02	0	7	0	0.00	0
Sweden	0	0.00	0	0	0	0.00	0
United Kingdom	9	3.16	0	9	6	2.11	6

Table 7 lists the frequency at which this problem arises by country and type of survey data after removing ‘false zeroes’ (e.g. those due to inappropriate rounding). As can be seen, all of the zeroes except one arise for the category corresponding to perceived or expected deflation. These are corrected manually and are typically assigned the value of 0.1%.

Another problem of the CP method arises in the instances in which it assumes constancy for response thresholds belonging to the distribution of inflation perceptions, as this can give rise to paradoxical movements for the perceived inflation measure (Pesaran, 1984; Batchelor, 1986; Nardo, 2003). An example of this is a situation in which over 50% of the respondents believe prices have increased whilst there is a shift of responses from the “stayed about the same” to the “fallen” category. This leads to a paradoxical *increase* in the quantified inflation perceptions. Intuitively, the distribution function needs to ‘flatten’ to accommodate such a move (i.e. a variance increase). Such a result can prove

³² Note that 27/28 instances of zeroes in perceptions data and all 14 of those in expectations data for Italy arise before January 1985. This is because the earlier data was provided directly by ISAE, with rounding at the nearest integer point rather than at the nearest decimal point, as appears to be the case thereafter for the European Commission’s data.

problematic if such situations are quantitatively large and common in the data. Furthermore, time-varying response thresholds are supported by signal-extraction theory (Batchelor, 1986) – suggesting, for instance, that response thresholds will be set higher when inflation uncertainty is increased.

The other assumptions concerning the response thresholds, both for the derivation of perceptions and expectations, is that they are symmetric and equal for all individuals. There is no specific justification for the former (Berk, 1999), and the latter has never been addressed in the literature (Nardo, 2003a).

The Pesaran approach has the advantage that it avoids problems linked to zeroes in the data and the counterintuitive movements of the CP expectations measures in the situations just noted. However, some arbitrariness is involved in specifying the basic regression on which the method relies. The evidence of residual autocorrelation and general misspecification clearly argues against the model, and we are reduced to a weak ‘yardstick’ interpretation of this relation to go forward. It is obviously a bad policy to impose an AR correction to the error term and proper alternatives would mean a radical re-working of the model. In a sense then, the quantitative expectations so-derived are a *“function of a specific regression model rather than a specific probability distribution”* as was the case with CP (Smith and McAleer, 1995).

The structure of the Pesaran methodology raises a further, potentially rather serious concern for its application in the specific context of the European Commission’s

Consumer Survey. As can be seen from Table 1, the expectations question is expressed *relative* to the previous 12 months, whereas the perceptions question is phrased in an *absolute* manner. This creates some doubts as to the legitimacy of estimating the relationship between official inflation data and the survey data of question 5 before applying that same relationship to the data from question 6 to obtain expectations, given the qualitative difference between the questions.

From a theoretical perspective, the Seitz approach has the advantage of taking time-variation of the indifference thresholds seriously. However, the random walk assumption for β_t is both an arbitrary way of modelling indifference thresholds and makes the state space model unstable, making it impossible to determine when the Kalman filter reaches a steady state (Harvey, 1989).

Notice also that whereas the Batchelor and Orr methodology treated the processes of perceptions and expectations formation as being quite independent, both the Pesaran and Seitz approaches link them through the assumption of identical indifference thresholds in both situations. This represents a crucial difference between the methodologies.

4 COMPARING METHODS ON THE BASIS OF PREDICTIVE ABILITY

4.1 Introduction

It has been commonplace in the literature to contrast different quantification methods on the basis of their forecasting performance (e.g. Dasgupta and Lahiri, 1992; Smith and McAleer, 1995; Clavería González, 2003; Nolte and Pohlmeier, 2004). The statistics considered notably include the MAE, RMSE, the proportion of correct turning points and a non-parametric predictive performance test statistic (Pesaran and Timmermann, 1990).

We wish to stress however that in no case can these exercises constitute a scientific basis for choosing the quantification method that most accurately *represents* the actual underlying expectations. Drawing such a conclusion would only be warranted in certain worlds involving rational expectations³³ – an assumption we are unwilling to make. Indeed, it cannot be excluded a priori that such expectations may be biased, fit the data poorly and forecast inadequately. Nonetheless, the topic is of interest in its own right, and given its popularity in the literature it is worthwhile to see what choice would be made on its basis for the set of quantification approaches considered.

4.2 Results

We first provide the Root Mean Squared Error (RMSE) as a measure of the accuracy of the competing expectations measures. For the purposes of establishing a benchmark, we also provide the results of forecasts based on a “naïve” methodology, which simply uses the inflation rate over the previous twelve months as an expectation.

³³ With limited information, rational expectations could potentially underperform naïve expectations, particularly if one allows for *unpredictable* structural breaks (see Cukierman, 1986). In a sense, the statement depends on how rational we allow the ‘rational expectations’ to be.

This exercise is carried out using the full sample of *monthly* expectations observations available for each country, with sample starting points provided in Table 2 (page 18).

Inflation expectations data are available for all countries until July 2005.

TABLE 8: RMSE RESULTS FOR PREDICTING ACTUAL INFLATION

	France		Spain		Belgium		Germany	
Measure	RMSE	Rank	RMSE	Rank	RMSE	Rank	RMSE	Rank
EXPINFLA	0.0092	5	0.0160	4	0.0112	8	0.0118	2
EXPINFLPCP	0.0101	7	0.0202	7	0.0103	6	0.0127	7
EXPINFLPBIS	0.0088	4	0.0180	5	0.0103	5	0.0119	3
EXPINFLPBO1	0.0127	9	0.0221	8	0.0100	2	0.0138	9
EXPINFLPBO2	0.0121	8	0.0199	6	0.0101	4	0.0126	6
EXPINFLPBO3	0.0093	6	0.0366	9	0.0117	9	0.0124	4
EXPINFLAPES	0.0078	3	0.0124	2	0.0107	7	0.0117	1
EXPSEITZ	0.0076	2	0.0116	1	0.0100	1	0.0128	8
NAÏVE	0.0070	1	0.0127	3	0.0101	3	0.0124	5
Period	1987m4-2005m8		1987m6-2005m8		1986m10-2005m9		1981m10-2005m8	

	Italy		United Kingdom		Sweden		Netherlands	
Measure	RMSE	Rank	RMSE	Rank	RMSE	Rank	RMSE	Rank
EXPINFLA	0.0202	6	0.0202	7	0.0118	3	0.0099	4
EXPINFLPCP	0.0153	3	0.0163	2	0.0105	1	0.0101	5
EXPINFLPBIS	0.0164	4	0.0161	1	0.0107	2	0.0096	3
EXPINFLPBO1	0.0164	4	0.0214	9	0.0127	6	0.0117	8
EXPINFLPBO2	0.0122	1	0.0197	5	0.0126	5	0.0117	7
EXPINFLPBO3	0.0490	9	0.0180	3	0.0139	7	0.0145	9
EXPINFLAPES	0.0261	8	0.0182	4	0.0163	9	0.0091	1
EXPSEITZ	0.0247	7	0.0202	6	0.0155	8	0.0101	6
NAÏVE	0.0148	2	0.0205	8	0.0125	4	0.0092	2
Period	1983m1-2005m9		1982m11-2005m8		1997m1-2005m8		1987m4-2005m8	

As can be seen, the performance of each indicator in terms of ranking varies rather widely across countries. Table 9 summarises this information. It provides the average country rank computed from Table 8 as well as the overall average RMSE and its associated ranking and standard deviation.

TABLE 9: AVERAGE RANK & RMSE FOR PREDICTIVE SUCCESS

Measure	Avg. Cty Rank	Avg. RMSE	Rank	RMSE Std
EXPINFLA	5	0.0138	4	0.0044417
EXPINFLPCP	5	0.0132	3	0.0037507
EXPINFLPBIS	3	0.0127	2	0.0035566
EXPINFLPBO1	7	0.0151	8	0.0044705
EXPINFLPBO2	5	0.0139	5	0.0037511
EXPINFLPBO3	7	0.0207	9	0.0142651
EXPINFLAPES	4	0.0140	6	0.0059798
EXPSEITZ	5	0.0141	7	0.005786
NAÏVE	4	0.0124	1	0.0040757

The table shows that the Carlson-Parkin indicator EXPINFLPBIS, although outperformed in selected countries, is the leader in terms of predictive ability among the competing methodologies. Nevertheless, it marginally underperforms the naïve benchmark. The standard deviation of the RMSEs is roughly comparable across indicators, except in the case of EXPINFLPBO3, although this finding is largely driven by its poor performance in the context of Spain and Italy.

An interesting feature of the Mean Squared Error (MSE) is that it can be decomposed into three insightful components, as noted in Clavería González (2003). Let y denote inflation; bars placed over variables indicate averages, whereas hats $\hat{}$ indicate the prediction of the variable for the relevant period. Hence, the expectation error in period t can be written as $e_t = (y_t - \hat{y}_t)$ and we have:

$$MSE = \frac{1}{T} \sum_{t=1}^T e_t^2 = (\bar{\hat{y}} - \bar{y})^2 + (\hat{\sigma} - r_{\hat{y}y}\sigma)^2 + (1 - r_{\hat{y}y}^2)\sigma^2 \quad (36)$$

Where $r_{\hat{y}y}$ denotes the correlation coefficient between predicted and actual values, while $\hat{\sigma}$ and σ correspond to the standard deviations of predictions and observations, respectively. The three components on the right hand side can be expressed as percentages by dividing through by the MSE: $1 = U1 + U2 + U3$.

These components have an intuitive interpretation. *U1 refers to the proportion of MSE imputable to bias*, being the square of the difference between the mean of predicted values and that of actual values. Since *U2* depends on the difference between the standard deviations of predictions and actual values, it has an interpretation as the *proportion of MSE due to dispersion*, i.e. ‘regression error’. Finally, *U3* arises from the lack of correlation between predictions and actual values and therefore represents the proportion of MSE due to all the *factors that are unexplained or unaccounted for*.

In terms of this decomposition, the ideal outcome would be for the greatest weight to be achieved by the unexplained component of MSE, with a minimisation of systematic (*U1*) and regression error (*U2*). This provides us with an additional criterion on which to assess our quantification measures.

Table 10 makes clear that significant differences in the MSE decomposition arise across countries for any given measure. The average results are provided in Table 11. Measures such as EXPINFLPCP, EXPINFLPBIS, EXPINFLAPES and EXPSEITZ exhibit the most desirable properties in this respect.

Finally, we consider the ability of our expectations measures to detect *turning points* in the actual inflation series. That is, we examine whether the direction of monthly changes in annual inflation tend to be reflected by changes in consumers' expectations thereof.

Table 12 provides the results on this account.

TABLE 10: MSE DECOMPOSITION, DETAILS

	France			Spain			Belgium			Germany		
Measure	U1	U2	U3	U1	U2	U3	U1	U2	U3	U1	U2	U3
EXPINFLA	37.0	7.7	55.3	45.5	6.8	47.7	15.8	34.7	49.5	0.1	28.8	71.1
EXPINFLPCP	34.0	8.1	57.9	31.9	17.8	50.3	20.3	19.0	60.6	1.0	4.2	94.8
EXPINFLPBIS	44.5	2.2	53.3	39.5	9.8	50.8	16.6	22.5	60.9	0.5	7.6	91.9
EXPINFLPBO1	57.7	1.5	40.8	56.9	1.9	41.3	32.1	5.5	62.4	20.5	6.7	72.8
EXPINFLPBO2	65.6	0.5	33.9	66.5	0.2	33.4	29.9	9.0	61.1	23.6	12.4	64.0
EXPINFLPBO3	44.2	0.5	55.3	79.1	7.0	13.9	7.1	44.2	48.7	0.0	8.5	91.4
EXPINFLAPES	0.1	0.0	99.9	2.6	0.0	97.3	34.9	10.8	54.3	1.7	1.6	96.7
EXPSEITZ	0.3	14.9	84.9	9.7	1.3	89.0	13.0	24.0	63.0	8.8	0.4	90.8
NAÏVE	0.2	17.7	82.1	3.2	27.3	69.4	0.1	39.3	60.6	1.7	23.1	75.2

	Italy			United Kingdom			Sweden			Netherlands		
Measure	U1	U2	U3	U1	U2	U3	U1	U2	U3	U1	U2	U3
EXPINFLA	8.0	63.1	28.9	0.4	40.7	59.0	2.5	27.8	69.8	9.2	36.6	54.2
EXPINFLPCP	3.8	7.1	89.1	1.2	6.6	92.2	5.8	3.4	90.8	28.1	1.5	70.5
EXPINFLPBIS	8.5	36.1	55.4	1.6	9.8	88.6	7.3	6.1	86.6	26.1	0.2	73.7
EXPINFLPBO1	0.1	6.4	93.5	43.4	3.4	53.2	38.0	0.4	61.6	51.7	0.6	47.7
EXPINFLPBO2	4.3	10.9	84.8	44.3	1.1	54.6	37.9	0.1	62.0	49.8	0.0	50.1
EXPINFLPBO3	51.6	44.0	4.4	5.3	10.4	84.2	46.9	0.0	53.0	59.0	0.9	40.1
EXPINFLAPES	41.0	1.9	57.1	28.2	3.1	68.7	31.3	29.9	38.8	19.7	7.0	73.3
EXPSEITZ	34.0	5.4	60.6	7.5	29.9	62.6	24.0	35.2	40.7	0.5	38.6	60.9
NAÏVE	15.5	32.3	52.2	1.7	34.1	64.2	0.0	38.2	61.8	1.4	30.0	68.6

TABLE 11: MSE DECOMPOSITION, AVERAGES BY MEASURE

Measure	Avg. U1	Avg. U2	Avg. U3
EXPINFLA	14.8	30.8	54.4
EXPINFLPCP	15.8	8.5	75.8
EXPINFLPBIS	18.1	11.8	70.1
EXPINFLPBO1	37.5	3.3	59.2
EXPINFLPBO2	40.2	4.3	55.5
EXPINFLPBO3	36.7	14.4	48.9
EXPINFLAPES	19.9	6.8	73.3
EXPSEITZ	12.2	18.7	69.1
NAÏVE	3.0	30.3	66.8

TABLE 12: % CORRECTLY PREDICTED TURNING POINTS

Measure	France		Spain		Belgium		Germany	
	%	Rank	%	Rank	%	Rank	%	Rank
EXPINFLA	33.6	8	40.8	7	33.0	9	35.0	9
EXPINFLPCP	38.6	5	47.2	3	47.6	1	43.4	3
EXPINFLPBIS	40.9	2	47.2	3	47.6	1	43.4	3
EXPINFLPBO1	36.8	6	41.3	6	41.9	4	40.2	6
EXPINFLPBO2	39.1	3	42.7	5	41.9	4	40.6	5
EXPINFLPBO3	39.1	3	40.4	8	41.4	6	40.2	6
EXPINFLAPES	41.4	1	49.1	1	42.3	3	46.9	1
EXPSEITZ	35.9	7	47.7	2	38.8	7	44.1	2
NAÏVE	33.6	8	39.9	9	35.2	8	38.1	8
N° Observations	220		218		227		286	

Measure	Italy		United Kingdom		Sweden		Netherlands	
	%	Rank	%	Rank	%	Rank	%	Rank
EXPINFLA	44.6	7	36.6	9	36.9	8	39.1	8
EXPINFLPCP	43.3	9	43.6	5	47.6	1	46.4	3
EXPINFLPBIS	45.9	3	44.0	4	47.6	1	46.8	2
EXPINFLPBO1	45.5	4	45.1	3	44.7	6	40.0	7
EXPINFLPBO2	48.9	1	45.4	1	46.6	4	40.5	5
EXPINFLPBO3	48.1	2	45.4	1	45.6	5	40.5	5
EXPINFLAPES	43.8	8	40.7	6	47.6	1	45.5	4
EXPSEITZ	45.4	6	38.8	7	40.8	7	49.1	1
NAÏVE	45.4	5	37.0	8	33.0	9	35.5	9
N° Observations	270		273		103		220	

As can be seen, the percentages of correctly predicted turning points are all in the region of 0.3-0.5 – there is not a single case of overperformance relative to a fair coin flip. In fact, the concern is actually the opposite – in many cases, the measures significantly *underperform* chance according to binomial tests³⁴. Table 13 summarises the information above. On balance, EXPINFLPBIS, EXPINFLPCP and EXPINFLAPES are the best performers.

³⁴ As a benchmark, the critical % of correct turning points at the 5% level for 220 observations is roughly 44%

TABLE 13: AVERAGE RANK, SUCCESS IN TURNING POINT PREDICTION

Measure	Avg. Cty Rank	Avg %	Rank
EXPINFLA	8	37.5	8
EXPINFLPCP	4	44.7	2
EXPINFLPBIS	2	45.4	1
EXPINFLPBO1	5	41.9	7
EXPINFLPBO2	4	43.2	4
EXPINFLPBO3	5	42.6	5
EXPINFLAPES	3	44.6	3
EXPSEITZ	5	42.6	6
NAÏVE	8	37.2	9

In conclusion, two results stand out from the previous analysis. First, if we were to choose a quantification methodology on this basis, the best choices would be the most traditional – namely EXPINFLPCP, the Carlson-Parkin method as adapted by Batchelor and Orr (1988) and Berk (1999), or its close cousin EXPINFLPBIS suggested and used in Curto Millet (2004). EXPINFLAPES also seems to perform reasonably well. Second, however, it is now very clear that predictive ability is not an appropriate basis for such a decision. As argued previously, this would only be the case in the context of rational expectations – yet the survey measures contrasted here actually *underperform* (in RMSE terms) the naïve indicator used for benchmarking. Unless one takes the extreme view that inflation is a pure random walk, such results can hardly be described as bearing the hallmark of rational expectations.

Therefore, we will now turn to a benchmarking of our quantification methodologies with the two sources of comparable *quantitative* expectations data available in Europe. We first present these quantitative data in the next section, before using them to make a sounder assessment of our quantification methodologies in Section 6.

5 QUANTITATIVE SURVEY DATA

5.1 United Kingdom: the Gallup Survey

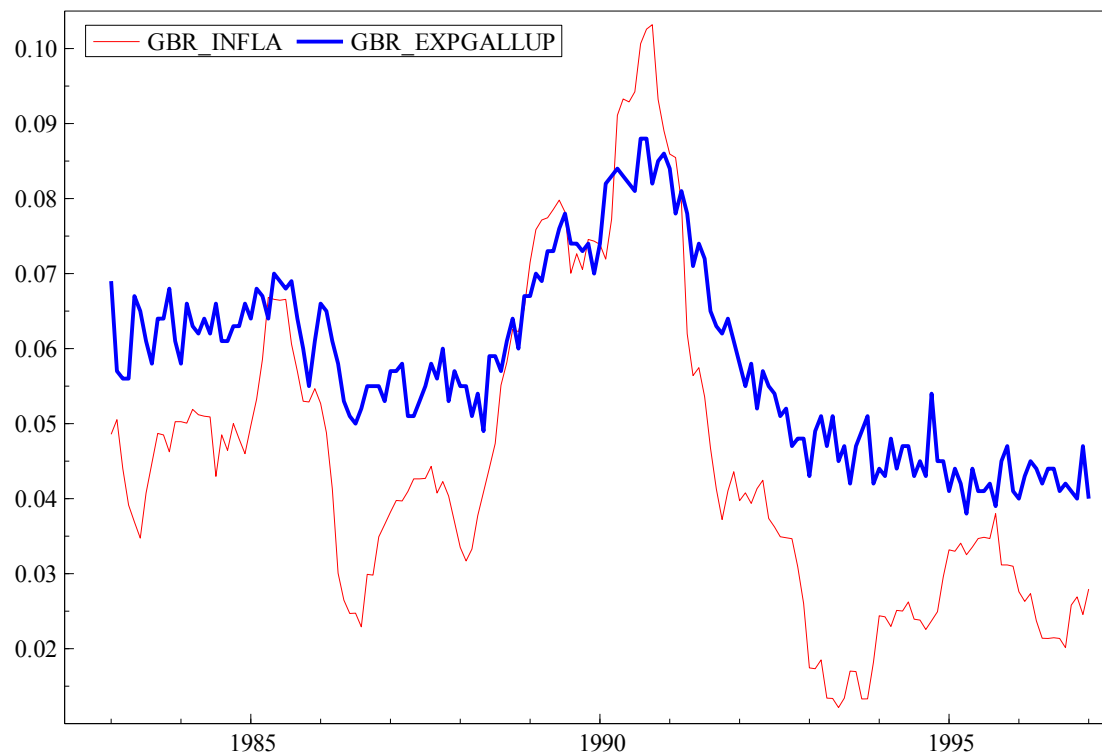
Social Surveys (Gallup Poll) Ltd carried out a monthly survey on a sample of employees in the United Kingdom continuously between January 1983 and January 1997. The results of this were published in the *Gallup Political Index* reports, subsequently renamed *Gallup Political and Economic Index* reports.

The special feature of this survey is that the responses required were of a *quantitative* nature. The precise question asked was:

Over the next twelve months, what do you think the rate of inflation will be?

The sample consisted of some 500 individuals, with roughly 80% of them addressing this question throughout the sample period. Gallup published the average survey response monthly, which is the series that will be used in what follows. Plotting these inflation expectations for the coming 12 months *at the time they are surveyed* against inflation yields Figure 11:

FIGURE 11: QUANTITATIVE GALLUP DATA & ACTUAL CPI INFLATION, UK



One remarkable feature of this picture is how quickly upward increases in inflation have fed into expectations, whereas it took a considerable amount of time for the low inflation periods of 1986-1988 and the 1990s to lower these. The graph also makes it clear that expectations exhibit extremely important ‘adaptive’ characteristics.

5.2 Sweden: the HIP Survey (SCB/GfK)

The *Households' Purchasing Plans* survey (*Hushållens Inköpsplaner* in Swedish, or *HIP*) has included quantitative questions on inflation expectations and perceptions on a quarterly basis between 1979q1 and 1992q4, the frequency becoming monthly from January 1993 onwards. With Sweden's entry into the EU and the subsequent questionnaire harmonisation that took place, the survey started to include the standard EC qualitative questions on inflation perceptions and expectations³⁵ (in addition to the quantitative ones) from January 1996.

The quantitative questions included in the HIP survey have the following phrasing³⁶:

TABLE 14: HIP SURVEY, QUANTITATIVE QUESTIONS

Concept	Question:
Perceptions	Compared with 12 months ago, how much higher in percent do you think that prices are now?
Expectations	Compared with today, by what percentage do you think that prices will go up (i.e. the rate of inflation 12 months from now)?

SOURCE: Konjunkturinstitutet, Hushållens Inköpsplaner – User Manual

The survey has been conducted by Statistics Sweden (SCB) since 1973. The sample size has been successively reduced since that time. In July 1979, it consisted of 6,600 households. In 2000-2001 the sample comprised 2,100 interviews. Starting in January 2002, the survey has been carried out by GfK Sverige AB with a total sample of 1,500 interviews of individuals aged 16 through 84³⁷.

³⁵ See TABLE 1, page 18.

³⁶ See Palmqvist and Strömberg (2004) for further details on the practical implementation of the survey

³⁷ See GfK (2003), Consumer Survey 2003.

Palmqvist and Strömberg (2004) noted that despite nearly identical questions, the average inflation perceptions and expectations were considerably higher in GfK's surveys relative to SCB's. They attributed part of the difference to sample stratification and to the probing procedure applied to respondents choosing an "about the same" response. However, part of the difference between the series remained unexplained. In the case of inflation expectations, the authors suggested using a dummy variable to ensure the series remains consistent. This suggests that some prudence should be exercised when considering samples that span the break related to a change in the institution collecting the survey³⁸.

The following figures plot the quantitative HIP data on perceptions and expectations against CPI inflation for the whole sample period:

³⁸ Indeed, Bryan and Palmqvist (2005) choose to end the sample they consider in 2001 due to this "significant break in the mean survey response at the beginning of 2002".

FIGURE 12: HIP PERCEPTIONS vs ACTUAL INFLATION, SWEDEN, 1979Q1-2005Q3

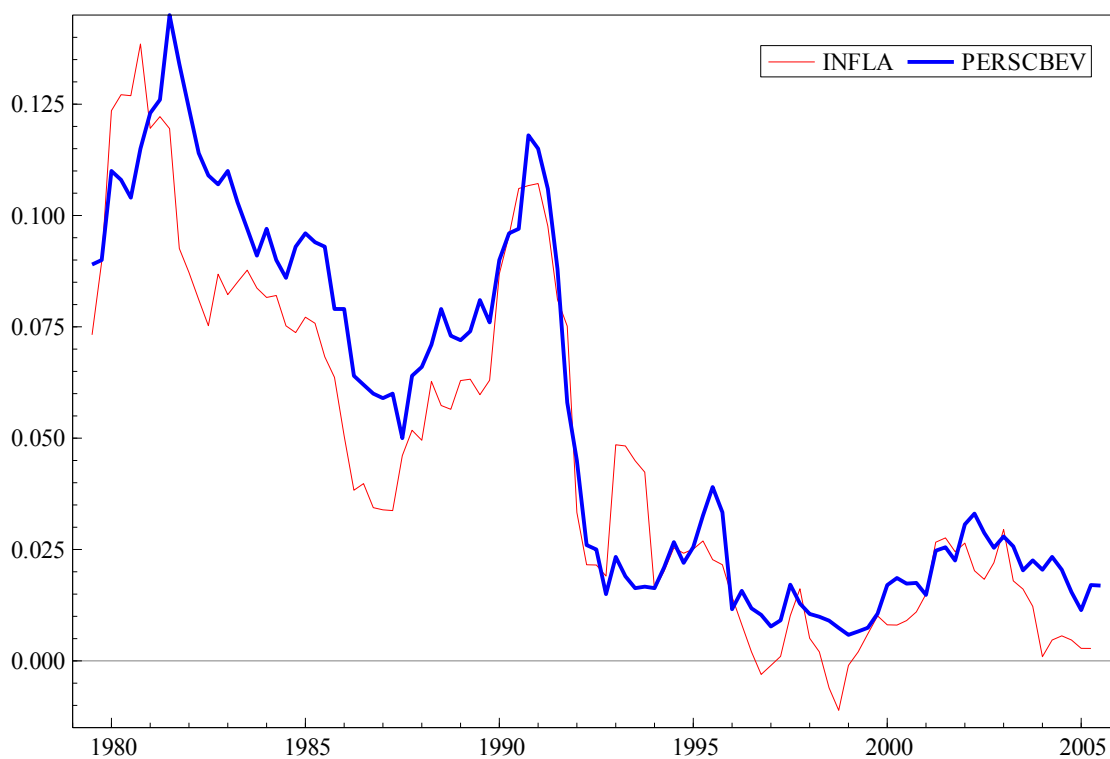
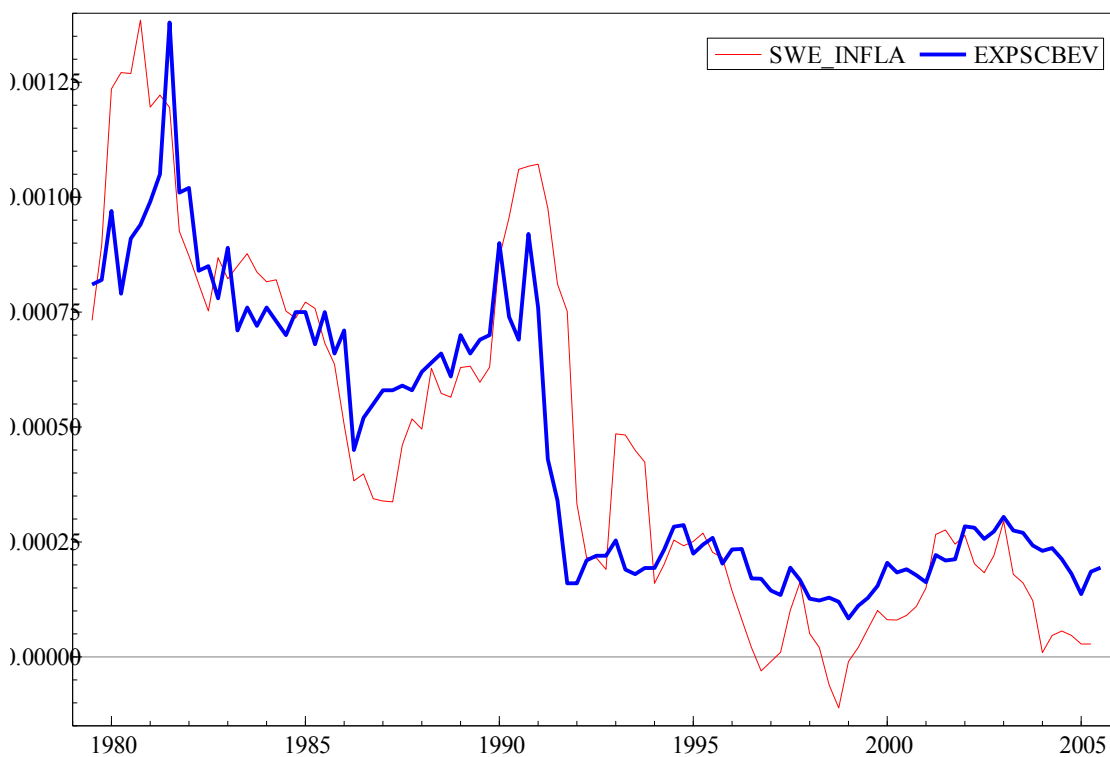


FIGURE 13: HIP EXPECTATIONS vs ACTUAL INFLATION, SWEDEN, 1979Q1-2005Q3



6 BENCHMARKING QUALITATIVE DATA ON QUANTITATIVE DATA

6.1 Predictive performance statistics for the quantitative measures

We present for completeness the statistics on predictive performance used in Section 4 for our quantitative expectations series. As can be seen, the results are broadly comparable to those of our other measures.

TABLE 15: PREDICTIVE PERFORMANCE OF QUANTITATIVE MEASURES

Country	Period	RMSE	U1 %	U2 %	U3 %	% Correct Turning Points
UK	1984m1-1998m1	0.0230	37.8	0.3	61.9	42.26
Sweden	1994m1-2005m8	0.0142	36.3	10.2	53.6	46.04

6.2 Descriptive statistics

More interestingly, we now contrast the quantified expectations from the EC Consumer Survey data with the quantitative data from Gallup surveys for the United Kingdom as well as Swedish HIP data, the latter two being treated as an approximation of the ‘true’ underlying expectations. Thus, we carry out for Europe an exercise in the spirit of Batchelor’s (1986) work on US data.

As noted in the previous section, the Gallup data are based on a survey of employees. Batchelor and Dua (1987) rightly remark that a sample of employees need not be equivalent to one of consumers. Indeed, employees may appear to be relatively more ‘rational’ in their assessments of inflation given their incentives to gather information on these matters for purposes of wage bargaining. We will nonetheless use this data as a benchmark to assess the output of quantification techniques applied to consumer surveys. We believe that this is likely to be a reasonable approximation. Furthermore, to our

knowledge, the only other quantitative evidence available in Europe on inflation expectations over a reasonable sample³⁹ concerns surveys of ‘experts’, which clearly constitute a very different group from that of consumers.

Table 16 presents simple correlations between our expectations measures and the quantitative ones. It is interesting to note how different the ranking of measures is in both countries. The reason underlying this finding is unclear. An immediate thought might be that there is simply a difference in the structure of expectations formation in both countries, for instance in terms of the behaviour of indifference thresholds, thereby explaining the differentiated behaviour of the various measures. This still prompts the question of what fundamental difference would lie behind this. The periods for which both quantitative and qualitative data are available for comparison (UK: 1983m1-1997m1; SWE: 1996m1-2005m7) are very different in terms of inflation behaviour – low and relatively stable for Sweden compared to relatively volatile, downward-trending behaviour in-sample for the UK. It may well be the case that different measures display different comparative advantages depending on underlying inflation behaviour.

TABLE 16: CORRELATIONS WITH QUANTITATIVE MEASURES

Measure	UK Corr	UK Rank	SWE Corr	SWE Rank
EXPINFLA	0.8755	4	0.7451	8
EXPINFLPCP	0.7763	7	0.9078	1
EXPINFLPBIS	0.8171	5	0.8806	5
EXPINFLPBO1	0.7957	6	0.8855	4
EXPINFLPBO2	0.8841	3	0.8531	6
EXPINFLPBO3	0.9027	2	0.808	7
EXPINFLAPES	0.6262	8	0.9066	2
EXPSEITZ	0.9086	1	0.8868	3

³⁹ The Bank of England/NOP Inflation Attitudes Survey also surveys the expectations of the general public, but is only available quarterly since November 1999. See Section 6.6.2 in Chapter 3.

The following table provides the same statistics in the context of this “matching” exercise as were presented in Section 4, with the addition of the mean error (ME) as an indicator of bias.

TABLE 17: ABILITY TO MATCH QUANTITATIVE DATA

Measure	UK					Sweden				
	ME	RMSE	U1	U2	U3	ME	RMSE	U1	U2	U3
EXPINFLA	0.0151	0.0215	68.5	26.5	5.0	0.0116	0.0129	81.8	8.9	9.3
EXPINFLPCP	0.0158	0.0184	83.5	6.8	9.8	0.0121	0.0124	93.7	2.3	3.9
EXPINFLPBIS	0.0153	0.0177	83.5	8.6	7.9	0.0123	0.0126	94.6	0.5	4.9
EXPINFLPBO1	0.0308	0.0318	95.4	0.2	4.4	0.0172	0.0179	92.0	5.6	2.3
EXPINFLPBO2	0.0281	0.0288	97.2	0.0	2.8	0.0171	0.0177	92.9	4.1	3.0
EXPINFLPBO3	0.0192	0.0199	91.8	4.2	4.0	0.0188	0.0196	92.2	4.6	3.1
EXPINFLAPES	0.0042	0.0129	45.3	23.8	30.9	0.0011	0.0045	6.3	63.6	30.1
EXPSEITZ	0.0097	0.0154	61.7	31.3	7.0	0.0024	0.0069	11.6	73.2	15.2
NAÏVE	0.0132	0.0173	67.9	26.1	5.9	0.0097	0.0122	63.5	24.0	12.5
Period	1983m1-1997m1					1996m1-2005m7				

As is readily seen, the best performers in terms of both RMSE and the MSE decomposition are EXPINFLAPES and EXPSEITZ. This is in large part due to a considerably lower bias of these measures relative to the alternatives. EXPINFLAPES also performs relatively well in terms of correctly indicating the direction of changes in the quantitative series, as do the EXPINFLPBO series.

TABLE 18: % CORRECTLY INDICATED TURNING POINTS

Measure	UK		Sweden	
	%	Rank	%	Rank
EXPINFLA	48.81	5	63.16	8
EXPINFLPCP	42.86	8	81.58	1
EXPINFLPBIS	42.86	8	81.58	1
EXPINFLPBO1	51.19	3	78.95	3
EXPINFLPBO2	51.19	3	75.44	5
EXPINFLPBO3	52.38	1	75.44	5
EXPINFLAPES	52.38	1	78.07	4
EXPSEITZ	45.83	6	74.56	7
NAÏVE	45.83	6	53.51	9
N° Observations	168		114	

It is interesting to note that for the United Kingdom, none of the measures are significantly superior to chance in predicting the direction of changes in the quantitative Gallup data (in fact, EXPINFLPCP and EXPINFLPBIS here are marginally significantly inferior at the 5% level in a binomial test). Considering this evidence alone, this would seem to reinforce the case against the popular claim that tendency surveys are more useful at indicating the *direction* of changes rather than levels or magnitudes. Indeed, it is consistent with past findings in the literature by Batchelor (1986) for monthly US SRC data over the period 1977-1984 and by Defris and Williams (1979) for a series of quarterly Australian data over 1973-1977.

However, the Swedish data used present a very different picture. Here, almost all of the measures perform *significantly* better than chance at the 0.1% level in capturing the direction of the underlying changes in the quantitative variable⁴⁰. Possible explanations for this finding might point to the difference in the respondent bases of EC survey and the Gallup one (respectively, consumers and employees) or to particularities of the relatively short Swedish sample with low and stable inflation that might facilitate good performance of the quantified survey measures.

This analysis is informative but cannot ultimately answer the question of what quantification method is to be preferred to the others for control purposes in regression equations (where mean biases are corrected for by the constant in the equation, for instance). This question is the focus of the remainder of this chapter.

⁴⁰ The critical % of correctly predicted points at the 0.1% level is roughly 64.5%.

6.3 Non-nested testing of quantification methods

We next proceed to a non-nested test of the quantified expectations series by regressing the quantitative expectations data for the UK and Sweden on these and a constant. The aim of this exercise is to verify whether one or a combination of these measures can be found to be dominating in terms of explanatory power of what we assume to be the ‘true’ underlying expectations.

The ambition of this exercise is limited by the strong collinearity between several of these series. Nevertheless, it may be helpful in identifying which broad ‘tradition’ (Carlson-Parkin, Pesaran or Seitz) performs best in this context. A further caveat concerns the regression for Sweden. As has already been noted, the sample here is limited to post-1996 data, a period characterised by relatively low and stable inflation. This suggests that the UK data will likely be more informative given its variation, and greater weight might accordingly be placed on its results.

The data used have a monthly frequency. The following tables present the results obtained after dropping insignificant and wrongly-signed variables⁴¹:

⁴¹ The initial regressions are subject to significant multicollinearity, notably visible through the large standard errors relating to the different variants within the Carlson-Parkin tradition. At such, it is not easy to infer a great deal from this first step, except that information from multiple traditions seems to be relevant. The particular qualitative assessment reached below thus appears to be robust and not a product of the selection process. Over the selection process, we find that the partial r^2 statistics of the retained variables rise overall, suggesting that the variables dropped acted primarily as substitutes of those retained. However, the rise of the partial r^2 s is non-monotonous, suggesting more complex patterns of complementarity and substitutability of information content. The coefficients of the retained variables overall do not change dramatically over the selection process, excepting that on the Carlson-Parkin tradition in Sweden, which exhibits the volatility to be expected from an initial situation of significant multicollinearity.

TABLE 19: NON-NESTED TESTING RESULTS, UK

Modelling GBR_EXPGLP; Estimation Sample: 1983m1-1997m1					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	0.0236163	0.001598	14.8	0.000	0.5681
GBR_EXPINFLPBO3	0.527728	0.07259	7.27	0.000	0.2415
GBR_EXPSEITZ	0.287384	0.03438	8.36	0.000	0.2962
sigma	0.00453826	RSS		0.0034188974	
R^2	0.867681	F(2,166) =	544.3	[0.000]**	
log-likelihood	673.504	DW		1.1	
no. of observations	169	no. of parameters		3	
mean(GBR_EXPGLP)	0.058213	var(GBR_EXPGLP)		0.00015289	
Specification Tests					
AR 1-7 test:	F(7,159) =	6.9588	[0.0000]**		
ARCH 1-7 test:	F(7,152) =	3.7945	[0.0008]**		
Normality test:	Chi^2(2) =	4.9277	[0.0851]		
hetero test:	F(4,161) =	1.7482	[0.1420]		
hetero-X test:	F(5,160) =	1.5044	[0.1913]		
RESET test:	F(1,165) =	8.9499	[0.0032]**		

TABLE 20: NON-NESTED TESTING RESULTS, SWEDEN

Modelling SWE_EXPSCBEV; Estimation Sample: 1996m1-2005m7					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	0.00881046	0.0004257	20.7	0.000	0.7942
SWE_EXPINFLPBIS	0.460251	0.08247	5.58	0.000	0.2191
SWE_EXPINFLAPES	0.291735	0.04274	6.83	0.000	0.2956
SWE_EXPSEITZ	0.116332	0.03700	3.14	0.002	0.0818
sigma	0.00194931	RSS		0.000421777819	
R^2	0.893355	F(3,111) =	309.9	[0.000]**	
log-likelihood	556.49	DW		0.882	
no. of observations	115	no. of parameters		4	
mean(SWE_EXPSCBEV)	0.0193617	var(SWE_EXPSCBEV)		3.43911e-005	
Specification Tests					
AR 1-7 test:	F(7,104) =	11.704	[0.0000]**		
ARCH 1-7 test:	F(7,97) =	4.2134	[0.0004]**		
Normality test:	Chi^2(2) =	7.8746	[0.0195]*		
hetero test:	F(6,104) =	2.2495	[0.0442]*		
hetero-X test:	F(9,101) =	7.1888	[0.0000]**		
RESET test:	F(1,110) =	28.131	[0.0000]**		

Several findings stand out from the econometric results.

First, the regressions perform badly in terms of specification statistics. This does not overly concern us in this context, as we are not attempting to *model* the quantitative expectations data. Nevertheless, this does indicate that our quantified measures are missing a certain amount of information that is present in the quantitative series.

Alternative specifications were considered in light of these issues. A lagged dependent variable was inserted in the equations; the aim of this was to control for differences in the timing of the quantitative and qualitative surveys in any given month, or the possibility that one survey targeted agents with a slower speed of adjustment to information, for instance. Indicator variables were also inserted to control for outliers (around 3 standard errors in size). It was possible to select a parsimonious and congruent specification for the UK; for Sweden, however, residual autocorrelation could not be sensibly eliminated, nor the rejection of the RESET test. The qualitative results emphasised below, however, were maintained – and therefore reinforced⁴².

Second, the constant in the equations is significant, confirming the earlier results that the quantified measures are mean-biased relative to the quantitative data. This is however not problematic for using these measures as controls in econometric work (see Chapter 2).

⁴² These further regressions are available from the author. In particular, the ranking of importance of the different methodological traditions was preserved.

Third, both regressions retain measures from more than one tradition – in fact, all three are retained in the case of Sweden. This suggests that each methodology may capture pieces of information missed by the others. We note that a homogeneity restriction was rejected in both countries by a Chi2 test. Nevertheless, we imposed this restriction for interest, notably as a possible way of compensating for potential downward bias generated by measurement error in the independent variables:

TABLE 21: NON-NESTED RESULTS WITH HOMOGENEITY RESTRICTION

Country	Quantification Tradition		
	CP	Pesaran	Seitz
United Kingdom	0.81 ***	-	{0.19}
Sweden	0.66 ***	0.34 ***	{0}

NOTE: *, **, *** respectively denote significance at the 10%, 5% and 1% levels. {} denote restricted variables.

Finally, we note that only measures in the Carlson-Parkin and Seitz traditions are retained in *both* equations above (Table 19 and Table 20). Furthermore, it can be seen that the coefficient on the Carlson-Parkin measure displays a significantly larger weight in both equations relative to the others⁴³.

The quantitative data available are insufficient to make general claims as to what combinations of measures would work best, or to speculate about the reasons underlying these results in terms of survey design or quantification approach. If a choice is to be made, one may tentatively argue that the results so far provide a nudge in the direction of

⁴³ In general, a comparison of partial r2 statistics sometimes leads to different conclusions than does a comparison of coefficients. However, we can note here the consistent performance of the Carlson-Parkin measure across both countries in terms of the partial r2.

a method in the Carlson-Parkin tradition. The following section considers an empirical application to verify whether this is confirmed in practical data analysis.

7 INFLATION EXPECTATIONS MEASURES IN WAGE EQUATIONS

It stands to reason that the modelling of wages constitutes one of the key applications for inflation expectations measures. As such, it is clear that a desirable characteristic for a candidate optimal quantification method is its ability to generate sensible wage equations, relative to theoretical sign priors. This is the test that we will implement in this section.

Our proposed equation is presented in Section 7.1 along with our sign priors. Section 7.2 provides some clarifications regarding the econometric approach taken. The results are then presented in Section 7.3.

Before proceeding, we should note here that wage equations of the type presented below form the backbone of Chapter 2, where they will be more carefully modelled, interpreted and analysed from a stability perspective. The present section *only* considers them in the context of selecting the optimal quantification method for qualitative expectations data. The results of both chapters are entirely consistent.

7.1 Model specification and sign priors

We consider the wage equation with an equilibrium correction structure proposed in Curto Millet (2004). The data used have a quarterly frequency. This equation has a long-

run solution that can be derived from a Nash bargaining model (Moghadam and Wren-Lewis, 1994), whilst the short-run dynamics are consistent with the staggered contracts approach. Full details of this are available in Chapter 2, Section 3.1.

We refrain from imposing dynamic homogeneity upon the equation ex ante, as there is little reason to believe that, in the short-run, workers will fully adjust wages to price changes. A homogeneity proposition is however much more reasonable as a feature of the long-run equilibrium implied by the model, and will be implemented. The resulting wage equation is presented below; Table 23 lists our sign priors.

TABLE 22: WAGE EQUATION STRUCTURE

$$\begin{aligned}
\Delta \ln COMPTH_t = & \alpha + a_1 INFLEXP_{t,t+4} + a_2 INFLEXP_{t-1,t+3} + a_3 INFLEXP_{t-2,t+2} \\
& + a_4 INFLEXP_{t-3,t+1} + a_5 INFLEXP_{t-4,t} \\
& + b_1 \Delta \ln CPI_t + b_2 \Delta \ln CPI_{t-1} + b_3 \Delta \ln CPI_{t-2} + b_4 \Delta \ln CPI_{t-3} \\
& + c_1 \Delta \ln COMPTH_{t-1} + c_2 \Delta \ln COMPTH_{t-2} + c_3 \Delta \ln COMPTH_{t-3} \\
& + c_4 \Delta \ln COMPTH_{t-4} + d_1 \Delta \ln U_t \\
& + e_1 \left[\ln \left(\frac{COMPTH}{P_c} \right)_{t-1} - f_1 \ln PRODIT_{t-1} \right. \\
& - h_1 \ln(1 - ITAX)_{t-1} - i_1 \ln(1 + CTAX)_{t-1} \\
& - j_1 \ln U_{t-1} - k_1 \ln BRR_{t-1} - l_1 \ln ACOV_{t-1} - m_1 \ln EPL_{t-1} \\
& \left. - n_1 \ln UDENS_{t-1} - o_1 \ln BD_{t-1} - p_1 \ln RELPMPC_{t-1} - q_1 \ln COORD \right]
\end{aligned} \tag{37}$$

COMPTH:	Hourly compensation of workers	BD:	Benefit duration index
INFLEXP:	Inflation expectations	ACOV:	Collective bargaining coverage index
PRODIT:	Trend productivity	UDENS:	Net union density
ITAX:	Income/direct taxation	EPL:	Employment Protection Legislation index
CTAX:	Consumption/indirect taxation	RELPMPC:	Ratio of import to consumer prices (Pm/Pc)
U:	Unemployment	COORD:	Bargaining coordination index
P _c :	CPI, All items		
BRR:	Benefit replacement rate		

TABLE 23: EMPIRICAL WAGE EQUATION, SIGN PRIORS

CORE VARIABLES	
$a_i \geq 0$ for all i	Higher inflation expectations should be reflected in higher wages. The magnitude of the coefficients should also be in line with the dynamics of the dependent variable.
$b_i \geq 0$ for all i	More inflation in the recent past should be expected to lead to a compensating adjustment of nominal wages, if such inflation was not reflected in the previous wage contract.
$c_i \leq 0$ or ≥ 0 for all i (variable)	Eq.(37) is constructed in first differences rather than the fourth difference $\Delta_4 W_t$. Given that our measure of inflation expectations has a yearly horizon, if the data were best described by a fourth order difference, we would have: $\Delta_4 W_t = f(x_t) \Leftrightarrow \Delta W_t = -\Delta W_{t-1} - \Delta W_{t-2} - \Delta W_{t-3} + f(x_t)$ With x_t the vector of explanatory variables. In this case, $c_i \leq 0$. This argument does not go through if we actually choose to express the dependent variable as $\Delta_4 W_t$, or if the dynamics are genuinely quarterly. In such cases, positive coefficients should be expected, as higher previous settlements could cause present ones to be inflated to preserve relativities. We let the modelling process determine the appropriate order of differencing.
$d_i \leq 0, j_i \leq 0$	Higher unemployment should lead to wage moderation in bargaining.
$e_i \leq 0$	The EqCM coefficient. It should be negative, showing adjustment to long-run equilibrium of the model, and significant to demonstrate cointegration.
$f_i \geq 0$	Higher trend productivity should have a positive effect on real wages; a plausible restriction would be the value one.

WEDGE VARIABLES	
$h_i \leq 0$	Direct/income taxes make workers demand higher wages, as they target the post-tax wage ⁴⁴ . The way this is coded, we expect a negative sign.
$i_i \leq 0$	Indirect/consumption tax rate increases widen the gap between producer and consumer wages ⁴⁵ . The way this is coded, we expect a negative sign.
$p_i \leq 0$	The net price received by the firm is affected by import prices of inputs.

INSTITUTIONAL VARIABLES	
$k_i \geq 0, o_i \geq 0$	Higher benefit replacement rates and duration make unemployment less unpalatable and lead to tougher union bargaining. See also Shapiro and Stiglitz (1984).
$l_i \geq 0$	Higher bargaining coverage of unions implies a smaller competitive part of the economy and higher wages through tougher bargaining.
$m_i \geq 0$	Employment protection legislation reinforces union bargaining power.
$n_i \geq 0$	Higher net union density reinforces union bargaining power.
$q_i \leq 0$	Higher coordination leads in the end to wage moderation due to internalisation of

⁴⁴ This shifts the Bargained Real Wage (BRW) curve up in the Carlin and Soskice (1990) framework.

⁴⁵ This shifts the Price Determined Real Wage Curve (PRW) down in the Carlin and Soskice (1990) framework.

The construction of the data series employed here is explained in full detail in Section 3.6 of Chapter 2. This step is taken as given for now.

7.2 Notes on the econometric approach

We adopt a General-to-Specific (*Gets*) modelling strategy, as advocated by Hendry (1995). The equations were checked for congruence at every stage of the selection process through tests provided by PcGive (residual autocorrelation, heteroscedasticity, autoregressive conditional heteroscedasticity, normality and the RESET test).

We take the model presented in Eq.(37) to be the GUM. We then select our model on the basis of an informal ‘Bayesian approach’, in the manner of Aron et al. (2004). Emphasis is placed on the long-run solution providing sensible results, leading us to impose the restrictions arising from economic theory detailed above if required. The lag structure of the model is essentially data-determined.

The nature of certain variables in the equilibrium correction term makes them clearly integrated, and for standard inference to hold we would require them to cointegrate. This will be signalled by a significant coefficient for the speed of adjustment term e_1 , whose distribution is non-standard but relatively close to a Student t-distribution (Hendry and Juselius, 2000).

Arguably, the model presented in the previous section could be further generalised, notably through additional delta effects. In the actual modelling, further such lags have been considered (e.g. for the unemployment variable) where appropriate.

Section 3.3 of Chapter 2 provides further details and motivation for our model selection algorithm and general econometric approach.

7.3 Results

We have carried out 64 general-to-specific model selection procedures in total⁴⁶. We emphasise that all the equations obtained as a result are congruent, except where mentioned explicitly. We will start by commenting on the summary results, before providing a discussion of the specific results in each country.

Table 24 indicates the retained lags of the expectations variable (if any) and their significance.

TABLE 24: WAGE EQUATIONS, RETAINED LAGS OF EXPECTATIONS

	FRA	ESP	BEL	DEU	GBR	SWE	ITA	NLD
EXPINFLA	X	[0***]	[0*, 2**]	X	[0,3***]	-	[0***, 2**]	X
EXPINFLPCP	X	[0***,4**]	[2***]	[4**]	[3***]	-	[0**]	[3**]
EXPINFLPBIS	X	[3***]	[2**]	[4***]	[3***]	-	[0**]	[3**]
EXPINFLPBO1	X	[4**]	[2**]	[4*]	[3***]	-	X	[0***,4]
EXPINFLPBO2	X	[4**]	[2***]	[4**]	[3***]	-	[0**]	[0***,4*]
EXPINFLPBO3	X	[0***]	[2**]	[4*]	[3***]	-	[0**]	[0***,4***]
EXPINFLAPES	X	X	[3**]	X	X	-	X	[0***]
EXPSEITZ	X	X	[1**]	[4]	[0***,3*]	-	X	[0***]
EXPGLP	-	-	-	-	[0*,3***]	-	-	-
EXPSCBEV	-	-	-	-	-	[0**]	-	-

⁴⁶ 17 of these concern the UK, as the process was carried out both for the full sample and the sub-sample over which quantitative data is available.

Notes: X denotes the absence of expectations in the estimated wage equation; - indicates that the regression was not or could not be estimated; *, **, *** denote significance at the 10%, 5% and 1% levels, respectively. Retained lags and their significance in brackets.

Several general characteristics of the findings are especially salient.

First, it can be seen that for France it was not possible to find evidence of expectations affecting the wage equation. One explanation for such a result might be that in certain countries with strong unions, these can be confident in their ability to recover the losses from any inflation occurring within the contract period ex post, and therefore bargain as if these were irrelevant in a forward-looking manner. A similar result was found by Aron, Muellbauer and Smit (2004) for South Africa. However, more detailed discussion below will suggest that other reasons are more plausible in the case of France.

Second, it was found that there was sometimes a significant amount of collinearity between the actual inflation dynamics and the inflation expectations variables. This complicates the interpretation of some of the equations. For example, the retained inflation expectations for Germany are lagged four quarters, which would seem a priori to be a rather long lag, especially given that it is isolated and not part of a lag structure for expectations. Could this variable be proxying for lagged annual inflation perhaps? Augmenting the model for that possibility did not change the modelling outcome, however. The other possibility, of course, is that other lags in expectations which would have made the interpretation easier were dropped due to collinearity.

Third, there are some instances in which several lags of expectations are retained. This has interesting economic interpretations in terms of contract staggering or in terms of “inattentive” agents only updating expectations irregularly and possibly making decisions on the basis of outdated expectations, as in Mankiw et al. (2003). The issue of multicollinearity we raised may mask the real prevalence of these situations. Lags in expectations and other variables could also be reflecting the fact that negotiations between unions and employers may be taking some time to be completed, yet be based on the information set acquired at the beginning of the bargaining round.

Finally, even at this level of generality, it can be seen that two measures are problematic in terms of their tendency to be dropped when others are retained. These are the measures based on the Pesaran and Seitz traditions – especially in the case of the former. This is a cause for suspicion, although it is not sufficient to make a case against them, as such a finding could conceivably reflect the unimportance of the ‘true’ expectations in the wage process whilst other measures are retained as proxies for inflation dynamics, for instance. This issue will be resolved by the more detailed comments below.

We now provide more detailed comments on the individual country results. The Swedish specification is briefly presented in the next sub-section to build intuition as to the properties of our specification. For the benefit of concision, the remaining set of results are presented descriptively; the finalised and selected models are offered in full detail in Chapter 2; all other results are available from the author in their entirety.

7.3.1 Sweden

Modelling SWE_D12lCOMPTH; Estimation Sample: 1983q1-2005q1					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-13.4651	2.200	-6.12	0.000	0.3391
SWE_D4lCOMPTH_1	0.301122	0.06065	4.97	0.000	0.2525
SWE_EXPSCBEV	0.288568	0.1191	2.42	0.018	0.0744
SWE_D4lCPI	0.466969	0.09037	5.17	0.000	0.2678
SWE_D4lCPI_4	0.830359	0.07022	11.8	0.000	0.6570
SWE_D4lCPI_8	0.802234	0.09408	8.53	0.000	0.4990
SWE_D4lU	-0.0157295	0.009014	-1.74	0.085	0.0400
SWE_D4lU_4	-0.0281434	0.01090	-2.58	0.012	0.0837
SWE_D4lU_8	-0.0737631	0.007703	-9.58	0.000	0.5568
I1999	-0.0221632	0.004721	-4.69	0.000	0.2319
I1995:1	-0.0225431	0.008142	-2.77	0.007	0.0950
I2002:1	0.0235681	0.007949	2.96	0.004	0.1075
SWE_lRCOMPTH_12	-0.701336	0.04742	-14.8	0.000	0.7497
SWE_lPRODIT_12	0.514141	0.07215	7.13	0.000	0.4102
SWE_lU_12	-0.0710404	0.009310	-7.63	0.000	0.4437
SWE_lACOV_12	2.54092	0.5211	4.88	0.000	0.2457
sigma	0.0075778	RSS	0.00419188561		
R^2	0.987712	F(15,73) =	391.2	[0.000]**	
log-likelihood	317.079	DW	1.8		
no. of observations	89	no. of parameters	16		
mean(SWE_D12lCOMPTH)	0.166093	var(SWE_D12lCOMPTH)	0.00383307		
Specification Tests					
AR 1-5 test:	F(5,68)	=	1.6675	[0.1543]	
ARCH 1-4 test:	F(4,65)	=	0.66158	[0.6209]	
Normality test:	Chi^2(2)	=	3.2054	[0.2014]	
hetero test:	F(27,45)	=	0.75711	[0.7774]	
RESET test:	F(1,72)	=	0.14568	[0.7038]	

The selected wage equation for Sweden has an especially interesting dynamic structure. The dependent variable modelled here is the three-year change in the hourly compensation of workers, D12lCOMPTH. The equilibrium term is accordingly lagged by three years. This is unique among the countries in our sample and is reflective of very interesting idiosyncratic features of the Swedish labour market. This finding in fact

corresponds to an institutional reality: the usual duration of collective agreements in Sweden is of three years (European Commission, 2004).

The equation exhibits a sensible long-run solution, with strongly significant productivity and unemployment effects. The speed of adjustment coefficient is found to be 70%. Long-run homogeneity for productivity cannot quite be imposed, however. In terms of institutional variables, only bargaining coverage could be retained, although it is estimated imprecisely given the limited variation that is typical of this series (in this case, a single continuous increase over the years 1990-1995).

Unemployment dynamics over the past three years are found to be strongly significant. The inflation dynamics are equally strong in this equation and the coefficients for inflation in the previous two years could (almost) be restricted to one, and remain high for the more recent changes within the past year. This could have been consistent with a bargaining framework in which expectations do not matter *ex ante* – because, for instance, unions are strong enough or otherwise confident to obtain compensation *ex post*. Nevertheless, the equation also displays a forward-looking component as evidenced by the significance of contemporaneous inflation expectations.

The specification tests show the model to be congruent in all aspects, with an excellent fit (sigma of 0.76%). That there is no significant autocorrelation when modelling a three-year change in compensation over the period of the last twenty years is rather remarkable. Furthermore, a battery of stability tests (including 1-step, breakpoint and

forecast Chow tests) confirms the stability of the model over the entire period examined. Individual recursive estimates of the coefficients reveal no break, further confirming the appropriateness of the model.

Further details on the wage equation for this country are available in Section 4.1.8 of Chapter 2. We now turn to a descriptive presentation of the results for the remaining countries.

7.3.2 United Kingdom

The wage equation for the United Kingdom models the yearly change in hourly compensation, with an equilibrium correction term lagged four quarters.

The model including the quantitative expectations EXPGLP is estimated over the period 1984q1-1996q4. Institutional variables are absent from the long-run solution, which includes productivity, unemployment and the relative price of imports to CPI. There are significant unemployment and inflation dynamics. Interestingly, expectations enter the model at two different lags – zero and three, although the former is only significant at the 10% level. This is a finding that is impossible to reproduce when using our quantified measures for modelling purposes on this reduced sample. The closest equations to that of EXPGLP over this sample period are those using EXPINFLPBIS, EXPINFLPCP and EXPINFLPBO1,2,3 – all of which have a matching long-run equilibrium and expectations at a lag of three.

The qualitative data are available over the longer sample 1983q1-2005q1. The long-run structure is identical to that of the model for EXPGLP, with the addition of a significant coefficient for either employment protection legislation or bargaining coverage depending on the equation. The great similarity in structure between the equations provides little information to form a preference for one measure over another (except for rejecting EXPINFLAPES, which is found to drop out in the modelling).

7.3.3 France

The French wage equation is expressed in terms of the quarterly change in hourly compensation, with an equilibrium correction term lagged three quarters; the model is estimated over the period 1986q1-2005q2. It is especially interesting to note that the French specification includes a significant effect for changes in working hours in the dynamics of the equation, as was already emphasised in Curto Millet (2004).

The remarkable feature of the selected French models is the absence of inflation expectations. Nevertheless, the models are all congruent with the data, and all specification tests are passed with ease. We argued earlier that in countries with strong unions, these may be able to recover the losses from any inflation occurring within the contract period *ex post*, and therefore bargain as if this was irrelevant in a forward-looking manner. Could this argument be applicable to the case of France? This seems unlikely. Although inflation dynamics are indeed present in an important way and highly significant, the same is true even more dramatically in other countries such as Sweden, which *do* exhibit a significant role for expectations. Furthermore, whilst there are good

reasons to expect important backward-looking behaviour in the case of France (e.g. indexation of the *SMIC* minimum wage; see Chapter 2, Section 4.1.2), these are by no means exceptional in our sample nor would be expected to affect the entire labour market in such a way as to completely discard a role for expectations. This therefore suggests that a different interpretation may be relevant here.

Alternatively, it may be that French unions are able to set a premium in such a way that it only needs to reflect partly the prevailing economic conditions. Perhaps consistent with this observation, the model for France includes highly significant effects for *both* benefit duration and the benefit replacement rate – a unique finding in our sample of countries. However, the unemployment coefficient is also strongly significant, refuting the idea that bargaining is completely insensitive to the economic situation.

A much more straightforward and perhaps more plausible explanation is simply that the inflation expectations data are of significantly worse quality than those of the other European countries given the measurement issue exposed in Section 2, caused by the phrasing of the French questionnaire. This would explain why France is an exception in the sample of countries considered, and suggests that special care should be taken by research using this data.

7.3.4 Spain

The Spanish model has the yearly change in hourly compensation as its dependent variable, with a long-run equilibrium term lagged four quarters. The model is estimated over the period 1987q3-2005q2.

The productivity data in the case of Spain are of somewhat lower quality than those of the other countries, as they were only available at a half-yearly frequency. This has translated into imprecise estimates in *some* of the equations and the need to impose the constraint of long-run homogeneity (easily accepted statistically by the data) to obtain a sensible outcome of the modelling process in those instances.

The outcomes of the selection process are reasonably close in all cases, except for the drop of the EXPINFLAPES and EXPSEITZ variables. The way in which EXPINFLA is retained in that equation is also dubious due to the associated absence of any inflation dynamics.

The long-run typically includes significant effects for unemployment, employment protection legislation and the relative price of imports to the CPI. The most sensible equations from an economic perspective are those for EXPINFLPCP, EXPINFLPBIS and EXPINFLPBO2. The equation for EXPINFLPCP displays an interesting lag structure of inflation expectations, without it being at the expense of some inflation dynamics. Its long-run solution uniquely includes the benefit replacement rate but suspiciously drops lagged unemployment. The equations for EXPINFLPBIS and EXPINFLPBO2 are similar and have a more conventional long-run (including unemployment) but simpler expectations dynamics. The EXPINFLPBIS model overall appears marginally more sensible, as the four quarter-lagged expectations of EXPINFLPBO2 seem rather long.

7.3.5 Belgium

The Belgian wage equation has the two-year change in hourly compensation as its dependent variable, and the equilibrium correction term is lagged eight quarters. The estimation is made over the period 1986q1-2005q2. This is consistent with the finding that the usual duration of collective agreements in Belgium is of two years (European Commission, 2004).

The resulting parsimonious equations are almost identical in all cases, with minor differences in the inflation and expectations dynamics. The long-run solution invariably involves unemployment, the benefit replacement rate, union density and the relative price of imports to the CPI. In all cases, long-run homogeneity for productivity can be imposed.

All the expectations measures seem to be performing adequately in the case of Belgium. Interestingly, EXPINFLA leads to the retention of two lags of expectations, contrary to what is the case for the other measures.

7.3.6 The Netherlands

The Dutch wage equation models yearly changes in the hourly compensation of workers over the period 1974q3-2005q2, with an equilibrium correction term lagged four quarters. The long sample period available in this instance poses particular challenges given the economic volatility present in the 1970s. This appears in the Dutch equation through the significance of a volatility index, both as a differenced variable and as part of

the long-run solution. The rationale for its presence is that in a volatile environment, inflation surprises can hurt workers' real compensation during the contract period, and they may therefore seek to insure against that possibility *ex ante*. Volatility is captured by the following moving average measure:

$$MA[abs(D4ICPI - D4ICPI_{-4})]$$

The equations paint a largely similar picture in terms of the long-run solution. The models of EXPINFLPBO1/2/3, EXPINFLAPES, EXPINFLPCP and EXPINFLPBIS all seem rather reasonable, although there is in several cases evidence of more or less significant heteroscedasticity we have found difficult to eliminate. This is not wholly unexpected given the sample period and historical circumstances.

The models for EXPINFLPBO3 and EXPINFLPBIS in particular stand out in terms of the completeness of the equilibrium correction term. In both, long-run homogeneity can be imposed for productivity and both tax terms are retained. Homogeneity can also be imposed for indirect taxation, thereby making producer prices the relevant price concept in equilibrium. The models also include significant unemployment, coordination and volatility EqCM effects, as well as a benefit replacement rate or benefit duration term. EXPINFLPBO3 has a more interesting lag structure of expectations, although this comes suspiciously at the expense of actual inflation dynamics, which can be retained in the EXPINFLPBIS equation (which seems overall a preferable choice).

7.3.7 Italy

The Italian wage equation is modelled over 1974q1-2005q2 in terms of the two-year change in hourly compensation, with the long-run equilibrium being lagged eight quarters. This is consistent with a usual duration for collective agreements of two years, as recorded by the European Commission (2004). Unlike for the Dutch wage equation, controls for inflation volatility were found to be unnecessary in this equation.

Three models are especially sensible in terms of their long-run solutions, namely those for EXPINFLPCP, EXPINFLPBIS and EXPINFLPBO1. Long-run homogeneity can be imposed for productivity; the direct taxation variable has a coefficient that is not negligible but is nevertheless insignificant at the 10% level. The equation for EXPINFLPBIS is however superior in that it gives significant roles to both unemployment and bargaining coordination. Furthermore, although all the models agree that the speed of adjustment is very low in Italy, the coefficient is often so low as to become implausible. The equation for EXPINFLPBIS provides a more reasonable (but still low) estimate of 15%. This issue is discussed further in Section 4.1.1 of Chapter 2.

In dynamic terms, the wage equation displays highly significant structures of inflation and unemployment dynamics. The EXPINFLPBIS term itself is highly significant, although rather small quantitatively (around 4%).

7.3.8 Germany

We model the half-yearly change in hourly compensation with an EqCM term lagged by two quarters over the period 1991q3-2005q2.

The selected models have a remarkably similar structure, even in the two cases where the expectations variable is dropped (EXPINFLAPES and EXPINFLA), and all satisfy our specification tests with ease. The long-run is characterised by very strong effects from both taxation variables, as well as the presence of the benefit replacement rate and bargaining coverage. Unemployment has an insignificant long-run effect and instead tends to appear (modestly) as part of the dynamic structure of the equation.

Despite a great similarity in outcomes, the wage equation for the measures EXPINFLPBIS and EXPINFLPBO2 appear especially sensible, as they display larger speeds of adjustment – which seem more reasonable given a usual duration of collective agreements of 1-2 years in Germany (European Commission, 2004) – as well as highly significant productivity effects. The equation for EXPINFLPBIS in addition displays a marginally insignificant union density variable at the 10% level, which is not apparent in other equations.

8 CONCLUSION

This chapter has presented a range of competing quantification methodologies for qualitative survey data on inflation expectations, with the aim of determining the optimal approach to this problem. These quantification methodologies arose from three distinct traditions in the literature, identified respectively with Carlson-Parkin, Pesaran and Seitz. The theoretical discussion highlighted shortcomings in all the approaches, making an empirical examination decisive in the selection of the optimal method.

The methodologies considered were assessed on a number of criteria. First, the competing approaches were contrasted in terms of their predictive ability. Although we argued that this exercise was in general inappropriate for the purposes of selecting the method closest to the ‘true’ expectations, the question of predictability is of interest in its own right and has often been the focus in the literature. A number of descriptive statistics were considered: the RMSE, the decomposition of the MSE and the proportion of correctly predicted turning points. The measures EXPINFLPBIS, EXPINFLPCP and EXPINFLAPES proved to be the most successful in this respect – although this success is very relative, as they were found to underperform the naïve expectations benchmark considered.

Second, the methodologies were compared in terms of their ability to reproduce the quantitative data available for the United Kingdom and Sweden. EXPINFLAPES and EXPSEITZ performed better than the other measures in RMSE terms due to relatively low bias. The picture was less clear in what regards the ability to indicate the direction of

change of the underlying variables, as the EXPINFLPBO measures also performed well here. Non-nested tests selected for both countries a combination of measures which included one in the Carlson-Parkin tradition and one in the Seitz tradition, with the Pesaran measure also being retained in the case of Sweden. This suggests that each tradition captures pieces of complementary information, although the largest coefficients were found to be weighing the information from the Carlson-Parkin approach.

Finally, we investigated the results of using these measures in practical econometric modelling in the economically interesting context of wage equations. This selection criterion was much more decisive, with the measure EXPINFLPBIS demonstrating the best performance consistently and EXPINFLPBO2 also offering sensible outcomes overall relative to our sign priors, especially in regard to the long-run equilibrium. Both measures are modifications of the Carlson-Parkin tradition suggested here, the former of which had already been used in Curto Millet (2004).

Interestingly, it was found that in the case of one country – France – inflation expectations did not seem to matter. Although potential institutional explanations were considered for this result, it is very plausible that measurement error caused by the difference in wording of the French questionnaire bears some responsibility, suggesting that special care should be taken in the quantification and use of these data. Apparently sound equations in terms of the specification tests were also found in other cases where the expectations terms were dropped, suggesting that these are not “critical” to a modelling exercise in this particular respect. Their importance is more subtle, however,

as the equations including inflation expectations terms often had greater economic interpretability and the presence of such terms helped in achieving a more complete description of the long-run equilibrium.

This study therefore provides support for the use of measures in the Carlson-Parkin/Batchelor and Orr tradition that use information on perceptions in the quantification and further recommends allowing for the expression of long-run trend changes in the underlying thresholds. Further investigation of more appropriate structural alternatives in the Seitz tradition may also be profitable, as would further empirical exercises contrasting these measures in other contexts.

A more important subtext to this chapter, however, is the necessity for the economics profession to notice the maturity of survey-based expectations measures, both in terms of the quantification procedures and their usability due to the sample sizes now available. These measures have the potential to make important contributions to econometric modelling and more generally to economic debates – including, but not limited to, that surrounding the mantra of rational expectations and credibility theory.

This study hopes to establish these claims in the following chapters.

CHAPTER II:

The Impact of European Monetary Unification on Labour Market Flexibility

“Continued wage moderation in recent years partly reflects high unemployment. But it could also suggest the possibility of a structural change in the area’s wage setting ...the advent of the euro may also have affected wage setting in the euro area, as there is no longer any possibility for sharp wage increases being accommodated later on by currency realignments”. (OECD, 2002a, page 22.)

1 INTRODUCTION

1.1 Motivation and chapter plan

EMU has been a reality for 11 European Union countries since 1 January 1999, and for Greece since 1 January 2001. The remaining “old” EU countries (the UK, Denmark and Sweden) have so far decided to opt out, whilst newer EU members are currently queuing to enter this monetary union⁴⁷.

Both the sign and magnitude of the economic impact of membership are hotly contested issues (Forder and Huhne, 2001), a debate that is usually framed within the theory of Optimum Currency Areas (henceforth OCA; see Mundell, 1961). Ideally speaking, an OCA is such in size and structure so as to equate the marginal benefits to the marginal costs of altering its shape. We will here focus on the cost side of the assessment.

The two sticking issues on which the *theoretical* case⁴⁸ against EMU membership critically rests are, first, the problems caused by asymmetric shocks and second, those due to asymmetric responses to symmetric shocks. Both imply differential economic developments across countries that would optimally require differing monetary policies.

Rather than debating the likelihood of asymmetric shocks (e.g. Krugman, 1993; Viñals and Jimeno, 1996) or the broad empirical idea that OCAs are endogenous (Frankel and

⁴⁷ Slovenia was admitted into the Eurozone on 1 January 2007. Cyprus and Malta aim to adopt the euro on 1 January 2008, and Slovakia the following year.

⁴⁸ As opposed to the *practical* case, which criticises – rightly or wrongly – the design of the EMU framework as it stands (e.g. the ECB’s independence, its two-pillar monetary policy strategy, the Stability and Growth Pact etc.)

Rose, 1998), this study examines the potential for stabilisation provided by labour market flexibility. There are several justifications for this focus.

First, there exist significant structural and institutional differences across Europe. Maclennan et al. (1998) have made this clear with respect to housing and financial market institutions; structural differences in labour market institutions are similarly well-documented (e.g. Nickell, 1997; Cadiou and Guichard, 1999a) and further avenues for asymmetric developments are not hard to come by (e.g. trading structure, government financing). One therefore expects that monetary transmission itself will display asymmetric features. VAR evidence to the contrary looks unpersuasive in this light, and can be criticised as deficient for omitting relevant variables (e.g. Gerlach and Smets, 1997; Barran et al., 1997) or being subject to the Lucas critique (e.g. Britton and Whitley, 1997). In other words, non-negligible asymmetric tensions are clearly on the cards – at least in the short run – and national adjustment mechanisms are an important issue.

Second, other adjustment mechanisms (labour mobility, inter-regional fiscal transfers) are possible but have their own problems and limitations (Buiter, 1995; Obstfeld and Peri, 1998), and given their lack of relevance will be ignored here (HM Treasury, 2002). Labour market flexibility thus stands out as the most desirable and realistic adjustment mechanism to reconcile asymmetric developments in a way that would avoid the generation of unemployment.

Labour market flexibility is a multifaceted concept. We will here be particularly interested in its wage aspects. For instance, it is often hoped that bargaining partners realise that wage increases in excess of productivity growth under monetary union will likely translate into a deterioration of competitiveness and employment, without the possibility of a bail-out through currency devaluation. This in turn should lead to a greater emphasis on preserving competitiveness in wage claims, possibly translating econometrically into a structural break in wage equations. This is an example of the type of evidence we are looking for in this chapter.

Economic theory allows a number of conjectures as to the prospects for finding enhanced wage flexibility as a consequence of EMU. This theoretical material is reviewed in Section 2.1. We have however more on which to base our expectations, given the existence of relevant historical precedents to EMU, most notably the EMS system established in 1979. More recent empirical studies have also sought to anticipate the impact of EMU itself. This empirical material is surveyed in Section 2.2.

Section 3 then presents the theoretical basis for the wage equation developed in this study. This is essentially a synthesis of the Layard et al (1991) bargaining approach and the staggered contracts literature. The empirical application considered here suggests particular care ought to be taken before confronting the data. Indeed, EMU constitutes a momentous regime shift – or, as Burda (1999) put it: “*if the Lucas Critique has any relevance at all, it had better be here and now*”. One of the most significant shifts that can be expected with EMU relates to the expectations mechanism, given the change in

the monetary framework and the début of the European Central Bank (ECB). Controlling for this will be crucial if we hope to detect changes in the structure of wage setting. Accordingly, the inflation expectations series selected in Chapter 1 as “optimal” is a key component of our wage equation structure. Similarly, a database of institutional variables for use in this study is carefully compiled, and their influence controlled for in our equation. The data and the econometric approach taken are also presented in Section 3.

Section 4 presents EqCM wage equation results for France, Germany, Italy, Spain, Belgium, the Netherlands, the United Kingdom and Sweden. The results are analysed and form the basis for a pooled mean group specification. Section 5 concludes.

1.2 The meaning of labour market flexibility

Labour market flexibility encompasses a vast range of issues, and its meaning has evolved over time (Lagos, 1994). HM Treasury (2003) distinguishes six concepts underlying this term, namely: relative wage flexibility (across labour market segments), real wage flexibility, nominal wage flexibility, geographic labour mobility, employment flexibility and functional flexibility.

Our study will focus on real and nominal wage flexibility as these are the aspects that matter most in the EMU debate. Real wages are said to be flexible if they are responsive to unemployment in a way that contributes to restoring equilibrium. Nominal wage flexibility is clearly helpful for this aim and contributes to securing international competitiveness in the face shocks, thus performing a stabilisation function.

In practical terms, in a model of the nominal wage in equilibrium correction form (Hendry, 1995), we will be especially concerned with the long-run elasticity of the real wage with respect to unemployment and with the speed of adjustment to equilibrium. ‘Large’ values of these coefficients should ensure a rapid return to equilibrium in the face of shocks.

Also critical in the EMU context is the productivity orientation of wage policy, as reflected by the long-run elasticity of real wages with respect to productivity (Lauer, 1999). Homogeneity for this coefficient would notably discard labour market shocks caused by an over-expansionary wage policy, and is a requirement for a successful monetary union.

1.3 Suggestive evidence of labour market changes linked to EMU

There are several indications that large changes are at hand, and that a different set of forces may be acting on the labour market following the introduction of EMU.

First, it appears that social partners are keenly aware of the stabilisation challenges presented by EMU. This is visible, for instance, in reports from the *Conseil Central de l'Economie*, a central organisation of social partners in Belgium (CCE, 1998). An especially clear manifestation of such concerns took place in Finland, where a system of ‘EMU buffer funds’ was put in place by central labour market organisations, who also administer it (Calmfors, 1998, 2001; Alho, 2000). The idea is to accumulate such funds in good times – through higher employer social security contributions – so as to draw

upon them in recessions, thereby avoiding the need to raise employer contributions counter-cyclically (as was the case in the face of the early 1990s recession). Increased concerns about competitiveness at the doorstep of EMU may also be part of the motivation for Belgium's law on the "preventive safeguard of competitiveness".

Second, there are some direct signs that EMU might have had consequences on the operation of labour markets. Trade union concerns with possible downward competition have been especially visible through initiatives aiming at improved cross-border coordination, such as the "Doorn declaration" by German and Benelux unions (1998) or sectoral responses such as the "European coordination rule" of the European Metalworker's Federation (EIRO, 2000a).

In turn, these observations highlight the need to control for institutional factors to identify flexibility developments linked with EMU.

2 THEORETICAL BACKGROUND AND LITERATURE REVIEW

2.1 Theoretical foundations and expectations

The theoretical literature supports a number of diverging attitudes with regard to the impact of EMU on wage flexibility.

One purely theoretical view of monetary unification regards it as little more than an exercise in sophisticated veil manufacturing. According to this view the monetary regime is largely irrelevant for the functioning of the real economy – and hence so is EMU. This diagnosis could for instance be reached from a Real Business Cycle standpoint, ignoring nominal rigidities. It is because of a belief that the latter are of consequence that the issue matters at all (Burda, 1999). Wage rigidity is well-documented in survey evidence (Bewley, 1999) and actual raw data. Recently, Dessy (2002) examined the European Community Household Panel (ECHP) and found evidence of a pronounced spike at zero in the distribution of one-year nominal wage changes for all European countries – a finding that is also typical of other surveys⁴⁹. Ignoring this fact of life does not seem helpful for our purposes.

Another theoretical approach stems from Heckscher-Ohlin trade theory and more concretely, from the factor price equalisation theorem. That is, the result that under incomplete specialisation and homogeneous technology, the equalisation of product prices through trade leads to factor price – and therefore wage – equalisation. Trade *de facto* substitutes for factor mobility, with the consequence that shocks will be spread across the currency area. The assumptions behind this result are obviously very stringent,

⁴⁹ See, *inter alia*, Card and Hyslop (1997) for evidence from the PSID survey in the US.

so that the relevant question has to be whether they hold *approximately*. In this respect, although European product market integration has clearly risen in recent years and can be expected to increase due to EMU itself (Ruehmann, 2002), the area of literal applicability of Hecksher-Ohlin theory still remains distant.

A more common attitude is that EMU will lead to an increase in wage flexibility, and that this matters for reasons familiar from the OCA literature. This belief is often held in a heuristic fashion, as it is not obvious to substantiate theoretically due to the absence of a generally accepted explanation for money wage rigidity (Calmfors, 1998, 2001).

Several arguments can nevertheless be advanced.

2.1.1 The sources of nominal wage rigidity and their implications for EMU

Keynes, social norms and coordination failure

The now classical explanation for nominal wage rigidity was provided by J.M. Keynes (1936). He argued that employees are concerned by wage relativities, leading them to oppose money-wage reductions unless cuts are coordinated – thus generating downward wage rigidity. Survey evidence in the United States seems to contradict this idea, as employees are found to have “*little systematic knowledge of pay rates at other firms*” (Bewley, 1999). However, Swedish survey evidence from personnel managers does attribute an important role to inter-firm comparisons (Agell and Lundborg, 1999). Boeri et al. (2001) suggest that this discrepancy may be due to the higher degree of unionisation found in Sweden and the role of unions in spreading pay information,

making this factor potentially relevant for other European economies. It is difficult to see how EMU would affect this factor of rigidity.

Keynes' explanation also nests two modern ideas of the sources of rigidity.

First is the *behavioural* nature of Keynes' argument (employees "care" about relativities). More recent literature has emphasised that wage cuts may conflict with social norms such as "fairness", affecting morale and ultimately cutting productivity (Akerlof, 1982; Akerlof and Yellen, 1990; Bewley, 1999). Again, it is difficult to imagine the extent to which EMU could impact these considerations⁵⁰.

Second, Keynes' argument points to the idea of *coordination failure*. Wage cuts are strategic complements, potentially leading to a situation with multiple equilibria (Ball and Romer, 1991). With very large or very small shocks, adjustment and non-adjustment respectively are not in doubt for the individual firms. In the grey area in between, adjustment may be individually worthwhile only if it is general. Presumably, increased macroeconomic variability would lead to more adjustments in this environment (Calmfors, 1998), although this depends very much on the definition of 'large' in the previous sentence.

⁵⁰ Keynes regarded wage rigidity as a fact of life that institutional reform could do little to change. He expressed such views even in the context of the Great Depression, in the Committee of Economists appointed by the government. See Moggridge (1992) and Hicks (1974).

The contracting approach

Contractual explanations provide an alternative view; these hold that rigidity arises from the negotiation of wages at discrete points in time, and that agents choose contract length optimally so as to equate its marginal costs and benefits.

Negotiations consume the scarce time of managers and union officials and risk ending in costly industrial conflict for both parties. On the other hand, long contracting periods lead to additional variations in output, employment and profits given the inability to adjust wages in the face of shocks. This source of rigidity takes its full importance when considering that wage contracts are *staggered* (Taylor 1979, 1980). The relative wage structure then comes to play a part in the contracting process and can considerably increase the adjustment lag in the economy.

Insofar as macroeconomic variability increases under EMU, the marginal cost of wage-contract duration increases, providing an incentive to reduce contract length. This intuition has been more rigorously modelled in Calmfors and Johansson (2002), who solve numerically a simple model for the choice of optimal contract duration. Under their assumptions, EMU creates incentives for significant reductions in contract length, leading to a large increase in price variability while failing to fully counteract the increase in employment variability⁵¹.

⁵¹ Similar results hold when the wage flexibility choice is expressed as one over the optimal degree of indexation (i.e. EMU strengthens the incentives for indexation).

A closely related view of the causes of rigidity was recently spelt out in Mankiw and Reis (2002), who posited that information diffuses slowly in the economy. This idea is modelled by letting a proportion λ of firms in each period obtain new information and compute new optimal paths for prices on this basis, whilst the remaining firms continue to base their decisions on outdated information⁵². In our context, inertia would then effectively arise because some wage setters are basing their current negotiations on outdated expectations. Insofar as the microfoundations for such behaviour arise from a trade-off between some costly activity (e.g. information acquisition) and its benefits in wage setting, higher macroeconomic variability in EMU should provide further incentives for information acquisition.

The bargaining approach

Holden (1994) placed the focus on the bargaining game behind wage setting and noted that the options facing bargaining actors are not limited to agreement or costly conflict. Indeed, a common practice is for production to continue with the same conditions as under the previous contract; in some countries such a practice is mandated by law. A consequence of this may be nominal wage rigidity, as moderate changes in nominal demand will not make labour market conflict worthwhile for either party. Higher demand fluctuations under EMU may thus lead to more frequent adjustments on this count.

⁵² See Section 6.2 of Chapter 3 for a more detailed examination of this model.

Overall assessment

Notwithstanding the above, Calmfors (1998, 2001) pointed out that there are also reasons to expect higher nominal wage rigidity, should EMU usher Europe into an era of lower inflation than hitherto. This would indeed provide an incentive for longer wage contracts to be signed, for instance.

There are some empirical indications of the significance of this factor. Using US PSID data, Card and Hyslop (1997) found that the spike at zero in the distribution of hourly wage changes increases from 7% in a 10% inflation environment to 15% when inflation falls to 5%, both for salary and wage earners. These are however large inflation changes, and given the general emphasis placed on price stability in the Europe of the 1990s and the Maastricht convergence criteria, the change in this respect should be minimal – not to mention the fact that average Eurozone CPI inflation has actually been slightly higher since the ECB took over.

One possible danger is that there will be an asymmetric development of wage flexibility (Calmfors, 2001). Some of the explanations highlighted above focused on downward wage rigidity, which suggests that this may be stronger than the upward variety. Recent asymmetric indexation clauses in Spain and Finland may be seen as a reflection of this. EMU could then lead to a development of flexibility that is skewed upwards. Temporary upturns could then result in wage increases that are later problematic to reverse, thus aggravating downswings.

Our overall assessment, however, is that on the basis of the theories outlined above, it is quite possible that EMU will lead to increased wage flexibility, although the strength of the effect is in doubt. Empirical verification of this is a more complicated matter, and will be dealt with econometrically in this study.

2.1.2 Real wage rigidity: sources and expectations

Mainstream economic theory provides several explanations for *downward* real wage rigidity⁵³. This may be due to the impact of real wage levels on the mix of applicants for vacancies when reservation wages are positively correlated with ability – a problem of adverse selection (Akerlof, 1970; Weiss, 1980). Similarly, the level of real wages may play a role in persuading workers to exert effort under conditions of imperfect monitoring⁵⁴ – a problem of moral hazard (Shapiro and Stiglitz, 1984). These stories are not of great relevance here insofar as they do not affect the *responsiveness* of wages to economic conditions, other than to place a floor to their possible levels.

Of greater relevance is Burda's (1999) speculation that real wages would become more flexible after EMU, reflecting his expectation of greater nominal wage flexibility coupled with increased price rigidity. One reason for the latter is that cost pressures will primarily arise from Euroland labour markets, thus decreasing the importance of the exchange rate for pricing decisions. For small open economies, devaluations lead to marginal cost increases that provide incentives for price adjustment, a mechanism whose relevance was decreased by the formation of the Eurozone.

⁵³ These are arguments about real wages if we assume workers do not suffer from money illusion.

⁵⁴ Speaking more rigorously, it is the unemployment thereby created that ultimately provides the incentives.

Overall, it seems reasonable to expect that real wage flexibility will increase in line with nominal flexibility.

2.2 Guidance from the EMS and other pre-EMU empirical studies

Two cautionary warnings are in order before forming an expectation of labour market developments under EMU from previous fixed exchange rate periods – most notably the EMS experience.

First, realignments were possible within the EMS, thereby decreasing the need for a break in wage behaviour. Indeed, small realignments frequently took place before the dawn of the “hard EMS” era in 1987, after which no (voluntary) realignments were implemented (Eichengreen and Wyplosz, 1993). In August 1993, following a series of speculative attacks, the EMS widened its fluctuation bands to $\pm 15\%$ ⁵⁵. As a matter of fact, the implicit bands were actually much narrower than the official bands (Labhard and Wyplosz, 1996), and overall the post-1993 system was characterised by considerable stability. Nonetheless, it clearly falls short of the EMU experiment⁵⁶.

Second, the presence of some exchange rate uncertainty, less-than-full capital market integration and possibly lower “transparency” under the EMS system imply a smaller impact than EMU on trade and product market integration. This suggests relatively less pressure was being exerted on wages and in fewer sectors. More generally, EMU being

⁵⁵ Except for the Dutch Guilder/German Mark exchange rate, which operated within $\pm 1\%$ bands.

⁵⁶ We note the devaluations of the Peseta and the Escudo, by 7% and 3.5% respectively, in 1995. This constituted the last pre-EMU change in the parity grid.

unique, the Lucas Critique applies with particular force (Lucas, 1976). It is likely nonetheless that valuable lessons may be gathered from previous studies.

2.2.1 Country case studies

A number of country case studies have examined the link between exchange rate arrangements and the flexibility of wages and prices (Eichengreen, 1998).

Two studies of small, open economies provide some encouragement. Focusing on Ireland, Kremers (1990) finds evidence of credibility effects from the EMS when modelling inflation expectations and using rational expectations: the coefficient on ERM inflation gains weight, as the one on UK inflation becomes insignificant, and lagged competitiveness becomes significant. Kremers further suggests that this in turn moderated wage settlements (see also Giavazzi and Pagano, 1988). Such effects underline the importance of controlling for expectations in our work. For Austria, Hochreiter and Winckler (1995) argue that the hard currency strategy followed turned the country into an OCA with Germany, and document how this went hand-in-hand with real wage flexibility, allowing the fixed exchange rate to be maintained in the face of asymmetric shocks.

Studies on the disinflationary experiences of two larger countries however provide little support. For Italy, Gressani et al (1989) showed that the EMS played a crucial role in disinflation over the 1980s by using conditional simulation on the Bank of Italy's econometric model but find small inflation precommitment effects. In the case of France, Artus and Salomon (1986) found that German inflation entered significantly in the

French nominal wage equation, but that its effect did not vary much over the 1980s as would be required by credibility theory. The reduced influence of domestic variables seems to point in the opposite direction, but the wage equation remains stable.

2.2.2 Multi-country econometric studies

In the context of the EMS, Artis and Ormerod (1991) distinguished between credibility effects and changes in the structure of the wage setting mechanism, proceeding in two steps. First, they modelled inflation expectations as an AR(4) process including relevant dummies and argued that their results showed credibility effects in the cases of France and Italy (German inflation switched to significance in those equations over the sample period). They then used the predictors from this first step in a quarterly wage equation with equilibrium correction form:

$$\Delta_4 l w = \alpha + \beta_1 \Delta_4 l w_{-1} + \beta_2 E \Delta_4 l p + \beta_3 X D_t + \beta_4 l(w/p)_{-i} + \beta_5 D_i + \beta_6 t \quad (38)$$

Where ‘l’ denotes the logarithm of variables, w and p are wages and consumer prices respectively, XD is a proxy for excess demand (unemployment), t denotes a time trend and D_i is a sequence of dummy variables based on institutional information. Their coefficients were essentially stable (although several years of “disturbance” were excluded in the cases of France and Italy) except in the case of Belgium, which exhibited a break in 1981. This reflected the disappearance of a previously significant time trend and is interpreted by the authors as an “EMS effect”. The dynamics present are however somewhat restrictive, miss important variables such as productivity and the tax wedge, and the estimation is carried out with an imposition of dynamic homogeneity, which is

arguably an excessively strong assumption. Furthermore, institutional changes are controlled for through simple dummies⁵⁷, which can make one nervous when attributing the break found to an ‘EMS effect’. Indeed, this was a period of intense institutional change in Belgium, including a temporary elimination of indexation in 1981 and the introduction of a ‘competitiveness norm’ in 1983. Whether such reforms can be credited indirectly to ‘EMS pressure’ would necessitate a separate argument.

In a follow-up paper, Artis and Ormerod (1996) reconsidered their results and extended their sample but maintained the same wage equation. They excluded the first oil shock years from the analysis, arguing that they had biased the earlier results; this weakened the evidence for a credibility effect, as only a marginal strengthening of the German inflation term in the French inflation expectations equation remained. The authors also decided to modify their dividing date between the ‘pre-EMS’ and ‘post-EMS’ periods from 1979 in their earlier paper to 1983 – arguing that the real watershed in the counter-inflationary evolution of the EMS only came at that point. The equilibrium correction wage equations once again exhibited no strong evidence of coefficient changes except, as before, in the case of Belgium. The authors accordingly concluded that there was overall little evidence that the EMS had forced through structural changes in labour markets, except for Belgium – although a detailed argument to support this interpretation of the break was not provided.

⁵⁷ Institutional information was considerably more scarce at the time than it is now.

Anderton and Barrell (1995) considered a wage equation derived from a theoretical framework based on wage bargaining in the context of ten European countries. This is a direct extension of earlier work presented in Anderton et al. (1992)⁵⁸. Inflation expectations were instrumented using the Pagan (1984) procedure. Their model is of the equilibrium correction type, and a general-to-specific modelling strategy was adopted. The authors used quarterly data and the following specification:

$$\begin{aligned} \Delta \log(COMP / WORK) = & a + b_1[\log(COMP(-1) / WORK(-1)) \\ & - \log CEDITAX(-1) - APROD(-1)] + b_2U(-1) + b_3PE + \\ & b_4TWEDGE(-1) + [\Delta \log COMP(-i), \Delta \log EMP(-i), \\ & \Delta \log HRSI(-i), \Delta \log CED(-i), \Delta U(-i), \Delta TWEDGE(-i)] \end{aligned} \quad (39)$$

COMP:	Compensation of employees, whole economy	HRSI:	Hours index, manufacturing
WORK:	Person hours worked	CED:	Consumers' Expenditure Deflator
EMP:	Employees in employment	CEDITAX:	CED, adjusted for indirect taxes
APROD:	Backward moving average of output per person (in logs)	TWEDGE:	Tax wedge for direct (t_2) and indirect (t_3) taxes
PE:	Expected inflation one quarter ahead		$[\log((1 - t_2)/(1 + t_3))]$

There was evidence of significant structural change for the Netherlands (1979), Italy (1982) and Spain (1986, towards the end of the sample) only. In Italy, the change made both the unemployment and expectations coefficients significant, coinciding with the dismantling of backward indexation, confirming the earlier results of Anderton et al. (1992). In the Netherlands, the absolute value of the unemployment coefficient increased. The authors made an assumption of rationality and constrained direct and indirect taxation to have the same effect; this assumption however does not hold empirically, as

⁵⁸ Note however that the specification in Anderton et al. (1992) did include a unionisation variable which is missing from the later study.

will be verified later in this chapter. Furthermore, as in Artis and Ormerod's work, institutional variables were omitted from the analysis and expectations were only controlled for through a proxy assuming "rationality". As a result, this makes it once again difficult to claim the identification of instability as an EMS effect; in any event there is little evidence of this, although the authors did speculate that the EMS may have facilitated reforms in Italy.

Overall, despite a few positive results, the work on the EMS period can provide little clear-cut support for the contention that labour markets are much affected by changes in the exchange rate regime, at least over the time scale considered. Nonetheless, the exceptional nature of EMU and improvements in methodology leave the issue open.

TABLE 25: STUDIES DEALING WITH THE EMS AND OTHER PRE-EMU EXPERIENCES

Study	European Countries	Sample	Econometrics	Results
Kremers (1990)	Ireland	1965q1-1986q4	IV estimation	Credibility effects
Gressani et al. (1988)	Italy	1979q4-1986q4 (but varies)	OLS and conditional simulation	EMS crucial for disinflation, but pre- commitment policy to target inflation rates had little effect
Artus and Salomon (1986)	France	1979q1-1991q3	OLS	Stable nominal wage equation
Hochreiter and Winckler (1995)	Austria	-	-	Argue real wage flexibility allowed Austria to maintain its fixed exchange rate in the face of asymmetric shocks.
Artis and Ormerod (1991)	FR, IT, NL, BE, DE	1972q1-1989q2	2SLS	Credibility effects: FR, IT. Break: BE (1981)
Artis and Ormerod (1996)	FR, NL, BE, DE	1975q1-1991q4	2SLS	Credibility effects: FR. Break: BE(1983)
Anderton et al. (1992)	DE, FR, UK, IT	1970q1-1991q1 (but varies)	IV estimation	Structural change: IT (1982)
Anderton and Barrell (1995)	DE, FR, IT, UK, NL, BE, DK, AU, IR, SP	197?q?-199?q? (unspecified)	IV estimation	Structural change: NL (1979), IT (1982), ES(1986)

2.3 EMU and labour market flexibility: econometric studies

More recent literature has directly focused on the degree of market flexibility in view of EMU. Several approaches have been rehearsed.

First, Lo Cascio (2001) followed the VAR tradition, applying her model to eleven European countries on quarterly data. She regressed the standardised effects of shocks on output – obtained from a three-variable structural VAR (GDP, CPI, short-term interest rate) – on a set of variables including measures of real wage rigidity and hysteresis, and found significant positive coefficients. The conclusion matches economic intuition and theory, and underlines the role labour markets can play under EMU. However, given the few variables used in the VAR and the time sample over which this reduced form was estimated, the validity of the coefficients obtained can be questioned – a recurrent criticism of VARs when structural change is ongoing.

A second approach has more appealingly relied on estimating wage equations in equilibrium correction form.

Lauer (1999) highlighted the history of high wage inflation and the associated competitive devaluations of Southern European countries. The author's descriptive analysis showed a tendency towards wage moderation in the years preceding EMU, which she argued could be accelerated by monetary union. She estimated a wage equation based on a simple ADL model in error-correction form, and found that a

dummy D_{fix} characterising periods of relative exchange rate stability dampened wage increases. Her model, based on annual data, was as follows:

$$\begin{aligned} \Delta w_t = & a_0 + a_1 \Delta p_t + a_2 \Delta \gamma_t + a_3 \Delta U_t + a_4 \Delta w_{t-1} \\ & + \alpha(w_{t-1} - a_{00} - a_{10} p_{t-1} - a_{20} \gamma_{t-1} - a_{30} U_{t-1} - D_{\text{fix}_{t-1}}) + \varepsilon_t \end{aligned} \quad (40)$$

All variables except the unemployment rate are in logarithms and γ denotes labour productivity. Her estimates showed significant money illusion in the long run solution, which may be due to omitted variable bias from not controlling for the tax wedge in her equation. Furthermore, actual labour productivity was used rather than the *trend* thereof, which would arguably be more appropriate. Structural change was controlled for by simple dummies without considering breaks in slopes. However, although her particular estimates are open to criticism, her descriptive argument does correctly highlight Southern European countries as an especially relevant population for EMU-related structural breaks – a priori, certainly more so than countries such as the Netherlands which have maintained a credible peg with the DM for over two decades. Significant effects in the case of the latter would be strong evidence for the claim that monetary union has added value with respect to fixed exchange rate regimes.

Scheremet (2000) considered an error-correction model over the period 1979-1997 for the *acceleration* of wages, which included price, trend productivity and unemployment variables. The degree of differencing was selected to ensure co-integration, and the specification allowed for hysteresis effects in unemployment – although these will tend to be confused with short-run unemployment dynamics. Concretely, Scheremet used the following equation:

$$\Delta^2 w_{t,i} = \alpha'_{0,i} + \alpha_{1,i} \Delta^2 p_{t,i} + \alpha_{2,i} \Delta^2 q_{t,i} + \alpha_{3,i} (1 - \gamma_2) U_{t,i} + \alpha_3 \gamma_2 \Delta U_{t,i} - \lambda_t [\Delta w_{t-1,i} - \alpha_{11,i} \Delta p_{t-1,i} - \alpha_{12,i} \Delta q_{t-1,i}] + \varepsilon_{t,i} \quad (41)$$

All variables except the unemployment rate are in logarithms. As in Lauer (1999), the tax wedge term was omitted, which may help explain some of evidence of non-neutralities. The author found significant heterogeneity among the countries considered in terms of their long-run productivity effects, price effects, unemployment effects and hysteresis, as confirmed by Wald tests of coefficient restrictions. Structural breaks were detected for Portugal (1988) the UK (1988) and the Netherlands (1985), the former two being of minor importance according to a Chow test. The Dutch break notably implied increases in the magnitude of the unemployment and productivity coefficients. The author further argued that the wage process changed in Italy in 1993 due to the elimination of the Scala Mobile⁵⁹ and increased bargaining coordination, despite insignificant test results. However, there appeared to be little overall evidence of appropriate labour market changes ahead of EMU.

Barrell and Dury (2001) considered a quarterly equilibrium correction specification for the real producer wage which included both forward- and backward-looking *consumer* price deflator (CED) terms; these were assumed to matter only in the short run. In contrast, the long-run was specified in terms of the real *producer* wage (W/P) and featured unemployment and productivity (with homogeneity imposed for the latter):

⁵⁹ The Scala Mobile was the Italian wage indexation system. Please refer to Section 4.1.1 for a discussion of these arguments.

$$\Delta \ln\left(\frac{W}{P}\right) = \alpha + \sum_{i=0}^4 \beta_i \Delta \ln CED_{t-i} + \sum_{i=1}^4 \gamma_i \Delta \ln\left(\frac{W}{P}\right)_{t-i} + \left(1 - \sum_{i=0}^4 \beta_i + \sum_{i=1}^4 \gamma_i\right) \Delta \ln CED_{t+1} \\ + \lambda U_{-1} + \varepsilon \left[\ln\left(\frac{W}{P}\right) - \ln(Prod) \right]_{t-1} \quad (42)$$

Inflation expectations were obtained through instrumenting future inflation, and a general-to-specific modelling strategy was adopted. The imposition of both dynamic and long-run homogeneity may be considered unnecessarily restrictive. Furthermore, the assumed lack of relevance of the CED in the long-run is a strong assumption on the weakness of unions that one might hesitate to impose given the sample period covered. We also note that the model omits tax wedge terms, which is consistent with their insignificance in Anderton and Barrell (1995); however, that finding was suspect due to the constraint imposed on the tax variables discussed earlier. Consequently, their omission from this equation may have biased the results. As in Scheremet (2000), countries were found to differ significantly, notably in their speed of adjustment, unemployment effects and degree of forward-lookingness – although some country groupings were statistically acceptable. The authors found evidence of structural change in Germany after reunification (1991), with the moderating role of unemployment increasing, for Italy (1993) – explained in terms of the abolition of the *Scala Mobile* – and for Spain (1987) after it joined the ERM.

Finally, Cadiou et al. (1999b) considered a quarterly model for the wage rate in equilibrium correction form including the GDP deflator (P), consumer prices (Pc), productivity (Q/N), unemployment (U) and a dummy for ERM participation (ERM):

$$\begin{aligned}
\Delta \ln(W_{it}) = & a_{0i} + (1 - b_{1i} - b_{2i})\Delta \ln(W_{it-1}) + b_{1i}\Delta \ln(P_{it}) + b_{2i}\Delta \ln(P_{it-1}) \\
& + c_{1i}\Delta \ln(Pc_{it} - P_{it}) + c_{2i}\Delta \ln(Pc_{it-1} - P_{it-1}) \\
& + d_{1i}\Delta \ln(Q_{it} / N_{it}) + d_{2i}\Delta \ln(Q_{it-1} / N_{it-1}) \\
& + e_{1i}\Delta U_{it} + e_{2i}\Delta U_{it-1} + s_i ERM_{it} \\
& - f_i(\ln(W_{it-1} / P_{it-1}) - \ln(Q_{it-1} / N_{it-1}) - g_i U_{it-1})
\end{aligned} \tag{43}$$

Consistent with the other studies, the authors found significant differences across countries in terms of the speed of adjustment, the short and long-run responses to productivity and the long-run unemployment effect. An equation based on the employment rate produced broadly consistent results. After a simulation exercise on this basis, the authors reached the conclusion that these differences should not be problematic in terms of asymmetric developments, yet that flexibility is too low to compensate for the loss of the national monetary instrument. No breaks were mentioned in the study.

A third and final approach in the literature has been to use panel data without equilibrium correction.

Guichard and Laffargue (2000) based their specification on a 2-year staggered contracts bargaining model over expected wages for sixteen OECD countries. Their wage equation (annual data) also included productivity (Y/E), a measure of the tax wedge (WDG) and unemployment (UN):

$$\begin{aligned}
\ln(W_t) = & w_0 + w_1[r \ln(Y_{t-1} / E_{t-1}) + (1 - r) \ln(Y_t / E_t)] \\
& + w_2[r \ln(WDG_{t-1}) + (1 - r) \ln(WDG_t)] + w_3[r \ln(UN_{t-1}) + (1 - r) \ln(UN_t)] \\
& + r q \ln(P_{t-1}) + (1 - r) \ln(P_t) + r(1 - q) \ln(P_{t+1}) + r(1 - q)(\eta_{t-1,t} + \eta_{t,t+1}) \\
& + r \varepsilon_{t-1,1} + (1 - 2r) \varepsilon_{t,1} + r \varepsilon_{t,2}
\end{aligned} \tag{44}$$

This was found to perform poorly, leading the authors to argue that unemployment was an inadequate indicator of labour market tensions and to replace it with the employment rate. The authors adopted a panel approach and found that the best model had country-specific employment and productivity coefficients. Their specification unfortunately omitted the change in the unemployment rate, which is important for general dynamic and hysteresis considerations, and the possibility of structural breaks was not considered.

TABLE 26: ECONOMETRIC STUDIES ON EMU AND LABOUR MARKET FLEXIBILITY

Study	Countries	Sample	Econometrics	Structural Change/Breaks
Scheremet (2000)	DE, OE, NL, BE, FR, UK, IT, PT, SW, ES	1979-1997	SURE estimation	NL(1985), IT(1993), PT(1988), UK(1988)
Guichard and Laffargue (2000)	DE, OE, BE, CA, ES, US, FI, FR, EL, IT, IR, JP, NL, PT, UK, SW	1982-1997	GMM	-
Cadiou and Guichard (1999b)	DE, FR, IT, UK, NL, SW, DK, FI	1991q1-1997q4	Constrained simultaneous equations	-
Lauer (1999)	DE, FR, ES, PT, NL, IT	1960-1996	Engle and Granger (1987) procedure	DE (1989), FR (1990), NL (1984), IT (1994), ES (1988), PT (1976-77).
Barrell and Dury (2001)	DE, FR, ES, IT, NL, BE, PT, IR, FI, OE, UK	1970q1-1998q4	SURE estimation ; dynamic fixed effects	DE (1991), ES (1987), IT (1993)
Lo Cascio (2001)	AT, BE, DK, FI, FR, DE, IT, NL, ES, SW, UK	1979q1-1998q4	VAR for GDP, CPI, short-term interest rate	-

2.4 Interim conclusion

Having surveyed the relevant literature, a few general remarks are in order.

First, we note that although nearly all of these studies claim the intellectual heritage of Layard et al. (1991) as background, few seem to embrace its full implications. Most notably, institutional variables which are so prominent in this framework are absent from these empirical models (such as measures of union bargaining power). This may be expected to bias coefficients. Controlling for institutional changes will prove especially crucial if we intend to claim with some robustness that a given structural change in flexibility coinciding with the introduction of EMU (if any) is not simply picking up institutional change.

Second, with a laudable exception, the implications of the multi-sectored structure of the economy are ignored (Cadiou and Guichard, 1999a). One important distinction is that between the public and private sectors. The public sector does not today serve as a reference for private wage bargaining in our sample of EU countries any more⁶⁰, but may nonetheless matter through other channels. For instance, the use of aggregate data may lead to downward-biased estimates of productivity effects due to the softer budget constraint facing the public sector. Similar biases can be expected in other coefficients, such as those corresponding to unemployment or those capturing adjustment dynamics. The significance of this distinction is likely to differ across countries, as public sector employment accounts for between 14% (Netherlands) to 35% (Sweden) of the total in

⁶⁰ This was not always so. For instance, in Italy the negotiations for wages in public enterprises were often pace-setting for the rest of the economy prior to 1990 (Ochel, 2000).

Europe. It cannot however be dismissed *ex ante*, especially for purposes of cross-country comparisons. It is a factor that should be kept in mind when interpreting coefficients, and will be mentioned further below.

A similar issue arises with respect to the “informal sector”. According to the European Commission (1998) – and depending on the method used – the sector would constitute between 7% and 19% of total employment in the EU. This will be a source of measurement error. Most notably, the unemployment rate will fail to reflect the true underlying labour market tensions, as people counted as unemployed will find employment in the informal sector, thus failing to exert downward pressure on wages. This would in turn bias the unemployment coefficient downwards (Cadiou and Guichard, 1999a).

Third, dynamic homogeneity has sometimes been imposed but is unnecessarily restrictive. It is indeed easily conceivable and empirically verifiable that workers are not always fully compensated for inflation in the short run through wage changes. Long-run homogeneity is a much more plausible restriction to impose, as we would not expect to see any money illusion then.

Fourth, and perhaps most importantly, previous studies have either relied on proxying for inflation expectations, on rational expectations or altogether avoided the issue. Expectations are perhaps the most crucial problem in studying structural changes arising from a regime shift such as EMU. Controlling for their changes *appropriately* is essential before forming a judgement on whether the wage setting mechanism has undergone

changes itself. The issue cannot be ignored. Indeed, any instability found would otherwise run the risk of being misinterpreted, for instance.

Among the studies surveyed above, those attempting to control for expectations all relied on an assumption that expectations were consistent in some form or the other. Tests for stability were therefore, in reality, joint tests of this hypothesis and the hypothesis of no structural change (Anderton and Barrell, 1995, footnotes 10, 11; Chan-Lee, 1980).

Rational expectations are clearly too much of a heroic assumption to endow agents with in the present context. This hypothesis holds that “*expectations, since they are informed predictions of future events are essentially the same as the predictions of the relevant economic theory*” (Muth, 1961). In this view, agents’ subjective expectations of economic variables coincide with the objective mathematical conditional expectations of those variables: agents know the true underlying model of the economy and use it to inform their expectations. In a *stationary* world, this requires that agents condition their expectations on the correct information set I_t and have knowledge of the complete joint distribution of all variables in the economy, which is a demanding enough proposition. However, in a *non-stationary* world, agents also need to know the shape of that distribution *in the future* (Clements and Hendry, 1999), as is made clear mathematically through subscripting:

$$E(y_{T+h} | I_t) = \int_{-\infty}^{+\infty} z * f_{T+h}(z | I_t) dz \quad (45)$$

Were such an assumption to hold, this chapter would effectively be irrelevant, as both the author and the readers would have known in advance whether a structural break would

occur due to the euro, and the commonality of forecast failure would be inexplicable. Arguments for rational expectations in terms of ‘as if’ behaviour or convergence to them through learning processes fail in a non-stationary world. Furthermore, they have been rejected by the data both directly (e.g. Roberts, 1997) and indirectly (Ericsson and Irons, 1995) – see Chapter 3 for further details on this point.

Accordingly, we choose to rely on the inflation expectations measure developed in Chapter 1 on the basis of EC Consumer Survey data, and shown to be “optimal” among the set of quantification methodologies considered.

3 ECONOMETRIC SPECIFICATION

3.1 Theoretical foundations of our wage equation

In this section, we will propose a wage equation with an equilibrium correction specification. Such an econometric model has an attractive interpretation in terms of a long-run equilibrium and short-run dynamics. Accordingly, we will divide our theoretical framework in the same way, following Moghadam and Wren-Lewis (1994). In long-run equilibrium expectations are by definition verified, and so we will ignore them in our theoretical discussion for that time scale. They are however of crucial importance in the short-run.

3.1.1 Long-run determinants of wages

We consider a model of bargaining between firms and unions. Although there has been a tendency for union membership to decline in several countries, coverage rates are still

high and indeed exceed 75% of the workforce in most EU countries (EIRO, 2002a). Given this framework, we have in mind a long-run equilibrium solution for the economy of the following form:

$$\frac{W}{p_c}(1+t_1)\left[1+\frac{1-\beta}{\beta}\frac{\varphi}{1-t_2}\left(1-t_2-\frac{b}{p_c}\right)\right]=\left(\frac{p}{p_c}\right)\frac{Y}{L} \quad (46)$$

W denotes wages; p_c and p respectively denote consumer and producer prices; t_1 , t_2 and t_3 respectively denote the employment, direct and indirect tax rates; (Y/L) denotes productivity; b denotes unemployment benefits; and β represents the relative bargaining power of unions versus firms $(1-\beta)$.

It is possible to arrive to an equation such as (46) through a number of different routes. A possible microfoundation would be a right-to-manage model in which firms and unions bargain according to the Nash bargaining paradigm (for details, see Moghadam and Wren-Lewis, 1994; Curto Millet, 2004).

Economic theory suggests that several variables are implicit in this specification. The probability φ of becoming unemployed will be related to the unemployment rate in the economy UN . The relative bargaining power of unions and firms $\beta/1-\beta$ may depend on union density $UDENS$, bargaining coverage $ACOV$ and employment protection legislation EPL . The ratio of producer to consumer prices p/p_c will be a function of import prices p_m and indirect taxes t_3 .

Augmenting (46) by such considerations yields the following log-linearised equation:

$$\begin{aligned} \ln\left(\frac{W(1+t_1)}{p_c}\right) = & a_0 + a_1 \ln UN + a_2 \ln\left(\frac{b}{p_c}\right) + a_3 \ln(1-t_2) + a_4 \ln(1+t_3) + a_5 \ln(Y/L) \\ & + a_6 \ln\left(\frac{p_m}{p_c}\right) + a_7 \ln UDENS + a_8 \ln ACOV + a_9 \ln EPL \end{aligned} \quad (47)$$

A more general model with several periods would take into account the possibility of prolonged unemployment, which matters as benefit systems differ in their generosity over time. This will be incorporated into our empirical specification, through a benefit duration index *BD*. Furthermore, since this model represents the bargaining of one representative firm and union, it fails to capture the aggregate impact that “bargaining coordination” (variable *COORD*) would have. This will also be part of the empirical specification, and its importance is discussed in Section 3.6.3.

3.1.2 Dynamics of the wage equation

We adopt a staggered contracts view of the labour market to inform our dynamics. Namely, different groups of workers bargain their wages at different points in time, in such a way that their contracts overlap. Such a framework has the capability of generating significant nominal inertia, even under an assumption of rational expectations (Taylor, 1979, 1980). A criticism of such models is their lack of microfoundations for the existence of long-term contracts. Microfoundations are however not our concern here, as for our purposes it only matters that such contracts *do* exist in practice. In any event, the justifications in terms of, for instance, contracting costs and coordination failures are many (Phelps, 1990).

We show the implications of this framework by considering a situation in which at every quarter, one quarter of (otherwise identical) workers negotiate their annual contracts

(Moghadam and Wren-Lewis, 1994). Assume that each group seeks to achieve the same desired level of the real wage w_t^* for each period t . So for group 1 negotiating at t , for instance, the following relationship holds:

$$\frac{1}{4} \sum_{i=0}^3 \ln w_{t+i}^* = \frac{1}{4} \sum_{i=0}^3 \ln \left(\frac{W_{t+i}^1}{P_{t+i}^{\text{exp}}} \right) \quad (48)$$

Given that wages are set annually ($W_t^1 = W_{t+1}^1 = W_{t+2}^1 = W_{t+3}^1$), this implies:

$$\ln W_t^1 = \frac{1}{4} \sum_{i=0}^3 \ln w_{t+i}^* + \frac{1}{4} \sum_{i=0}^3 \ln P_{t+i}^{\text{exp}} \quad (49)$$

The other groups will have similarly negotiated in periods $(t-1) \dots (t-3)$. Aggregate nominal wages are therefore given by:

$$\begin{aligned} \ln W_t &= \frac{1}{4} \sum_{j=1}^4 \ln W_t^j \\ &= \frac{1}{4} \left[\frac{1}{4} \ln P_{t+3}^{\text{exp}} + \frac{1}{2} \ln P_{t+2}^{\text{exp}} + \frac{3}{4} \ln P_{t+1}^{\text{exp}} + \ln P_t^{\text{exp}} + \frac{3}{4} \ln P_{t-1}^{\text{exp}} + \frac{1}{2} \ln P_{t-2}^{\text{exp}} + \frac{1}{4} \ln P_{t-3}^{\text{exp}} \right. \\ &\quad \left. + \frac{1}{4} \ln w_{t+3}^* + \frac{1}{2} \ln w_{t+2}^* + \frac{3}{4} \ln w_{t+1}^* + \ln w_t^* + \frac{3}{4} \ln w_{t-1}^* + \frac{1}{2} \ln w_{t-2}^* + \frac{1}{4} \ln w_{t-3}^* \right] \\ &= \frac{1}{16} \left[\ln P_{t+3}^{\text{exp}} + 2 \ln P_{t+2}^{\text{exp}} + 3 \ln P_{t+1}^{\text{exp}} + 4 \ln P_t^{\text{exp}} + 3 \ln P_{t-1}^{\text{exp}} + 2 \ln P_{t-2}^{\text{exp}} + \ln P_{t-3}^{\text{exp}} \right. \\ &\quad \left. + \ln w_{t+3}^* + 2 \ln w_{t+2}^* + 3 \ln w_{t+1}^* + 4 \ln w_t^* + 3 \ln w_{t-1}^* + 2 \ln w_{t-2}^* + \ln w_{t-3}^* \right] \end{aligned} \quad (50)$$

Thus, the wage is shown to depend on a combination of forward and backward elements in prices and desired real wages.

The framework above sets out clearly some of the dynamic implications of a staggered worldview whilst retaining the advantage of expositional simplicity. Richer models have been built by Taylor (1979, 1980), who considers the effects of the relative wage structure on the contract being negotiated (x_t). In particular, he makes wages depend on

the settlements reached by other bargaining groups, both in the past and in the future, weighted by the amount of time these settlements overlap with the contract being negotiated. Wages also depend upon the expected state of excess demand (y) over the contract's lifetime. Taylor's (1980) model can be written:

$$x_t = \sum_{s=1}^{N-1} b_s x_{t-s} + \sum_{s=1}^{N-1} b_s x_{t+s}^{\text{exp}} + \frac{\gamma}{N} \sum_{s=0}^{N-1} y_{t+s}^{\text{exp}} + \varepsilon_t \quad (51)$$

$$y_t = -\beta p_t + v_t \quad (52)$$

The Taylor framework exhibits the same symmetric structure of forward/backward prices as the Moghadam and Wren-Lewis version. It however displays an even greater degree of inertia as the symmetry is centred around the lagged price level, because of the 'overlap discounting' applied to future contracts (Anderton et al., 1992).

'Desired real wages' can be postulated to depend upon the tax wedge, unemployment, productivity and institutional variables. However, it is unlikely that institutional changes (which tend to occur gradually, and over long time periods) will have much effect in a short-run horizon. A strict implementation of the framework above would however require modelling the structure of desired real wages both in its backward *and* forward components. This is not a straightforward issue. On a practical level, data on all such expectations are simply not collected. Moreover, little guidance is available as to whether it is more sensible to model tax wedge expectations, for instance, as rational or static (Moghadam and Wren-Lewis, 1994). The literature has, accordingly, largely ignored the issue. We will make the reasonable conjecture that although there are circumstances in which forward-looking elements of the above variables may matter, they are not likely to

be important for European countries at this point in time. For instance, expected productivity increases are likely to have been an important driving force of wages in a fast-growing economy such as that of Japan during much of the second part of the last century, with workers coming to expect certain rises in living standards. Even without data on such expectations, these effects might be reasonably well-captured by a linear trend, for instance. Our general argument is therefore that we do not expect such forward-looking components to be important, but that if this were wrong, it would be apparent through evidence of misspecification and could be corrected.

3.2 The empirical wage equation

Our wage equation has an equilibrium correction form and is a synthesis of the theoretical strands presented above. Indeed, its short-run dynamics are informed by the staggered contracts approach⁶¹, whilst its long-run solution arises from the Nash bargaining model proposed earlier.

We refrain from imposing dynamic homogeneity upon the equation *ex ante*, as there is little reason to believe that in the short-run workers will fully adjust wages to price changes. A homogeneity proposition is however much more reasonable as a feature of the long-run equilibrium implied by the model, and will be implemented.

The resulting wage equation was already presented in Chapter 1. It is reproduced here for convenience in Table 27, and so are the relevant sign priors (Table 28).

⁶¹ The wage equations accordingly condition on inflation expectations terms derived from surveys, and as such subject to sampling error. The standard errors associated with these coefficients do not take account of this, and could be refined via a bootstrap procedure. However, as only one variable is affected, it is unlikely to have an important effect on the results.

TABLE 27: WAGE EQUATION STRUCTURE

$$\begin{aligned}
\Delta \ln COMPTH_t = & \alpha + a_1 INFLEXP_{t,t+4} + a_2 INFLEXP_{t-1,t+3} + a_3 INFLEXP_{t-2,t+2} \\
& + a_4 INFLEXP_{t-3,t+1} + a_5 INFLEXP_{t-4,t} \\
& + b_1 \Delta \ln CPI_t + b_2 \Delta \ln CPI_{t-1} + b_3 \Delta \ln CPI_{t-2} + b_4 \Delta \ln CPI_{t-3} \\
& + c_1 \Delta \ln COMPTH_{t-1} + c_2 \Delta \ln COMPTH_{t-2} + c_3 \Delta \ln COMPTH_{t-3} \\
& + c_4 \Delta \ln COMPTH_{t-4} + d_1 \Delta \ln U_t \\
& + e_1 \left[\ln \left(\frac{COMPTH}{P_c} \right)_{t-1} - f_1 \ln PRODIT_{t-1} \right. \\
& - h_1 \ln(1 - ITAX)_{t-1} - i_1 \ln(1 + CTAX)_{t-1} \\
& - j_1 \ln U_{t-1} - k_1 \ln BRR_{t-1} - l_1 \ln ACOV_{t-1} - m_1 \ln EPL_{t-1} \\
& \left. - n_1 \ln UDENS_{t-1} - o_1 \ln BD_{t-1} - p_1 \ln RELPMPC_{t-1} - q_1 \ln COORD \right]
\end{aligned} \tag{53}$$

COMP_{TH}: Hourly compensation of workers
INFLEXP: Inflation expectations
PRODIT: Trend productivity
ITAX: Income/direct taxation
CTAX: Consumption/indirect taxation
U: Unemployment
P_c: CPI, All items
BRR: Benefit replacement rate
BD: Benefit duration index
ACOV: Collective bargaining coverage index
UDENS: Net union density
EPL: Employment Protection Legislation index
RELPMPC: Ratio of import to consumer prices (P_m/P_c)
COORD: Bargaining coordination index

TABLE 28: COEFFICIENT SIGN PRIORS

CORE VARIABLES	
$a_i \geq 0$ for all i	Higher inflation expectations should be reflected in higher wages. The magnitude of the coefficients should also be in line with the dynamics of the dependent variable.
$b_i \geq 0$ for all i	More inflation in the recent past should be expected to lead to a compensating adjustment of nominal wages, if such inflation was not reflected in the previous wage contract.
$c_i \leq 0$ or ≥ 0 for all i (variable)	Eq.(53) is constructed in first differences rather than the fourth difference $\Delta_4 W_t$. Given that our measure of inflation expectations has a yearly horizon, if the data were best described by a fourth order difference, we would have: $\Delta_4 W_t = f(x_t) \Leftrightarrow \Delta W_t = -\Delta W_{t-1} - \Delta W_{t-2} - \Delta W_{t-3} + f(x_t)$ With x_t the vector of explanatory variables. In this case, $c_i \leq 0$. This argument does not go through if we actually choose to express the dependent variable as $\Delta_4 W_t$, or if the dynamics are genuinely quarterly. In such cases, positive coefficients should be expected, as higher previous settlements could cause present ones to be inflated to preserve relativities. We let the modelling process determine the appropriate order of differencing.
$d_i \leq 0, j_i \leq 0$	Higher unemployment should lead to wage moderation in bargaining.

$e_i \leq 0$	The EqCM coefficient. It should be negative, showing adjustment to long-run equilibrium of the model, and significant to demonstrate cointegration.
$f_i \geq 0$	Higher trend productivity should have a positive effect on real wages; a plausible restriction would be the value one.

WEDGE VARIABLES	
$h_i \leq 0$	Direct/income taxes make workers demand higher wages, as they target the post-tax wage ⁶² . The way this is coded, we expect a negative sign.
$i_i \leq 0$	Indirect/consumption tax rate increases widen the gap between producer and consumer wages ⁶³ . The way this is coded, we expect a negative sign.
$p_i \leq 0$	The net price received by the firm is affected by import prices of inputs.
INSTITUTIONAL VARIABLES	
$k_i \geq 0 \quad o_i \geq 0$	Higher benefit replacement ratios and duration make unemployment less unpalatable and lead to tougher union bargaining. See also Shapiro and Stiglitz (1984).
$l_i \geq 0$	Higher bargaining coverage of unions implies a smaller competitive part of the economy and higher wages through tougher bargaining.
$m_i \geq 0$	Employment protection legislation reinforces union bargaining power.
$n_i \geq 0$	Higher net union density reinforces union bargaining power.
$q_i \leq 0$	Higher coordination leads in the end to wage moderation due to internalisation of externalities (see Calmfors and Driffill, 1988) ⁶⁴ .

Before turning to the econometric approach adopted in this study, we should mention an identification issue relating to the coefficient on inflation expectations, which has been raised in the New Keynesian Phillips Curve literature (Ruud and Whelan, 2001). In that context, it was noted that expectations respond to developments (e.g. oil shocks) which also raise prices and wages in backward-looking models of markup pricing, complicating inference regarding the degree of forward-lookingness. This would suggest in our context that prudence be exercised with regard to interpretations relying on the *level* of the inflation expectations coefficient.

⁶² This shifts the Bargained Real Wage (BRW) curve up in the Carlin and Soskice (1990) framework.

⁶³ This shifts the Price Determined Real Wage Curve (PRW) down in the Carlin and Soskice (1990) framework.

⁶⁴ Experimentation in the context of Curto Millet (2004) showed that it was difficult to evidence a ‘hump shape’ relationship on the basis of this data. On balance, and given this constraint, the discussion in Section 3.6.3 and the institutional information on coordination in European countries would lead us to expect a negative relationship.

3.3 Econometric approach

We adopt a General-to-Specific (*Gets*) modelling strategy, as advocated by Hendry (1995). The intellectual basis for this is the theory of reduction, which describes the conditions under which the actual (and complicated) Data Generation Process (DGP) can be reduced to a representation based on a (transformed) subset of the original variables with no information loss. That is, the process is such that everything about the parameters of interest can be learnt through the latter representation – termed the *Local* Data Generation Process (LDGP) –, which we seek to recover.

The *Gets* approach then consists in formulating a General Unrestricted Model (GUM). This is the most general model that can be postulated initially, on the basis of the current sample of data, previous empirical research as well as theoretical, institutional and measurement information. The GUM should contain as a special case the parsimonious representation at which the modelling exercise aims. A requirement for the GUM is that it be *congruent*, matching the data in all measured respects (e.g. homoscedastic, innovation errors) and is checked by testing for misspecification. A process of simplification can then be engaged, checking for congruence at every step.

The empirical model of Eq.(53) is essentially the ‘skeleton’ of our GUM. It embodies the theoretical contributions of the bargaining and staggered contracts literature and extends the past empirical work described in Section 2 by controlling for expectations and institutional variables.

Momentous advances have recently been achieved in what regards the process of simplification from the GUM to the final model, and have been enclosed in the PcGETS package as an algorithm (Hendry and Krolzig, 2001). PcGETS delivers a principled way to carry out multi-path simplification that is surprisingly free of ‘pre-test bias’ problems, selecting a final model on the basis of encompassing considerations⁶⁵. However, we choose not to follow this approach for the present modelling. Indeed, PcGETS considers zero-type parameter restrictions, whereas we may be interested in other possibilities, such as moving average restrictions (Aron et al., 2004). These can often yield more economically interpretable outcomes. A relevant example would be the moving average restriction Moghadam and Wren-Lewis (1994) impose on the structure of their inflation expectations terms *ex ante* on the basis of theory, and we may be interested in verifying whether our data support this. Furthermore, the version of PcGETS available at the time of writing did not yet allow for the possibility of sign restrictions.

We instead apply a ‘Bayesian approach’ informally, in the manner of Aron et al. (2004). Emphasis is placed on the long-run solution providing sensible results, leading us to impose the restrictions arising from economic theory detailed above if required. The lag structure of the model is essentially data-determined. The temporal position of the EqCM is informed by the short-term inflation dynamics, as negative short-run coefficients of the same magnitude as the speed of adjustment coefficient may suggest the need for a specific lag. For instance, referring to Eq.(53), if $e_1 = b_2 = b_3$, this would suggest an EqCM lagged three times⁶⁶. We are also especially attentive to the possibility of

⁶⁵ If encompassing reductions are exhausted and multiple terminal models remain over the iteration, information criteria are used instead. For details, see Hendry and Krolzig (2001).

⁶⁶ This restriction is taken to be indicative of a more general condition holding for the other EqCM terms.

restrictions that would suggest fourth difference effects in the dynamics. For instance, $b_1 = b_2 = b_3 = b_4$ would suggest a $\Delta_4 I CPI_t$ formulation. Such effects may be interpretable as the consideration of inflationary changes since the last negotiations in a framework of yearly bargaining, for example.

The nature of certain variables in the equilibrium correction term makes them clearly integrated, and for standard inference to hold we would require them to cointegrate. This will be signalled by a significant coefficient for the speed of adjustment term e_1 , whose distribution is non-standard but relatively close to a Student t-distribution (Hendry and Juselius, 2000).

Arguably, the model presented in the previous section could be further generalised, notably through additional delta effects. In the actual modelling, further such lags have been considered (e.g. for the unemployment variable). Other delta effects (e.g. for institutional variables) were not considered, as theoretically they appear to be more naturally features of the equilibrium rather than short-run influences on the wage bargain. These considerations were reinforced practically by the limited sample sizes at our disposal, and the empirical fact that many such variables hardly vary in the short-run, making such effects almost certainly irrelevant in a quarterly model such as ours.

3.4 The time aggregation problem

The time aggregation problem arises when the data used to describe a process have a lower frequency than that of the original process itself. Consider the following process referring to a monthly frequency:

$$\Delta \ln W_t = \alpha z_t + \varepsilon_t \quad (54)$$

Suppose that only data at the quarterly frequency are available. Given that:

$$\begin{aligned} \Delta_3 \ln W_t &= \alpha(z_t + z_{t-1} + z_{t-2}) + \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} \\ &= \alpha \bar{z}_t + \bar{\varepsilon}_t \end{aligned} \quad (55)$$

This has implications for an equation in *quarterly* differences:

$$\begin{aligned} (\ln W_t + \ln W_{t-1} + \ln W_{t-2}) - (\ln W_{t-3} + \ln W_{t-4} + \ln W_{t-5}) \\ = \alpha(\bar{z}_t + \bar{z}_{t-1} + \bar{z}_{t-2}) + [\bar{\varepsilon}_t + \bar{\varepsilon}_{t-1} + \bar{\varepsilon}_{t-2}] \end{aligned} \quad (56)$$

Thus, the error term would display serial correlation with a moving average structure.

The data we have collected are quarterly in nature. Given the bargaining framework we presented earlier, we expect that this should match the frequency of the process we are trying to model, and time aggregation issues should not be problematic. However, it will be important to test this assumption by verifying the white noise properties of the residuals. This will be systematically carried out in Section 4.1.

3.5 Correspondence between measures and concepts

An important question regarding our empirical implementation concerns the extent to which the coefficients earmarked for a possible structural break correspond to the concepts they purport to measure. This does not seem problematic with the ‘speed of adjustment’ coefficient, for instance, which has a clear meaning in an equilibrium correction framework. Real wage rigidity, on the other hand, can be measured in a number of different ways.

An alternative to the type of measure we employ has been proposed by Blanchard and Quah (1989) and notably used in our context by Viñals and Jimeno (1996) and Berthold et al. (1999). The approach consists in combining a labour market model of the Layard et al. (1991) type with a Structural VAR approach (SVAR). The advantage claimed is that it is not only the current effect of variables that is then taken into account, but the whole adjustment process, as the impulse response function serves to compute an index of real wage rigidities.

Berthold et al. consider a VAR modelling the change in real wages and unemployment⁶⁷. A distinction is made between technological shocks and wage-push shocks, the identifying assumption being that wage-push shocks have only temporary effects on real wages. Three further assumed restrictions enable the extraction of the pure shocks from the VAR residuals, allowing the authors to compute the measure of real wage rigidity implied by the model.

Several issues are however particularly problematic in this approach. First, the identification scheme is (as often with such models) open to criticism due to its ad hoc nature - is the distinctive feature of 'technology shocks' their permanence? Furthermore, the data contain many more 'shocks' than the two singled out here, some of them completely trivial and unrelated to either technology or wage-push shocks, such as measurement error. Second, VARs being reduced forms of the system of economic relationships, their stability is clearly conditional on the sample used (time period,

⁶⁷ They also consider a VAR comprising the change in the real wage and the change in unemployment due to order of integration issues.

countries etc.), and would be most inappropriate in a study focusing on stability such as this one.

In light of these shortcomings, we will maintain the approach outlined earlier. Not only does it appear to be more robust, but the EqCM structure enables insight into the dynamics of real wage rigidity all the same.

Another word of caution relates to the interpretation of the unemployment coefficient, as its relationship with actual flexibility is not fully monotonic. Indeed, in an extreme case, perfect wage flexibility would absorb shocks instantaneously such that unemployment remains constant and unemployment appears insignificant in the wage equation despite there being a clear relationship between the two (Goubert and Omey, 1996). The degree to which this argument applies is unknown, but likely to be of limited significance.

3.6 Data description and motivation

In order to draw sensible conclusions regarding the stability of wage equations, it is essential to be using the right data, and for these data to be of high quality. This section defines the variables used and explains the data work behind our study. The variables fall into three broad categories: (a) General macroeconomic variables, (b) Wedge variables and (c) Institutional variables.

3.6.1 General macroeconomic variables

Hourly Compensation of Workers (COMPTH)

The dependent variable in our specification is the hourly compensation of workers in the economy. This is computed on the basis of the (quarterly) aggregate compensation of employees from the OECD's Economic Outlook 77. The OECD defines employee compensation as:

“All payments by resident producers of wages and salaries to their employees, in kind and in cash, and of contributions, paid or imputed, in respect of their employees to social security schemes and to private pension, family allowance, casualty insurance, life insurance and similar schemes.”

The absence of Italian data from the electronic database attached to the OECD's Economic Outlook 77 led us to rely on those present in the Economic Outlook 76. This is without consequence for the sample length available for Italy in our regressions. For Spain and Belgium, quarterly data on this dimension are unavailable from the Economic Outlook and were sourced from the OECD's Quarterly National Accounts. The sample in the case of Belgium was limited to 2003q2. Subsequent values were extrapolated according to the growth rates of the (non-truncated) annual data. As could be expected, German data are only available from 1991 onwards.

To obtain the aggregate compensation of workers from these data (therefore including both employees and the self-employed) we follow Batini et al. (2000) by assuming that the productivity of self-employed workers is the same as that of employees on average. We can then obtain a measure of the compensation of workers by multiplying the OECD figures by a factor $A_t = (E_t + SE_t)/E_t$, where E_t is the number of employees and SE_t the number of self-employed.

Total hours worked in the economy can be extracted from the OECD Productivity Database⁶⁸. This database contains the average number of hours worked in the year per worker and a consistent measure of employment in the economy, allowing the derivation of total hours worked. The fact that the data are available on an annual basis should not be excessively problematic given the gradual behaviour typical of changes in this series⁶⁹. The hourly compensation of workers can then be derived by dividing the aggregate compensation measure by the total number of hours worked in the economy. Figure 14 plots these data in the case of the Eurozone countries in our sample, and Figure 15 does the same for Sweden and the United Kingdom.

⁶⁸ This data is compiled and made consistent by the OECD. A number of other sources are used: the OECD Employment Outlook, OECD Labour Force Statistics, OECD Annual National Accounts and a number of national sources. For details, refer to the Statistical Annex of the OECD Employment Outlook (2004a) and https://www.oecd.org/LongAbstract/0,2546,en_2825_30453906_29861101_1_1_1_1,00.html

⁶⁹ However, if business cycle effects caused hours worked to vary within the year, it is possible that the measurement error would correlate with cyclical terms in the regression – the unemployment variable. The results obtained do not suggest any obvious cause for concern, as the coefficients appear to be reasonable.

FIGURE 14: WORKER HOURLY COMPENSATION IN EUROS, EUROZONE

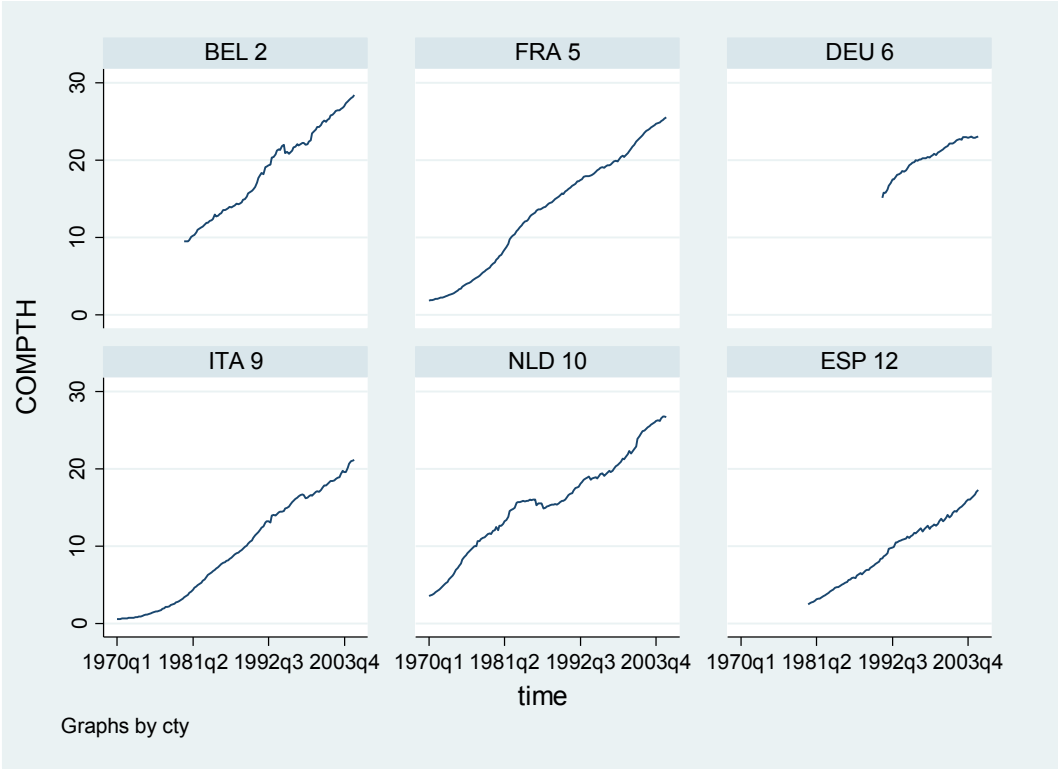
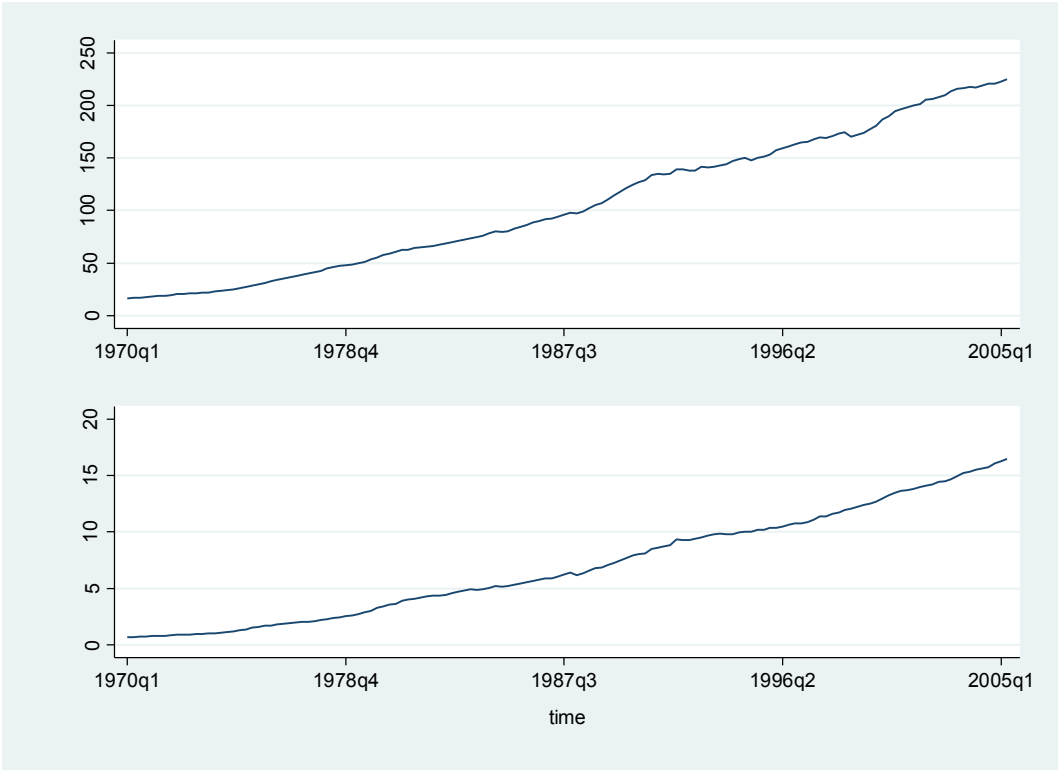


FIGURE 15: WORKER HOURLY COMPENSATION, SWEDEN (SEK) AND UK (£)



Productivity Trend (PRODIT)

These data are based on a Productivity Index referring to the whole economy, sourced from the OECD's Economic Outlook No 77. It is computed as the ratio of an output measure and the number of employed individuals. Thus, it is potentially sensitive to changes in working time. The series consists of quarterly data for all countries, except for Spain and Belgium. For those two countries, only annual data are available in the Economic Outlook 77. Semi-annual data are available for both from the OECD's Economic Outlook 73 up to 2003s2; subsequent values are extrapolated on the basis of the annual data in the Economic Outlook 77.

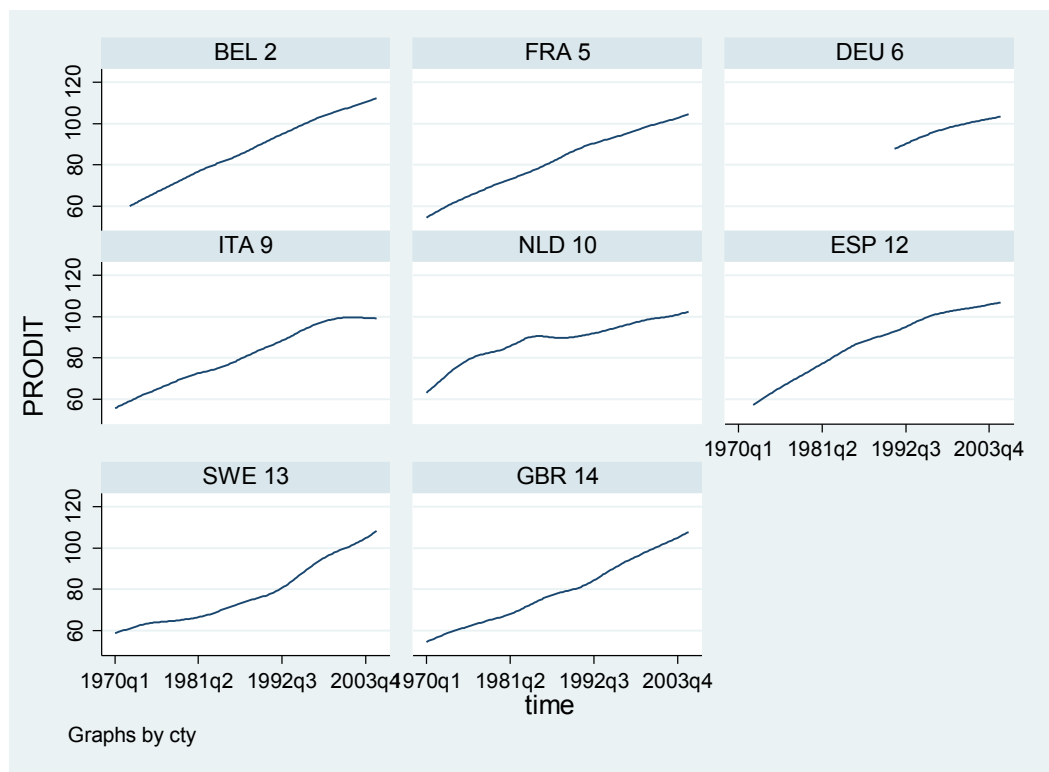
It is reasonable to expect that short-term variations in productivity will not overly impact wage bargains and that consequently the productivity *trend* is the more relevant measure in this case. Thus, we derive our productivity measure by applying a Hodrick-Prescott filter with lambda parameter 1600 ⁷⁰ to the productivity index provided by the OECD.

Consumer Price Index, All Items (Pc)

The consumer price index employed concerns all items and is provided by the OECD Main Economic Indicators database. Graphs of inflation are provided in Appendix 1.

⁷⁰ This is a standard choice for a series with quarterly frequency (e.g. Hodrick and Prescott, 1980, 1997).

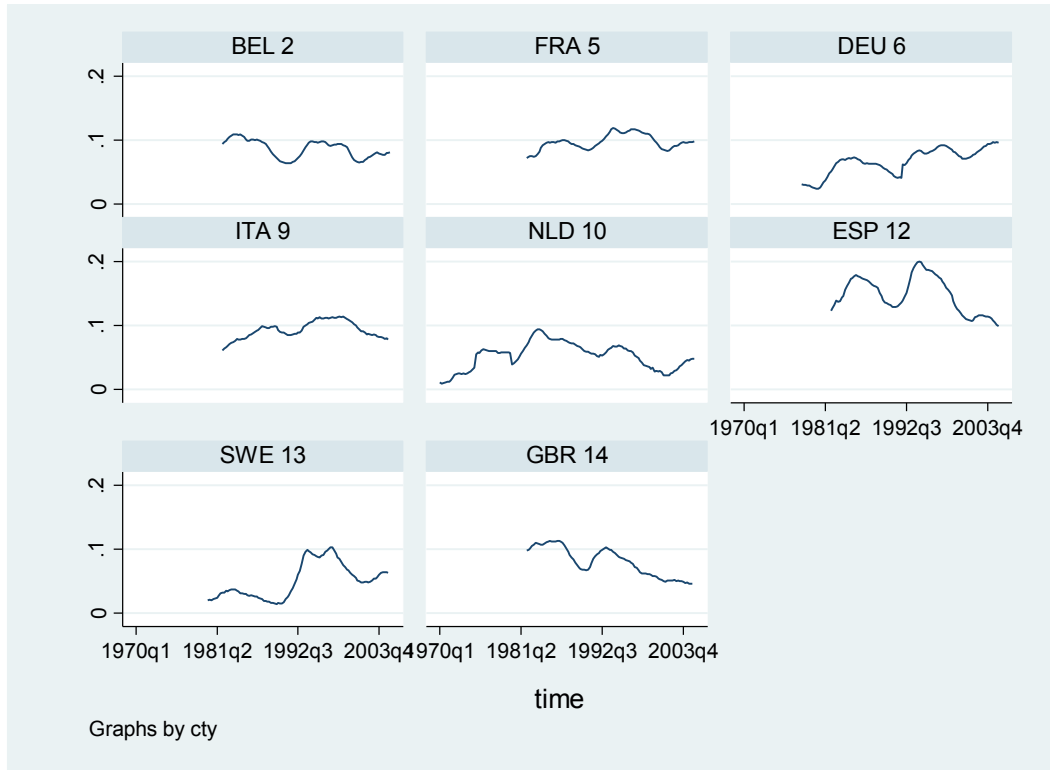
FIGURE 16: PRODUCTIVITY TREND INDEX, BY COUNTRY



Unemployment (U)

The OECD's Quarterly Labour Force Survey provides quarterly data on standardised unemployment rates (SUR). The OECD has adjusted these data to ensure comparability over time and to conform to the guidelines of the International Labour Office. For these reasons, it is used wherever possible in this study. In the case of Italy, however, such data are only available from 1982q1, which does not offer sufficient backward coverage for the sample period we wish to use. In this case, data compiled according to the national definition are used and sourced from the OECD's Economic Outlook 77.

FIGURE 17: STANDARDISED UNEMPLOYMENT RATES, BY COUNTRY



3.6.2 Wedge variables

Regarding the tax variables, we adopt the definitions used in Bell and Dryden (1996), Nickell (2003) and Nickell and Nunziata (2001) to quantify the components of the tax wedge that will be relevant for our purposes. Our data in this respect are an extension of the CEP-OECD dataset. The underlying data are taken from the OECD Revenue Statistics and OECD Economic Outlook 77.

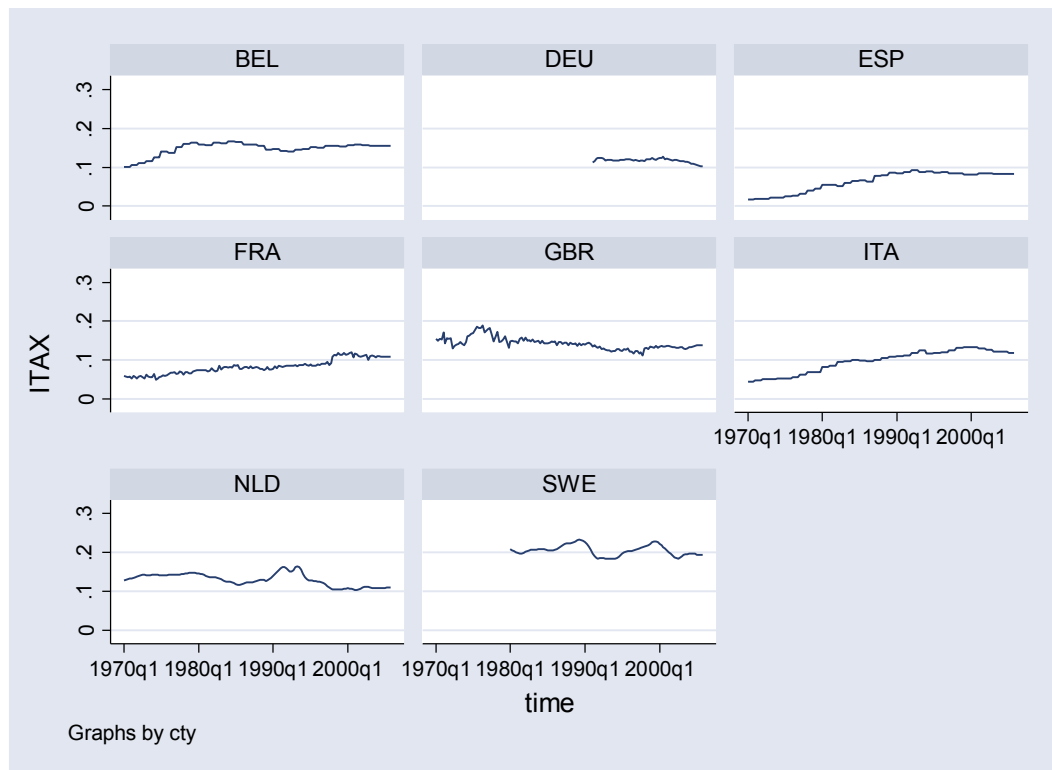
Income Tax Rate / Direct Taxation (ITAX)

The direct tax rate is obtained as:

$$ITAX = \frac{WC + IT}{HCR} \quad (57)$$

Its components are employees' social security contributions (WC), income taxes (IT) and household current receipts (HCR). The WC variable is only available on an annual basis (from the OECD's Revenue Statistics); the IT and HCR variables are available at a quarterly frequency from the OECD Economic Outlook 77, except in the cases of Belgium, Italy and Spain.

FIGURE 18: DIRECT/INCOME TAX RATE (%) 1970Q1-2005Q4



Consumption Tax Rate / Indirect Taxation (CTAX)

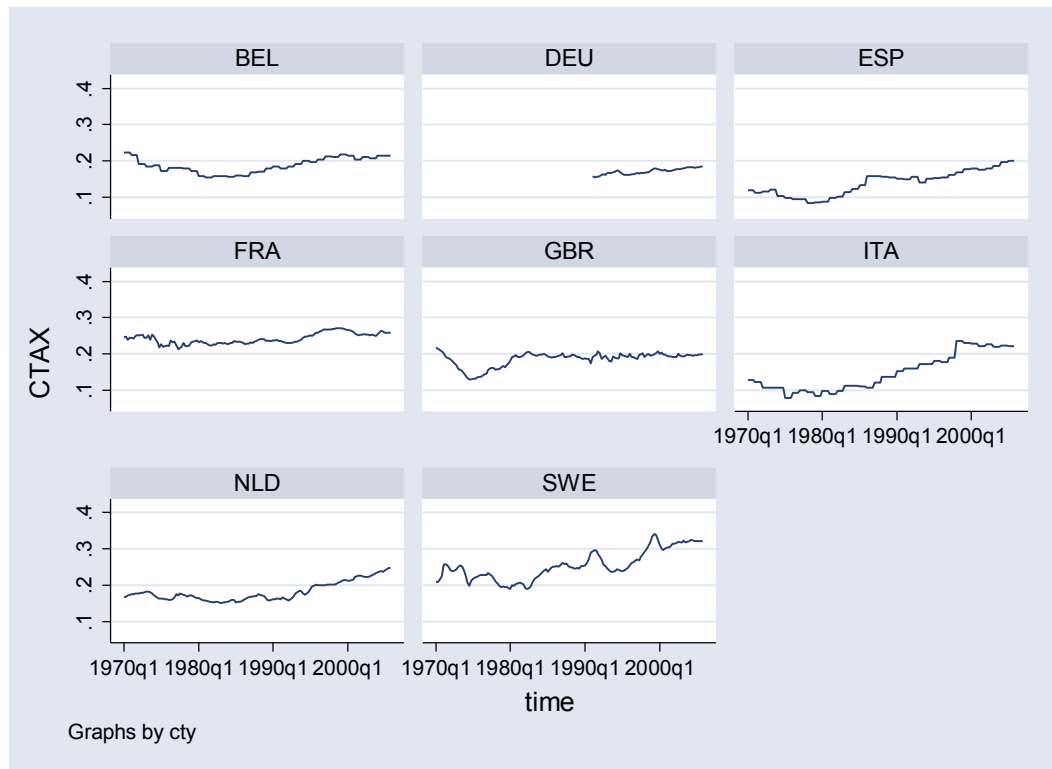
The indirect tax rate is obtained as:

$$CTAX = \frac{TX - SB}{CC} \quad (58)$$

That is, indirect taxes (TX) minus subsidies (SB) divided by private final consumption expenditure (CC). These variables are all available at a quarterly frequency from the

OECD's Economic Outlook 77, except in the cases of Spain, Belgium and Italy where only annual data are available.

FIGURE 19: INDIRECT/CONSUMPTION TAX RATE (%) 1970Q1-2005Q4

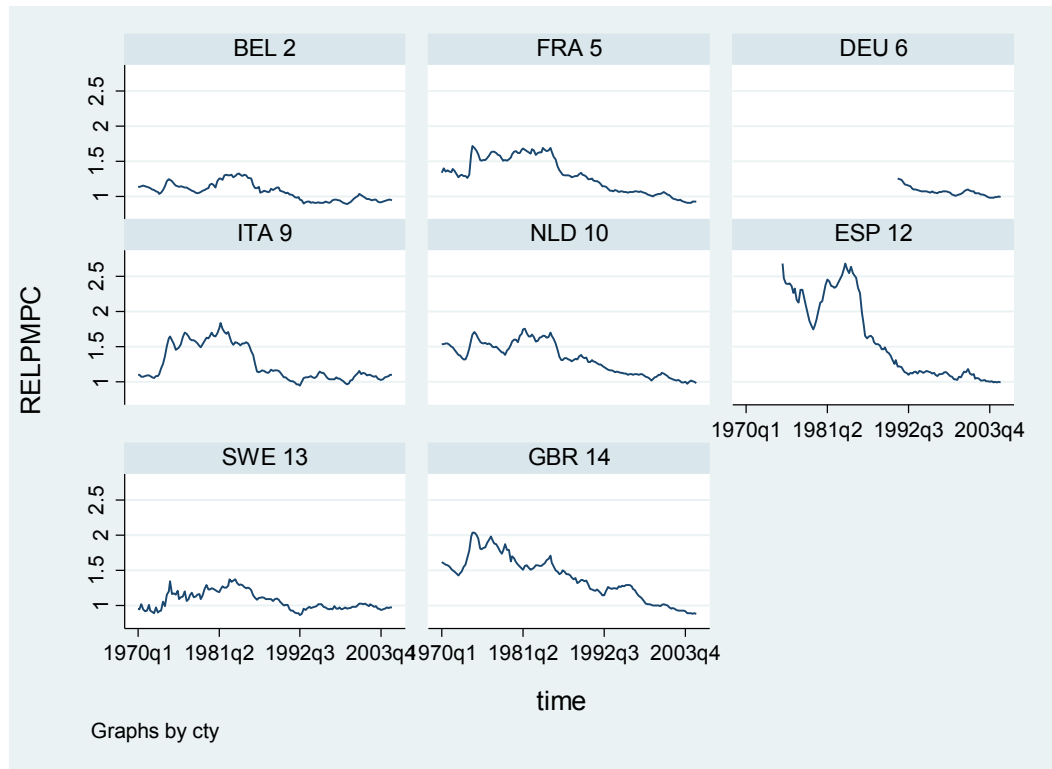


Ratio of Import to Consumer Prices (P_m/P_c) (RELPMPC)

CPI is as defined above and includes all items. The index of import prices refers to all goods and services, and is extracted from the OECD's Economic Outlook 77. These data are available at a quarterly frequency for all countries except Spain, which only has annual data available. It was found that the most satisfactory resolution of this issue was to use quarterly data from the OECD's ITCI Database (International Trade and Competitiveness Indicators) for Spain. The concept in this case concerns the import price of total goods; although a different concept, the time series movement of this and the Economic Outlook data were found to be almost identical. The ITCI data are available

until 2003q1, after which updating is realised on the basis of the Economic Outlook data by assuming a constant intra-annual growth rate.

FIGURE 20: RATIO OF IMPORT TO CONSUMER PRICES



3.6.3 Institutional variables

Institutional variables are important determinants of wages, and it is surprising to find many wage equations missing them.

The potential problems in terms of omitted variable bias are obvious when one considers the correlation between the unemployment and benefit replacement rates, for instance. One reason for the correlation comes from search models of unemployment, as the cost of this situation affects the search effort of the unemployed. A less obvious reason is that in models of optimal contract length, lower replacement rates would make the cost of

employment reductions higher from a worker perspective, thereby providing incentives for shorter contract lengths (Calmfors, 1998). It is not difficult to think of many such damaging correlations for each institutional variable below, highlighting the importance of their inclusion.

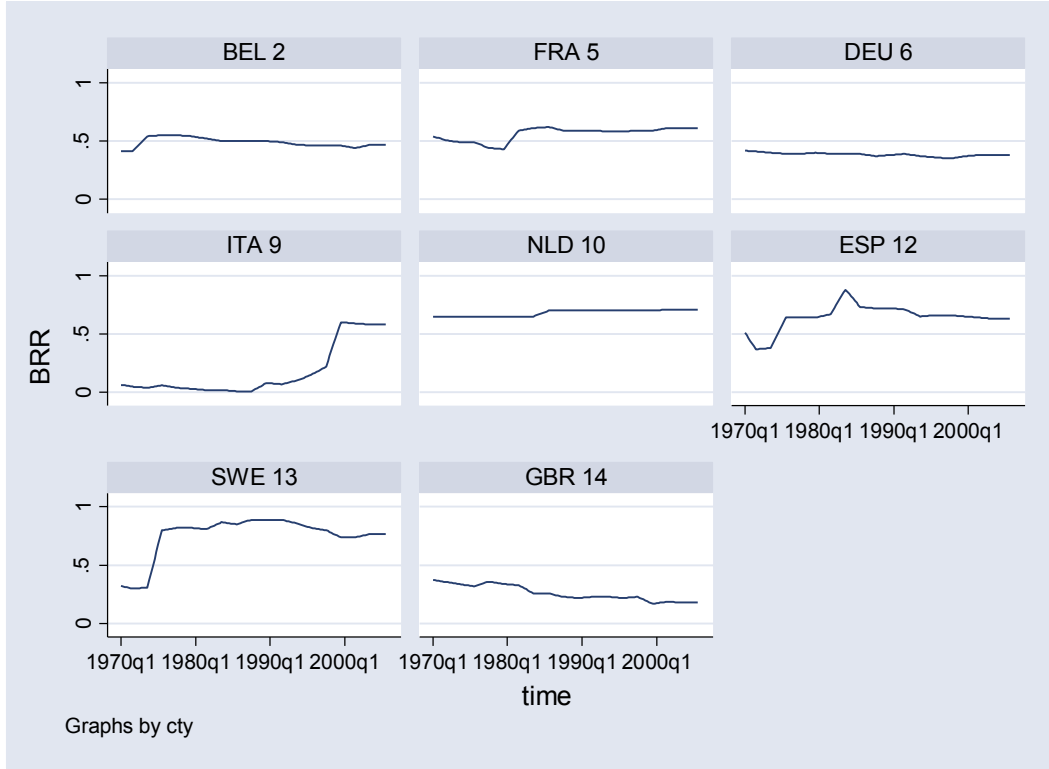
Benefit Replacement Rate (BRR)

The benefit replacement rate (BRR) is computed from data on the first year of unemployment benefits, averaged over family types of recipient (single earner/dependent spouse/spouse at work) (Nickell and Nunziata, 2001). The benefits are expressed as a proportion of average earnings before tax, and the data are collected every two years from 1961 to 2003 ⁷¹.

The data are assigned to the third quarter of every relevant year, since the replacement rates are computed to reflect the situation on July 1 of every year considered, as is clarified in the Appendices of the OECD's *Benefits and Wages* (2002, 2004). The series used is obtained by interpolation of these values. Post-2003 values are assumed to be constant, which is not unreasonable given the variation characteristic of this series.

⁷¹ It may be feared that the BRR would suffer from endogeneity, should wage increases be associated with reductions in the replacement rate. This would be the case if benefits did not rise in line with average earnings, which could be especially relevant in the short-run. However, the data underlying Figure 21 and the significant EqCM results obtained later for this variable do not suggest excessive cause for concern.

FIGURE 21: BENEFIT REPLACEMENT RATES 1970Q1-2005Q4



Benefit Duration Index (BD)

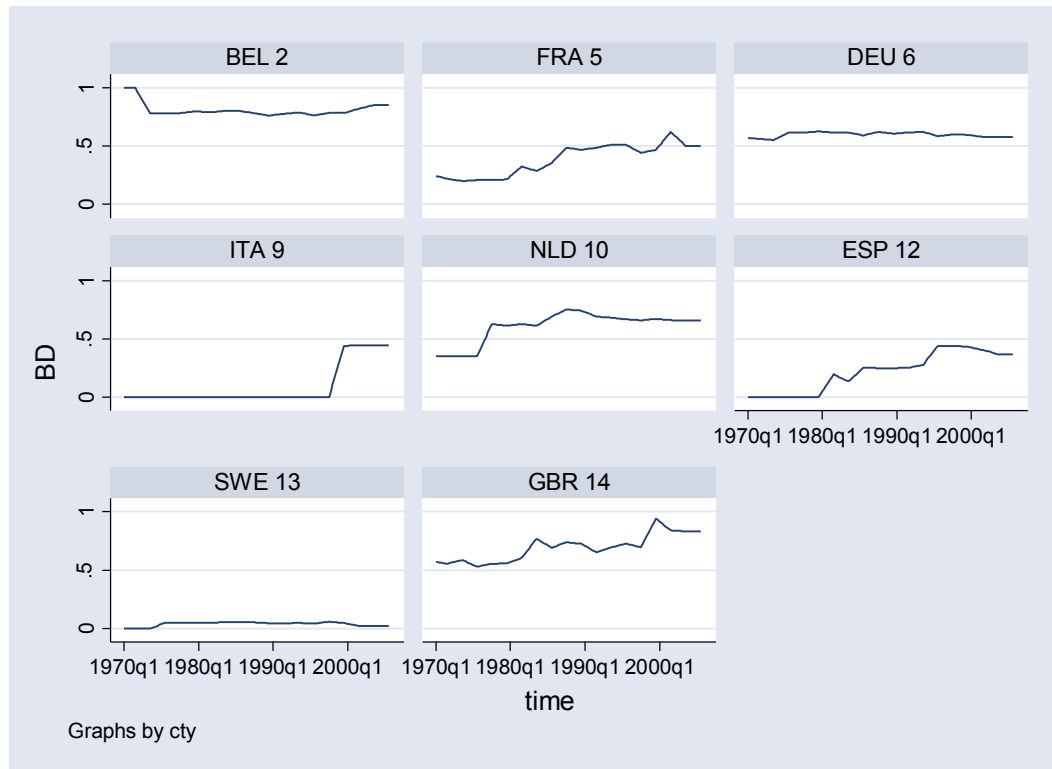
The index for benefit duration was constructed as a weighted average of the following benefit replacement rates (Nickell and Nunziata, 2001):

$$BD = \alpha \frac{BRR_2}{BRR_1} + (1 - \alpha) \frac{BRR_4}{BRR_1} \quad (59)$$

BRR1 is the benefit replacement rate received during the first year of unemployment, BRR2 is the replacement rate received during the second and third years of unemployment and BRR4 is the replacement rate received during the fourth and fifth years of unemployment. Note that more weight is given to the first ratio than to the second ($\alpha = 0.6$). If the benefits stop after one year, the index takes the value zero, whereas if the benefits are constant over the years of unemployment it takes the value one.

The index is rebased from the range $\{0,1\}$ to the range $\{1,2\}$ as the logarithm thereof will be used in the subsequent regressions.

FIGURE 22: BENEFIT DURATION INDEX 1970Q1-2005Q4



Net Union Density (Gainfully Employed) (UDENS)

The data on union density were extracted from the OECD's Corporate Data Environment (CDE) database.

There is widespread agreement that such figures should be based on *surveys*, wherever possible (OECD, 2005). The alternative is to use *administrative data*, which typically involves the compilation of membership statistics from questionnaires completed by individual unions or trade union federations. The ILO notes two disadvantages of the latter method in its World Labour Report 1997-98, namely: (1) difficulties in the

statistical coverage or the identification of unions, leading to under-reporting of non-affiliated and non-registered trade unions; small, newly-created unions are especially difficult to take into account; and (2) many figures are based on self-reporting, i.e. membership claims made by the unions. Self-reporting of membership may be influenced by certain administrative and political practices; notably, incentives for fully truthful revelation may not be in place. Despite efforts to check the data with a variety of sources for consistency, the approach may still yield incomparable and unreliable results.

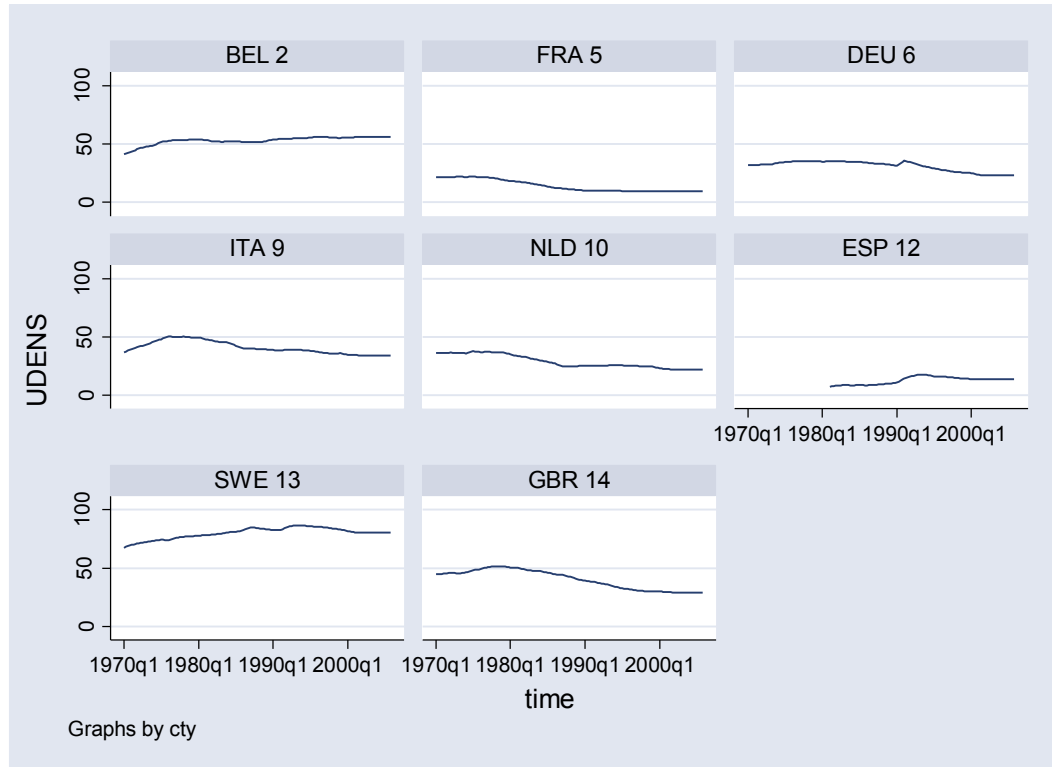
Unfortunately, the availability of survey data is very limited, forcing us to rely mainly on administrative data.

The concept we are interested in is *net* union density. Accordingly, our union density variable (UDENS) measures active union membership (excluding pensioners and students) as a share of the *gainfully employed* wage and salary earners (excluding the unemployed). The adjustments for non-active and self-employed members in the OECD data were carried out by Professor Jelle Visser, along the model used in the *Trade Unions in Western Europe* project (Ebbinghaus and Visser, 2000; see Ebbinghaus (2002) for some updated numbers).

Data are available on an annual basis up to 2002 for the UK, Sweden, The Netherlands, Italy and Germany; and up to 2001 for Spain, France and Belgium. Data points beyond that are assumed to show constancy in the series, which is not unreasonable given the historical variability displayed. All the data used are *administrative* in nature except for

the UK and Sweden, where administrative data are used at the start and updated following survey data when this becomes available (UK: post 1995; Sweden: post 1987).

FIGURE 23: NET UNION DENSITY (%) (1970Q1-2005Q4)



Adjusted Bargaining Coverage (ACOV)

The bargaining coverage variable (ACOV) is adjusted so as to refer to the number of employees covered by a collective agreement as a percentage of employees *equipped with the right to bargain* (hence the ‘adjustment’). In some countries, this criterion notably excludes the public sector. This is the relevant measure for our purposes, as it indicates the ‘governance capacity’ of a collective agreement within its own domain (EIRO, 2002a).

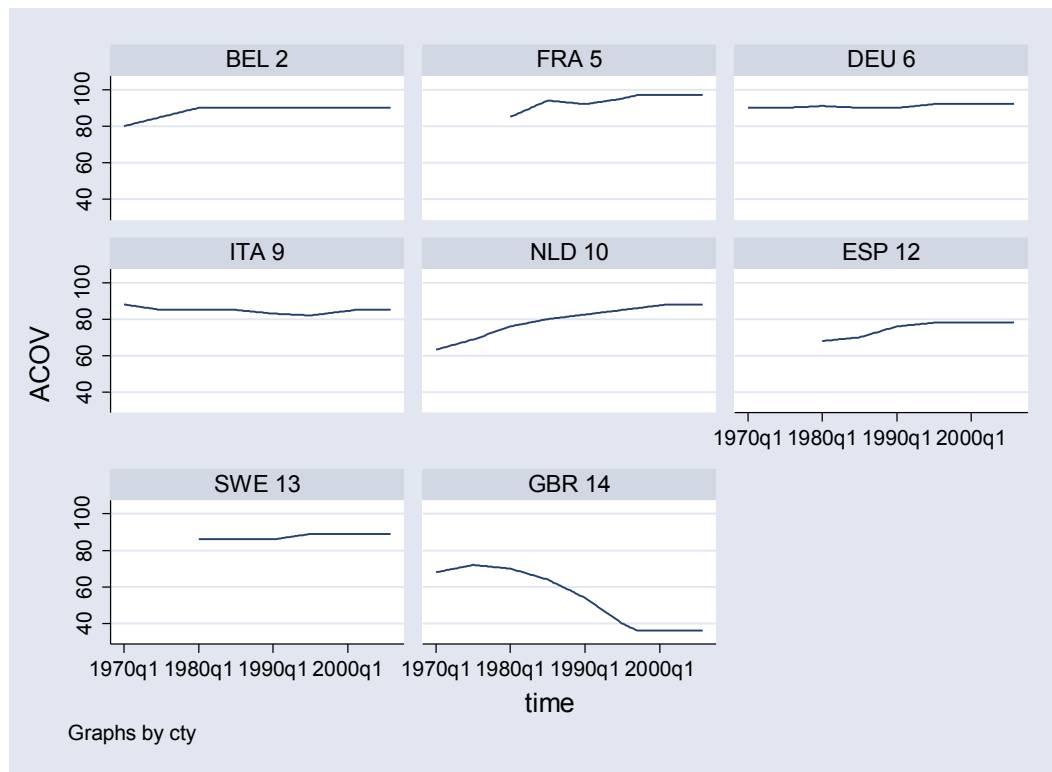
Our main source of data is Ochel (2001), as it offers the most comprehensive time series coverage available. The data generally start in 1960 and observations are provided every five years until the late 1990s. We used a number of sources to complete and update this data series as required. These complementary sources included: Traxler (1994), Traxler, Blaschke and Kittel (2001), OECD Employment Outlook 1997, EIRO (2002a) and Golden, Lange and Wallerstein (2001). Table 29 lists the samples available for each country and details the end observations resulting from our updating process. The series are assumed to remain constant following the last data point available, which is not unreasonable given the typically slow variation of this variable.

TABLE 29: BARGAINING COVERAGE DATA – SAMPLES BY COUNTRY

Country	Start of sample	End of sample
Belgium	1960	1996, 2000, 2001
France	1980	1997, 2000, 2001
Spain	1980	2001
Italy	1960	2001
Germany	1960	1995
United Kingdom	1960	1997, 2000, 2001
Sweden	1980	2000, 2001
The Netherlands ⁷²	1960	2001

⁷² An observation for 1990 is missing in this case, and was interpolated over. Only qualitative data was available in the Netherlands for the period 1965-1975. However, figures were available for coverage in the private and public sectors separately for those years. We weighted these by the proportion of people employed in each sector obtained from the OECD Economic Outlook 77.

FIGURE 24: ADJUSTED BARGAINING COVERAGE (%) (1970Q1-2005Q4)



Bargaining Coordination (COORD)

It has long been recognised that the type of bargaining framework in place has an effect on real wage realisations.

In a seminal article Calmfors and Driffill (1988; henceforth CD) argued that both highly centralised and highly decentralised systems would yield real wage moderation (the ‘hump-shape’ hypothesis). This is because firm-level bargainers face a high elasticity of product demand, restraining wage increases, whilst centralised bargaining internalises the externality imposed by sectoral/firm wage increases on the rest of the economy. At the industry level however, the elasticity of product demand is lower than that faced by firms and the internalisation effect weak, thus failing to effectively provide incentives for wage

restraint. This argument makes clear that such effects should be controlled for in our wage equation.

On the basis of their insight, CD compiled an index on the level of centralisation to be found in several OECD countries and set out to examine empirically its macroeconomic impact. However, Soskice (1990) provided a critique of this empirical approach. He argued that the theory put forward was one regarding the *behaviour* of bargaining partners rather than one about the *level* at which negotiations took place. For instance, although German wage bargaining takes place at the sectoral level, the institutional reality is one of ‘pattern-setting’, with the settlements negotiated by IG Metall serving as a reference for the rest of the economy. IG Metall realises the key role it plays in the German bargaining round, making the level of *coordination* in the economy high. CD had effectively made the confusion between what is *de jure* and what is *de facto* ⁷³.

Two alternative views of the wage bargaining framework can be identified and contrasted with the CD story (Traxler and Kittel, 2000; henceforth TK). First, the ‘corporatist’ tradition highlights the collective action problems burdening economy-wide coordination, which is unlikely to be delivered by fragmented bargaining systems. Hence, it tends to equate effective coordination with centralisation levels (Schmitter, 1979). Second, a ‘neoliberal’ approach emphasises the role of market competition in ensuring wage restraint and argues that the power of organised labour interferes with this, calling for deregulation (Siebert, 1997). The latter point is to be interpreted as concerning a move from collective agreements to individual labour contracts, and is to be properly

⁷³ This problem is also common in the central bank independence literature; see Forder (1996).

measured by collective bargaining coverage, for which we have a separate measure (ACOV).

This discussion has highlighted the potential role of several bargaining framework dimensions in wage determination. TK highlight an additional one, which we refer to as ‘capacity for implementation’. This refers to the ‘supportiveness’ of the institutional framework towards coordination, and is equated by TK with the legal enforceability of collective agreements in combination with a ‘peace obligation’ (which prohibits industrial action as long as collective agreements are in force). They argue that this dimension is an important prerequisite for coordination patterns with central-level peak associations voluntarily taking charge of coordination.

In light of the above, we extended an index compiled by Ochel (2000)⁷⁴ and contained in the Labour Market Institutions Database (Nickell and Nunziata, 2001). The index comprises the following dimensions, measured on the scales provided in parentheses:

- Centralisation (1-3)
- Coordination (1-3)
- Capacity for implementation (1-2)

Indices are computed for each of the above categories, the (subjectively weighted) average of which yields the summary index we use. ‘Centralisation’ ranges from firm-level (1) to central level (3). The ‘coordination’ index follows TK by distinguishing six types of coordination and marking them according to their effectiveness (from

⁷⁴ I am grateful for the help Wolfgang Ochel provided by sharing his data and commenting on my first round of updating.

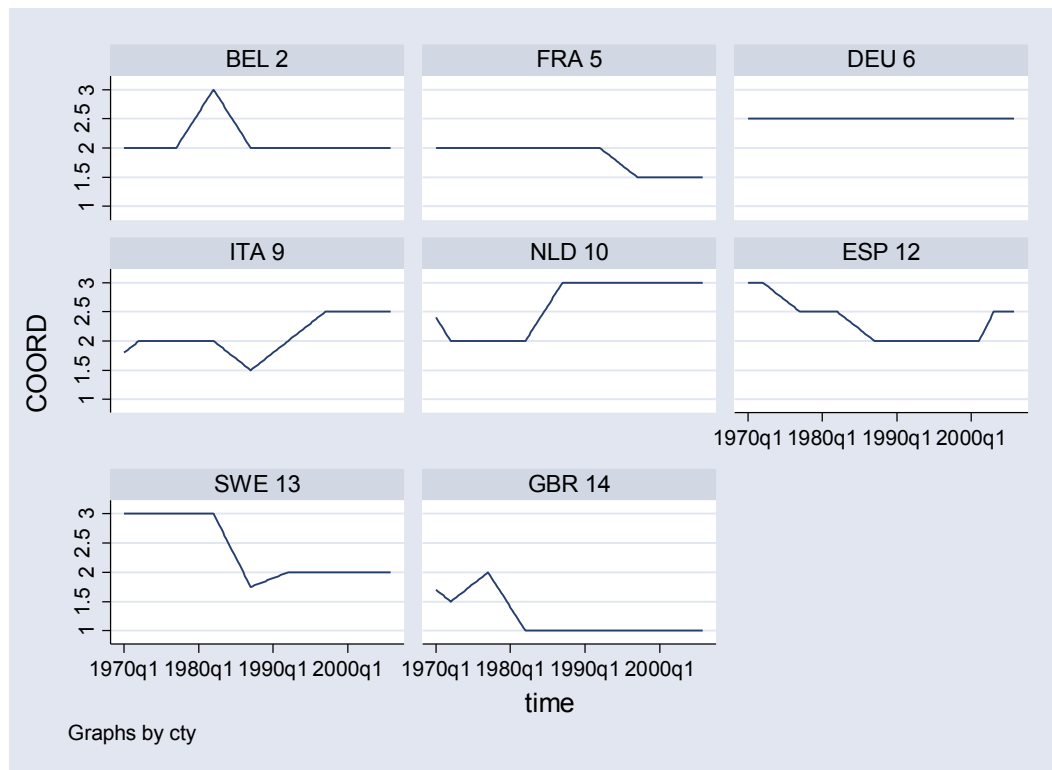
uncoordinated (1) to strong coordination (3)). The six types of coordination and the values attributed to them as a general rule are:

- No coordination (1)
- Pattern-setting (2; 2.5)
- Intra-associational coordination (1.5; 2; 2.5)
- State-sponsored coordination (1.5; 2; 2.5)
- State-imposed coordination (2.5; 3)
- Inter-associational coordination (2.5; 3)

Finally, capacity of implementation can be either low (1) or high (2).

The summary measure computed from the above indices has a scale (1-3). We extended the data provided by the earlier authors up to 2005 on the basis of EIRO Annual Reviews and EIRO National Reports for the period concerned.

FIGURE 25: BARGAINING COORDINATION INDEX, 1970Q1-2005Q4



Employment Protection Legislation (EPL)

Employment protection legislation is widely thought to be an important determinant of labour market inflexibility, although its actual empirical relationship with the overall unemployment rate provides mixed results (Nickell, 1997; Siebert, 1997).

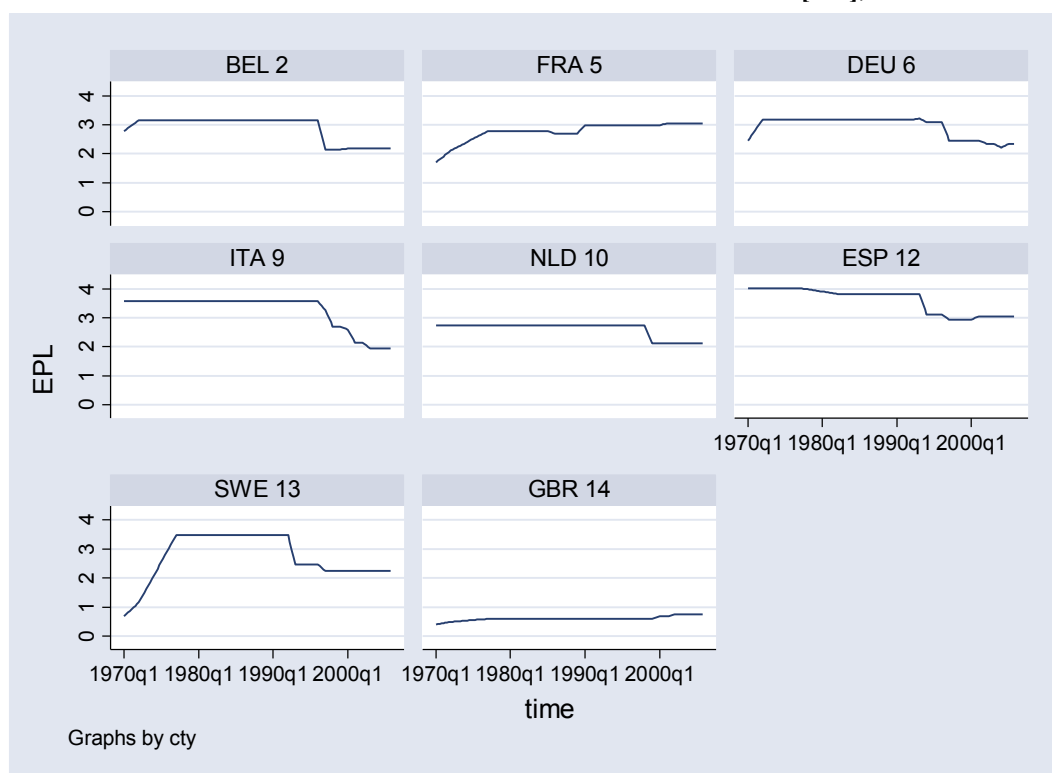
Data on this variable were until very recently extremely scarce (Nicoletti et al., 2001). This changed with the OECD Employment Outlook 2004, for which the OECD compiled an annual index of this dimension for the years 1985-2003⁷⁵. The version of the OECD indicator used here concerns both the regulation of permanent and temporary contracts, but excludes additional information regarding the regulation of collective dismissals, as this would considerably reduce the sample available. The indicator measures a range of

⁷⁵ See OECD Employment Outlook 2004 for details of the reforms underlying changes in the index; Young (2003) also provides a useful list of EPL-related reforms for a number of countries

legislative dimensions, notably including the difficulty of dismissal, procedural inconveniences, notice and severance pay provisions and restrictions placed on fixed-term contracts. Data beyond 2003 are assumed to display constancy.

For the period before 1985, we chained annual data from the LMIDB database to the OECD series (Nickell and Nunziata, 2001). These data came initially from Lazear (1990) and were based on the number of months of severance pay or notice for a blue collar worker with ten years of experience received upon termination without “cause”. The series is therefore less comprehensive than the OECD’s but represents the only available way of extending the data backwards.

FIGURE 26: EMPLOYMENT PROTECTION LEGISLATION [0-6], 1970q1-2005q4



3.6.4 General Remark

As a general remark for the institutional data presented, we note that there is often preciously little variation over the sample periods we will consider (mainly quarterly data over 1985-2005). This will make the identification of their effects problematic on single-country regressions, and may mean that even in panels most of the identification will be provided by a handful of countries, requiring caution in the interpretation.

4 EMPIRICAL RESULTS AND ANALYSIS

We report on the selected models below, including all acceptable restrictions. The graphs refer to the finally selected, parsimonious models prior to the imposition of long-run homogeneity restrictions so that the behaviour of the individual coefficients concerned may be examined. We also remind the reader that the coding of the tax wedge variables is such that they are to display negative signs to conform to our priors (with CTAX and ITAX as defined in Section 3.6.2):

$$\begin{aligned} ICTAX &= \ln(1 + CTAX) \\ IITAX &= \ln(1 - ITAX) \end{aligned} \tag{60}$$

We consider first the Eurozone economies examined in this study (Italy, France, Belgium, Spain, Germany, the Netherlands) and also report the results of our model when applied to two EMU-outsiders, the United Kingdom and Sweden.

We are able to consider the hypothesis that EMU increased labour market flexibility following its introduction in January 1999 for the six Eurozone countries in our sample. The extraordinary availability of expectations data for Italy and the Netherlands covering the 1970s also allows us to consider potential effects linked to their ERM membership.

A number of dates may be of particular relevance in this endeavour. The date of formation of the ERM (March 1979) may obviously be of interest. Anderton and Barrell (1995) proposed two further dates: 1982q1 (which the authors identify as corresponding to the formation of a ‘harder’ ERM) and 1983q2 (which follows a realignment of *all parities* in March 1983, and is considered by the authors as “*the last significant realignment*” in their sample period). However, there was another important realignment of all parities in July 1985 (during which the Dutch guilder appreciated by 2% and the Italian Lira depreciated by 6%). The Dutch Guilder was further appreciated in April 1986 and January 1987, by 3% each time (Fратиanni and von Hagen, 1992). Yet another possibility is the demarcation line drawn by Eichengreen and Wyplosz (1993) who emphasise the emergence of a “new EMS” from 1987 onwards, given the absence of any realignments between January 1987 and the 1992 crisis. In the interest of avoiding *a priori* judgements, none of these candidate dates is excluded in the analysis below.

The battery of stability tests⁷⁶ considered graphically from the PcGive output in what follows is composed of: the evolution of the RSS; the 1-step residuals; the standardised innovations; and the 1-step, breakpoint and forecast Chow tests (Hendry and Doornik, 2001).

⁷⁶ From an econometric perspective, we will notably consider stability at the end of the sample. In that case, the effect of EMU observations in updating the vector of coefficients will be a marginal one. However, the qualitative robustness of the findings to be presented, and their successful scrutiny under such a range of tests (e.g. both breakpoint and forecast Chow tests) applied at different sample dates, appears established. As a further note, we will be dealing with highly parametrised models, so that any structural breaks would likely be spread across many coefficients. Such general specifications are clearly to be preferred, although this effect may influence the small sample power of the tests applied to some extent.

4.1 Country-specific results

4.1.1 Italy

Modelling ITA_D8lCOMPTH; Estimation Sample: 1973q1-2005q2					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-0.537111	0.1352	-3.97	0.000	0.1207
ITA_D7lCOMPTH_1	0.855537	0.03356	25.5	0.000	0.8496
ITA_DlCOMPTH_8	0.102627	0.06348	1.62	0.109	0.0222
ITA_EXPINFLPBIS	0.0301217	0.01037	2.91	0.004	0.0684
ITA_D4lCPI_1	0.268759	0.04211	6.38	0.000	0.2616
ITA_D2lCPI_6	0.206081	0.06632	3.11	0.002	0.0775
ITA_D3lUNEMP	-0.0516128	0.01117	-4.62	0.000	0.1565
I1998:1	-0.0304326	0.007213	-4.22	0.000	0.1340
I1993:1	0.0516000	0.007206	7.16	0.000	0.3084
I1992:1	0.0273798	0.007190	3.81	0.000	0.1120
I1992:4	-0.0236381	0.007164	-3.30	0.001	0.0865
I2003:4	-0.0217680	0.007186	-3.03	0.003	0.0739
ITA_lRCOMPTH_8	-0.0812003	0.02281	-3.56	0.001	0.0992
ITA_lPRODIT_8	0.0812003*				
ITA_lUNEMP_8	-0.0226827	0.005725	-3.96	0.000	0.1201
ITA_lCOORD_8	-0.0241825	0.006274	-3.85	0.000	0.1144
sigma	0.00705685	RSS		0.00572690706	
R^2	0.997419	F(14,115) =	3175	[0.000]**	
log-likelihood	467.495	DW		1.89	
no. of observations	130	no. of parameters		15	
mean(ITA_D8lCOMPTH)	0.203195	var(ITA_D8lCOMPTH)		0.017071	
Specification Tests					
AR 1-5 test:	F(5,110) =	1.0605	[0.3863]		
ARCH 1-4 test:	F(4,107) =	2.0772	[0.0888]		
Normality test:	Chi^2(2) =	0.27360	[0.8721]		
hetero test:	F(23,91) =	1.4606	[0.1058]		
RESET test:	F(1,114) =	0.11312	[0.7372]		

SOURCE: own calculations; starred * coefficients have been restricted.

Greater availability of key data in Italy has allowed us to build a model for a considerably longer period than is possible in all other countries, except the Netherlands: 1973q1-2005q2. A sample extending back to the volatile period of the 1970s generates specific challenges, but also presents the opportunity of examining the potential link between *both* EMU *and* the EMS on the one hand, and labour market flexibility on the

other. A direct consequence of using this longer sample is that we are unable to rely on the OECD's standardised unemployment rate data (which are only available from 1982q1) and have to resort to the national definition for this variable, unlike for the other countries.

We model the two-year change in the hourly compensation of workers, with an equilibrium correction term lagged eight quarters. The dynamics for the dependent variable and the inflation terms in the equation indicated that this was the most appropriate specification. This is also consistent with *a priori* information available on the Italian wage bargaining process, whose collective agreements are known to have a usual duration of two years (European Commission, 2004).

The equation displays a speed of adjustment term of 8.1%. This is rather low, both in *absolute* and in *comparative* terms, when considering the rest of our sample countries. Curiously however, this is not a finding typical of the literature in the case of Italy. This issue is discussed in greater detail in Section 4.2 below. Suffice it to say for now that this is notably consistent with a view of the Italian labour market displaying sluggish adjustment, notably due to regional imbalances (Obstfeld and Peri, 1998).

A note should be made that the influence of inflation volatility on the wage bargain was examined, without any significant results despite the sample period covered. This issue is discussed in detail in the context of the Dutch wage equation, in Section 4.1.6.

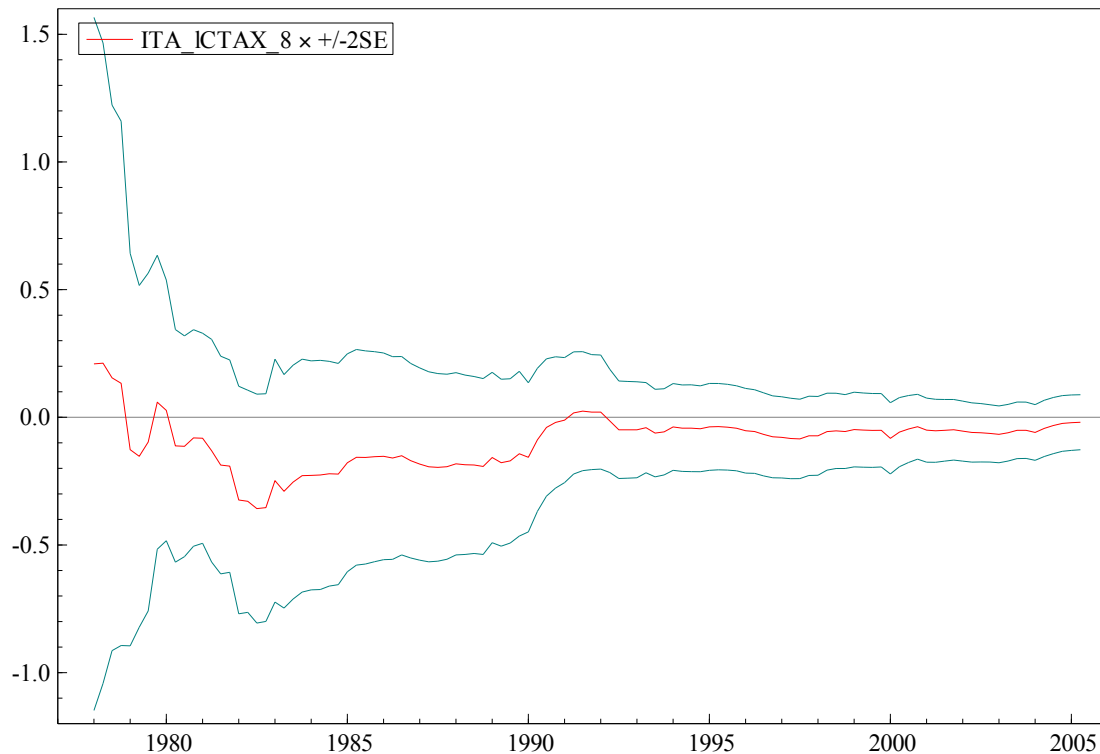
It was possible to impose the restriction that real hourly compensation rises one for one with productivity in the long-run⁷⁷. The equilibrium also features a strongly significant influence of unemployment in terms of dampening real hourly compensation.

However, we find that the relevant prices in equilibrium are consumer prices, since the indirect taxation variable CTAX was dropped. This represents a difference with the findings in Curto Millet (2004) and may be related to the much larger backwards sample available here. Indeed, Curto Millet found that the statistical significance of producer prices was a recent phenomenon in Italy, and that our ability to detect it was apparently linked with the increase in the standard rate of VAT by 1% in late 1997 (European Commission, 2005a) and the introduction of IRAP (*Imposta Regionale sulle Attività Produttive*) which may be responsible for the sharp increase in indirect tax receipts measured by CTAX for 1998. On the present sample and specification, the coefficient displays stable behaviour but does come closer to significance around 1997-1998. Interestingly, it also appears to be gaining importance around 1982. This is certainly linked to two important rises in the standard rate of VAT (up by 1% in January 1981 and 3% in August 1982, leaving the rate at 18%; European Commission, 2005a). The 1982 rise, in particular, was the most important to take place over our entire sample. The recursive development highlighted may also be linked to an increased relevance of producer prices in the wage bargain around that period, given other momentous labour market developments taking place at that time (see below).

⁷⁷ The restriction is easily supported by a Chi2 test.

Consistent with the insignificance of CTAX, we find the relative price of imports to be insignificant.

FIGURE 27: ITALY, RECURSIVE COEFFICIENT ON DIRECT TAXATION (LCTAX)



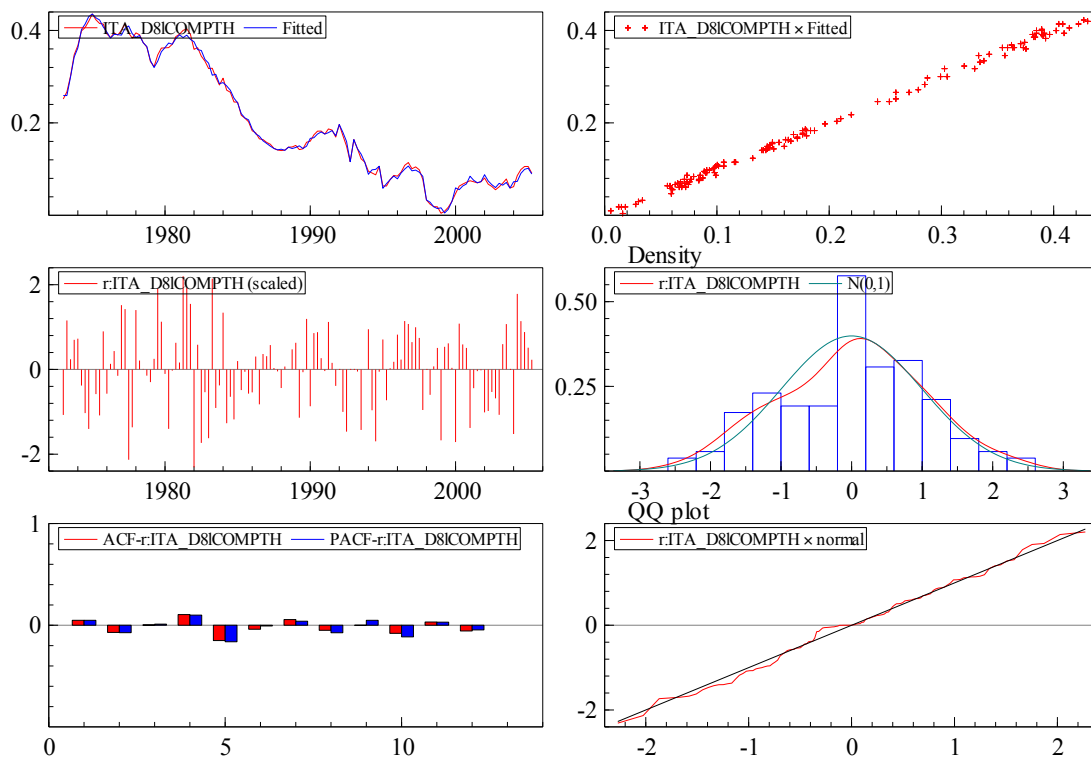
Bargaining coordination is found to have a significant role in equilibrium. The time variation of our index was displayed in Figure 25 and reflects a number of important changes. Bargaining coordination in Italy was mainly characterised by pattern-setting in the 1970s and much of the 1980s. Over the period, the state exerted some influence through its role as a public employer (with negotiations that were often pace-setting) and as a mediator in the reforms to the *Scala Mobile* (see below). However, its role was overall relatively secondary until the 1990s, when the coordination pattern shifted

towards state-sponsored coordination and the centralisation level shifted towards the central level with the key 1993 Tripartite agreement (Ochel, 2001).

The dynamics of the equation feature a highly significant coefficient for inflation expectations, confirming the finding of previous research that this variable plays an important role in Italy (e.g. Barrell and Dury, 2001). It is of particular interest to note that the variable enters without lags in the equation, unlike for most other countries in our sample. There are strongly significant terms relating to unemployment and past inflation dynamics, over the previous three and seven quarters respectively.

The equation is congruent as evidenced by the specification tests and provides an excellent fit in-sample (sigma of 0.71%). The fact that the model exhibits no signs of autocorrelation is particularly impressive given the sample period covered and the smoothness of the series. Figure 28 provides some graphical evidence confirming the congruence of our specification.

FIGURE 28: ITALIAN MODEL, GRAPHICAL ANALYSIS, 1973Q1-2005Q2



A close examination of some of the dummy variables introduced is especially interesting. The dummy variable concerning the first quarter of 1998 was necessary to account for a clear outlier in the series. This may reflect the tensions caused by the government's proposal to reduce the working week to 35 hours, hampering negotiations over the renewal of sectoral collective agreements. These talks resumed in the second quarter with the agreement of *Confindustria* (EIRO, 1998c).

As can be readily seen, we found it necessary to introduce three dummies in the period 1992-1993. This confirms the instability found around this time by previous research. Barrell and Dury (2001) found evidence of a significant structural break in 1993 and attributed it to the abolition of the wage indexation scheme known as '*Scala Mobile*'. This mechanism adjusted wages to inflation on a quarterly basis for all wage and salary

earners. Scheremet (2000) failed to locate a significant shift and, surprised by its absence, went on to argue that the recursive estimates nonetheless showed ‘clear changes’ and that the lack of significance may have been due to his limited sample post-1993. Lauer (1999) found a break in 1994 (unattributed), but discussed elsewhere in her paper the abolition of the *Scala Mobile* in 1993. In contrast, Curto Millet (2004) did not find evidence of instability in this period and argued that a detailed understanding of the history of the *Scala Mobile* made it an unlikely culprit for any instability. How can these findings be reconciled?

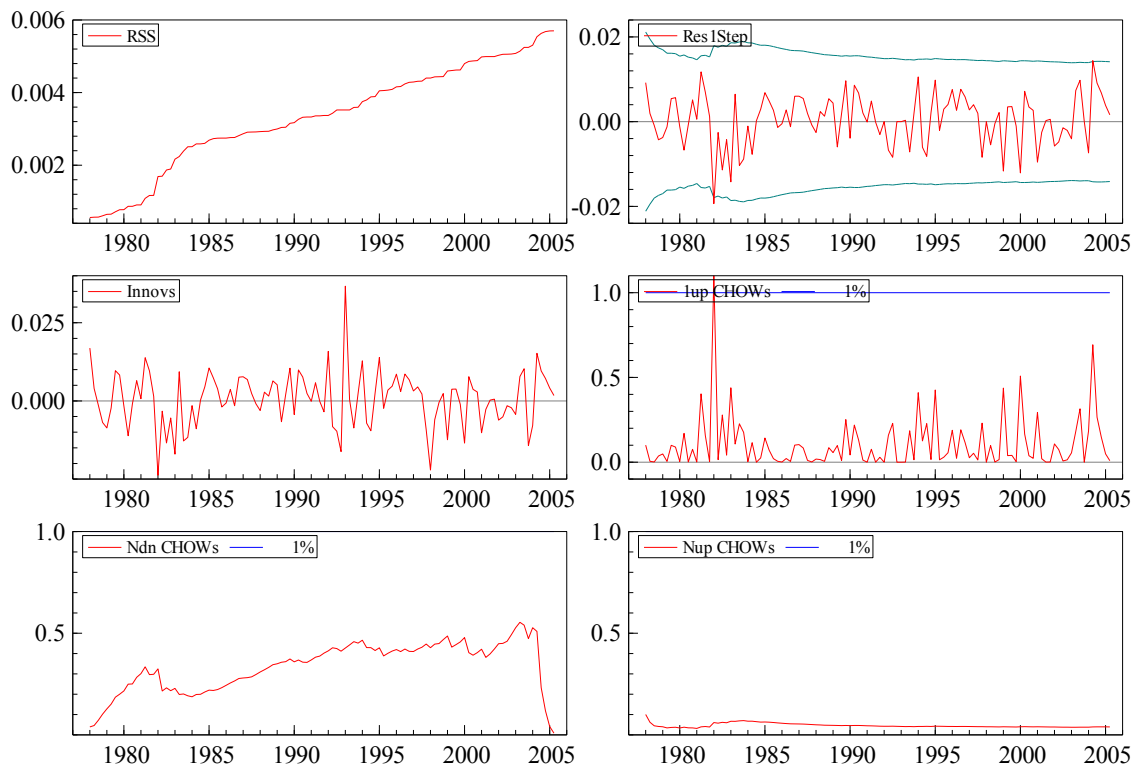
The *Scala Mobile* had been in force in Italy since the post-World War II period, but was importantly revised in 1975. The indexation provided was higher for low wages (Erickson and Ichino, 1993). The subsequent inflation meant that the mechanism quickly became a dominant influence on earnings increases and pay equality, even beyond the direct effects of collective bargaining; indeed, by the end of the 1970s *Scala Mobile* payments “constituted two-thirds to three-quarters of the earnings increases of Italian workers and had substantially compressed both inter-industry and inter-occupational wage differentials” (Flanagan, 1988). The story of its subsequent dismantling is one of a long and difficult process. Labour market tensions related to it soared in 1982, with *Cofindustria* repeatedly demanding the repeal of the mechanism (e.g. March 11, June 1) and the Bank of Italy’s Governor Ciampi weighing in to propose its stop (March 31). On 22 January 1983, the Italian government and the unions reached an agreement to revise the *Scala Mobile*, thereby reducing indexation by 15%. In 1986 a referendum succeeded in reducing indexation even further, making it “almost negligible” (Russo and Veredas,

2000). Thus, effects of the indexation mechanism should mainly be sought in the 1980s rather than in the largely *de jure* abolition of the early 1990s.

Much of the literature examining this period has found considerable evidence of instability in Italian wage equations. Anderton et al. (1992) found a structural break in 1982, with significant roles for expectations and unemployment *post-break* only; these findings were confirmed in Anderton and Barrell (1995). Artis and Ormerod (1991) found such disturbance in their regressions that they chose to exclude data in the period 1979q4-1981q4. Given the institutional background described, it is puzzling that findings of instability or coefficient changes are not replicated in other studies covering the period (e.g. Barrell and Dury, 2001; Scheremet, 2000; Lauer, 1999).

Our results here do provide evidence of some instability in the early 1980s: Figure 29 confirms this by an accelerated build-up of the RSS and the evidence from the 1-step residuals and 1-step Chow test. 1982q1 seems to be highlighted, in particular. The equation is otherwise remarkably stable; in particular, the stability tests below provide no evidence in favour of an “EMU effect” in the structure wage bargaining.

FIGURE 29: ITALIAN MODEL, STABILITY TESTS



The early 1980s was a period of momentous changes in the Italian economy, making it difficult to distinguish causally between the different potential sources of instability. Nevertheless, it was already noted that 1982q1 had been singled out by Anderton and Barrell (1995) as the start of a ‘harder ERM’. Could this be a sign of an “EMS effect” in the Italian wage structure? We propose to study some of the subcomponents of our equation to explore the issue.

Figure 30 plots the key recursive EqCM coefficients. As can be seen, there is a shift around 1981q4 / 1982q1 which led to equilibrium-correction behaviour being largely “turned off” in terms of the interaction between real hourly compensation and the productivity trend, for the remainder of the 1980s. These are hardly the consequences one would expect from the membership of a hard currency union; indeed, the reverse

developments would have been more logical. Instead, this resembles rather more closely the behaviour of a labour market working on “momentum” (i.e. the short-run dynamics in the equation), the *Scala Mobile* and the high-inflation environment being consistent with such a view. The re-instatement of equilibrium behaviour may then have taken some time to be identified econometrically.

It is true that the lagged unemployment coefficient regains significance from 1983q2 onwards, following Anderton and Barrell’s (1995) last “significant” realignment. However, given the above and the existence of subsequent realignments for Italy, it seems inappropriate to attribute this to an EMS effect of any kind.

FIGURE 30: ITALIAN MODEL, RECURSIVE EQCM COEFFICIENTS

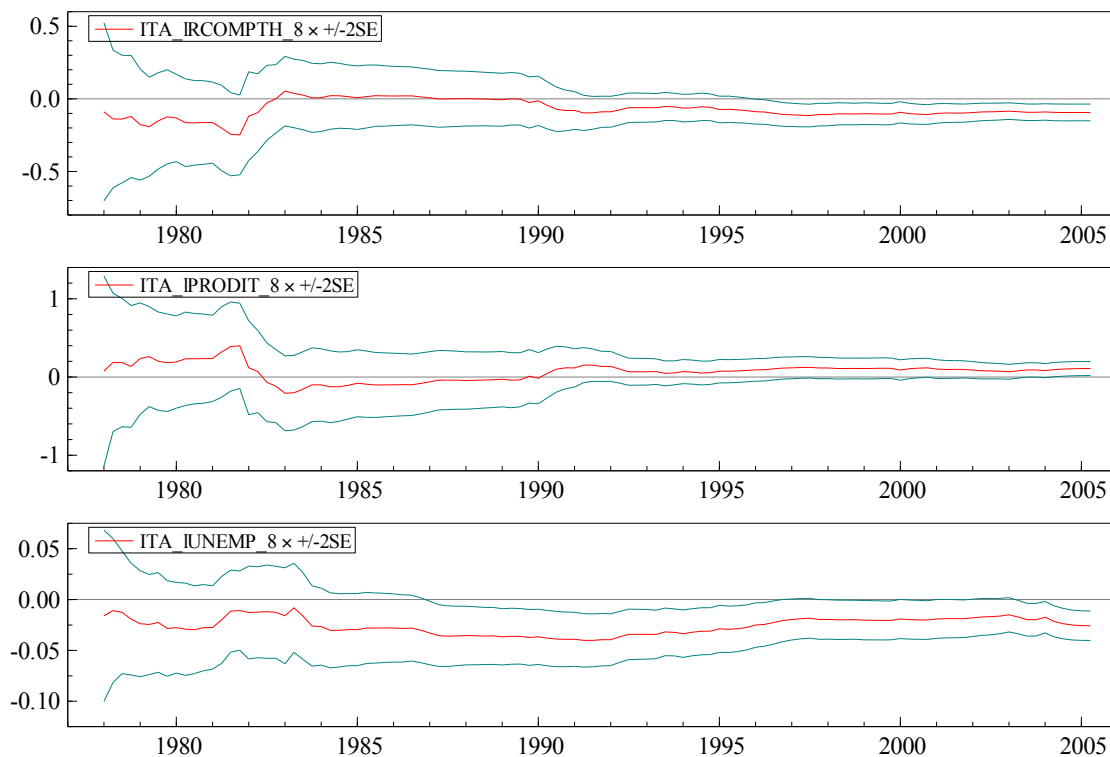


FIGURE 31: DEPENDENT VARIABLE AND UNEMPLOYMENT, RECURSIVE

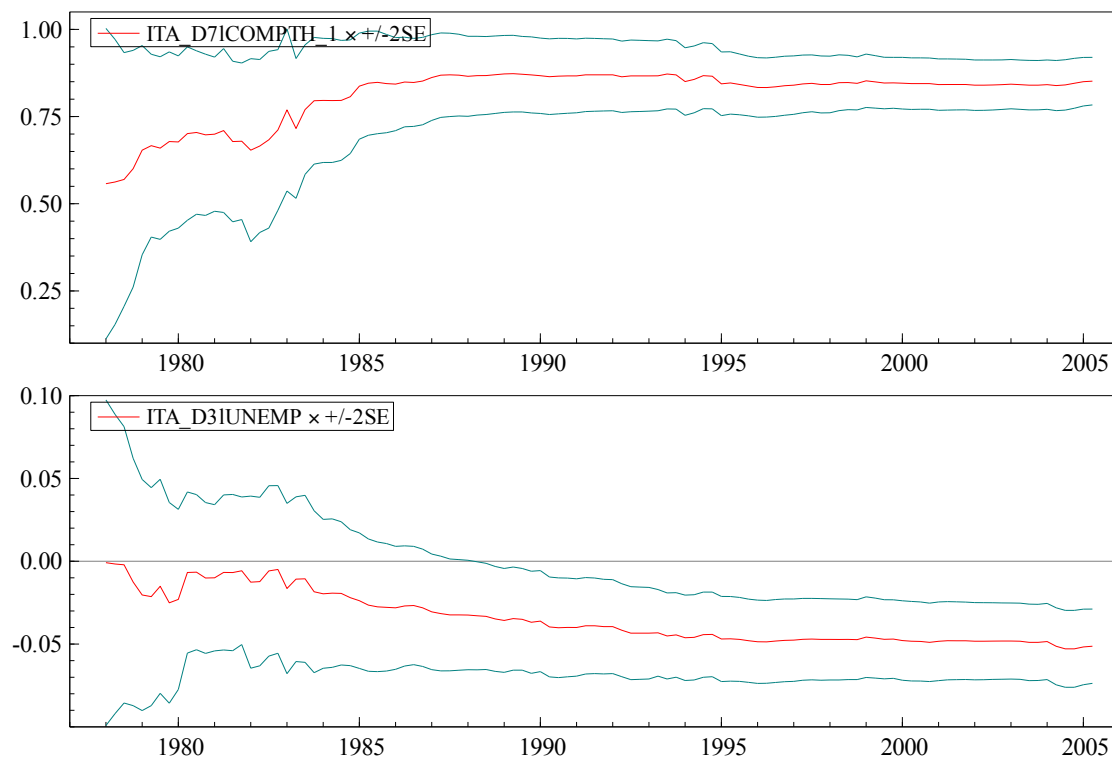
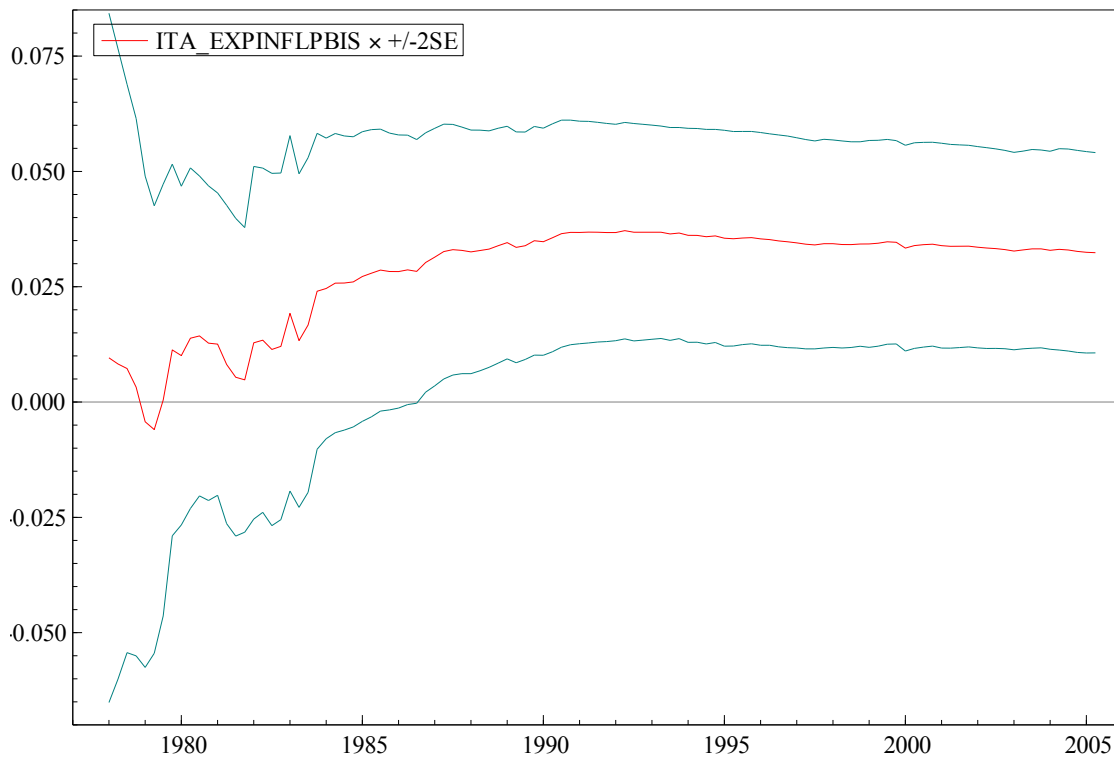


Figure 31 shows that the early 1980s saw a marked increase in the importance of dynamics in the dependent variable. The same is true for unemployment dynamics. Thus, our results for *both* the unemployment level and dynamics confirm the finding by Anderton and Barrell (1995) of a greater role for unemployment after the ‘break’ (although the authors only found this to be the case in level terms). Figure 32 is crucial and shows that the importance of inflation expectations in wage bargaining increased greatly over the period, coinciding with the dismantling of the *Scala Mobile*. In turn, this reinforces the argument that the situation in the Italian labour market in the early 1980s is to be principally related to this abolition and the surrounding circumstances.

FIGURE 32: ITALY, INFLATION EXPECTATIONS, RECURSIVE COEFFICIENT



Of course, it is possible to argue that the abolition of the *Scala Mobile* would have been politically impossible without ERM membership (Anderton and Barrell, 1995). It is indeed very plausible that the ERM was an important motivating factor in speeding its demise, especially given the hardening of the currency commitment and the need to contain ERM inflation differentials to preserve competitiveness. Nevertheless, this represents an additional layer of argument, which provides no evidence for a particular *direct* link between the exchange rate regime and labour market *flexibility*.

To close our consideration of the *Scala Mobile*, it is necessary to comment on the events surrounding its largely *de jure* abolition in the early 1990s that have generated such confusion in the past. In 1990, tension between employer organisations and unions over the cost of living index increased dramatically. In the face of the unions' (CGIL, CISL,

UIL) refusal to negotiate, *Cofindustria*'s president unilaterally terminated the *Scala Mobile* agreement, generating strikes. The government intervened by passing a law extending the validity of the cost of living index mechanism until December 1991 to provide room for negotiation (EMIRE, 2005). However, all the national labour contracts renewed in the period 1990-91 *already* did not include indexation, in anticipation of the upcoming social pact (ILO, 2005a). Negotiations for it were unfruitful at first, leading to the expiry of the law guaranteeing indexation. In the end, the mechanism was formally abolished in July 1992 by a national multi-industry agreement. This included provisions for workers to receive 20,000 lire per month in exchange for 13 months covering the 1992-93 period, as well as a freeze for the whole 1993 on company level wage increases. Following this, the key July 1993 tripartite agreement introduced the forecast rate of inflation and actual rate of inflation as new reference indicators for the protection of the purchasing power of pay, to be defined at industry-wide bargaining level. The agreement also dealt with numerous other issues (e.g. collective bargaining, employment policies).

As this institutional overview confirms, instability found around the years 1992 and 1993 cannot be related to the *Scala Mobile* directly, as the mechanism had effectively already been abolished. Notwithstanding this, the bargaining environment of the time can be described as one of “*mounting economic and political crisis*” (Pérez, 2000), some features of which have just been outlined. Such instability in labour market bargaining may be behind our dummy variables in the period.

Returning our attention to EMU, Figure 33 considers once more the issue of a possible “EMU effect”, this time by presenting forecasts covering the EMU period from 1999q1.

The forecasts for the entire post-EMU period appear to be entirely satisfactory. Similarly, the key recursive EqCM coefficients appear to be very stable. Thus, the data provide no evidence for an “EMU effect”. This may be seen as particularly damaging to the case for EMU increasing labour market flexibility. Indeed, given Italy’s history of competitive devaluations (Lauer, 1999), it should arguably have displayed the most glaring evidence in this respect.

FIGURE 33: ITALIAN MODEL, FORECASTS FROM 1999:1

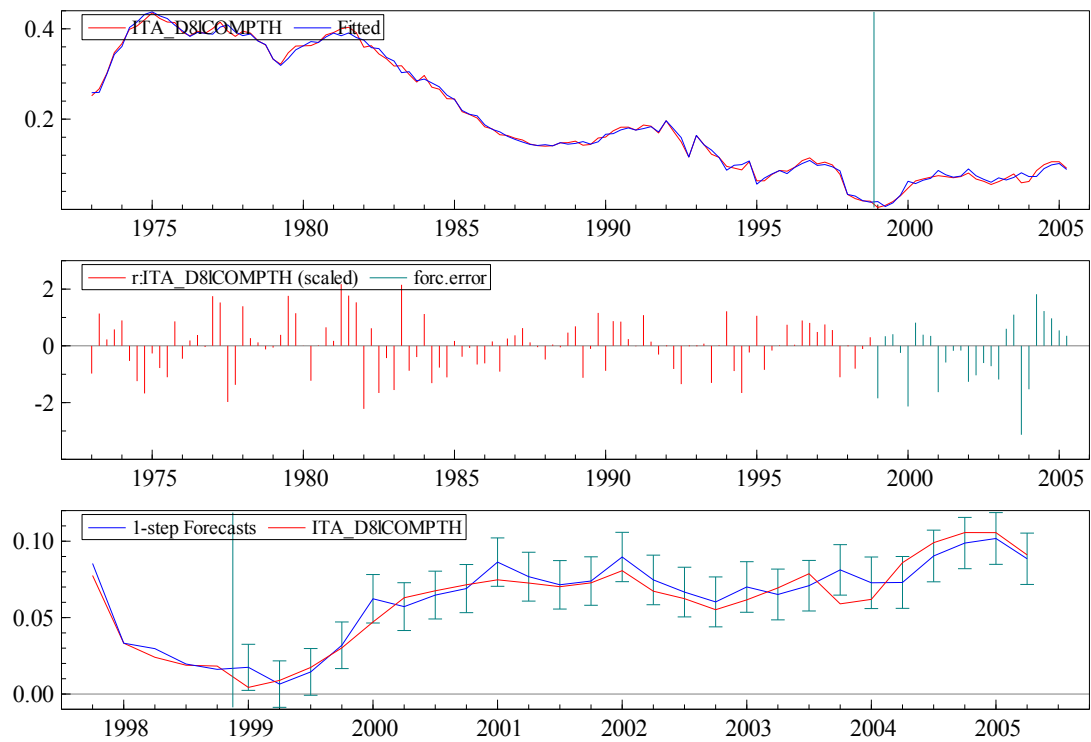
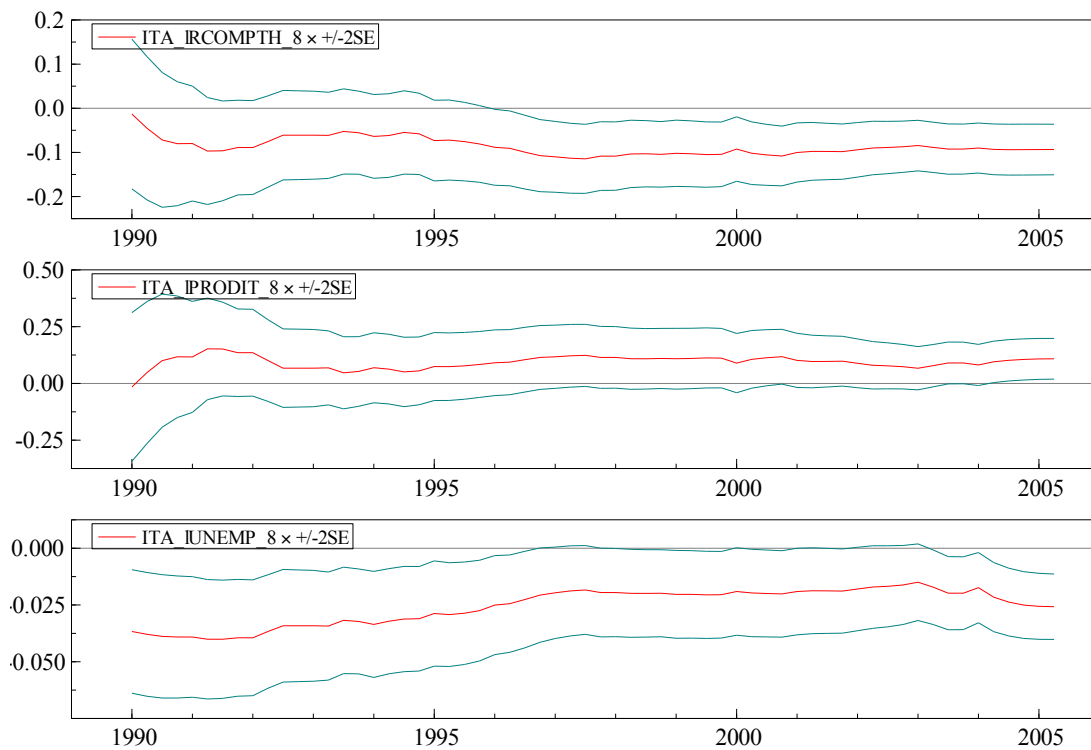


FIGURE 34: ITALIAN MODEL, RECURSIVE EqCM COEFFICIENTS



4.1.2 France

Modelling FRA_DlCOMPTh; Estimation Sample: 1982q4-2005q2					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-2.44914	0.4439	-5.52	0.000	0.2943
FRA_DlCOMPTh_1	-0.547508	0.08029	-6.82	0.000	0.3891
FRA_DlCOMPTh_2	-0.608160	0.07978	-7.62	0.000	0.4432
FRA_DlCOMPTh_3	-0.209784	0.05826	-3.60	0.001	0.1508
FRA_D3lCPI_1	0.282254	0.06885	4.10	0.000	0.1871
FRA_D2lU_1	-0.0846952	0.01137	-7.45	0.000	0.4318
DlFRA_HOURS	-0.803565	0.1738	-4.62	0.000	0.2265
FRA_DlPRODI	0.160979	0.08170	1.97	0.053	0.0505
I1993:1	0.00976081	0.003344	2.92	0.005	0.1045
I1990:1	-0.0138492	0.003330	-4.16	0.000	0.1915
I1984:3	-0.00802708	0.003455	-2.32	0.023	0.0689
I1998:1	-0.00976914	0.003240	-3.02	0.004	0.1108
FRA_lRCOMPTh_3	-0.308699	0.06296	-4.90	0.000	0.2477
FRA_lPRODIT_3	0.361997	0.06895	5.25	0.000	0.2741
FRA_lU_3	-0.0499963	0.006817	-7.33	0.000	0.4242
FRA_lUDENS_3	0.107004	0.01821	5.88	0.000	0.3212
FRA_lBD_3	0.0618324	0.02663	2.32	0.023	0.0688
FRA_lBRR_3	0.110637	0.05094	2.17	0.033	0.0607
sigma	0.00310698	RSS	0.000704693692		
R^2	0.838375	F(17,73) =	22.27 [0.000]**		
log-likelihood	406.348	DW	2.18		
no. of observations	91	no. of parameters	18		
mean(FRA_DlCOMPTh)	0.0100115	var(FRA_DlCOMPTh)	4.79126e-005		
Specification Tests					
AR 1-5 test:	F(5,68)	=	1.0090	[0.4192]	
ARCH 1-4 test:	F(4,65)	=	0.47070	[0.7570]	
Normality test:	Chi^2(2)	=	1.4606	[0.4818]	
hetero test:	F(30,42)	=	0.95701	[0.5438]	
RESET test:	F(1,72)	=	2.5260	[0.1164]	

SOURCE: own calculations; starred * coefficients have been restricted.

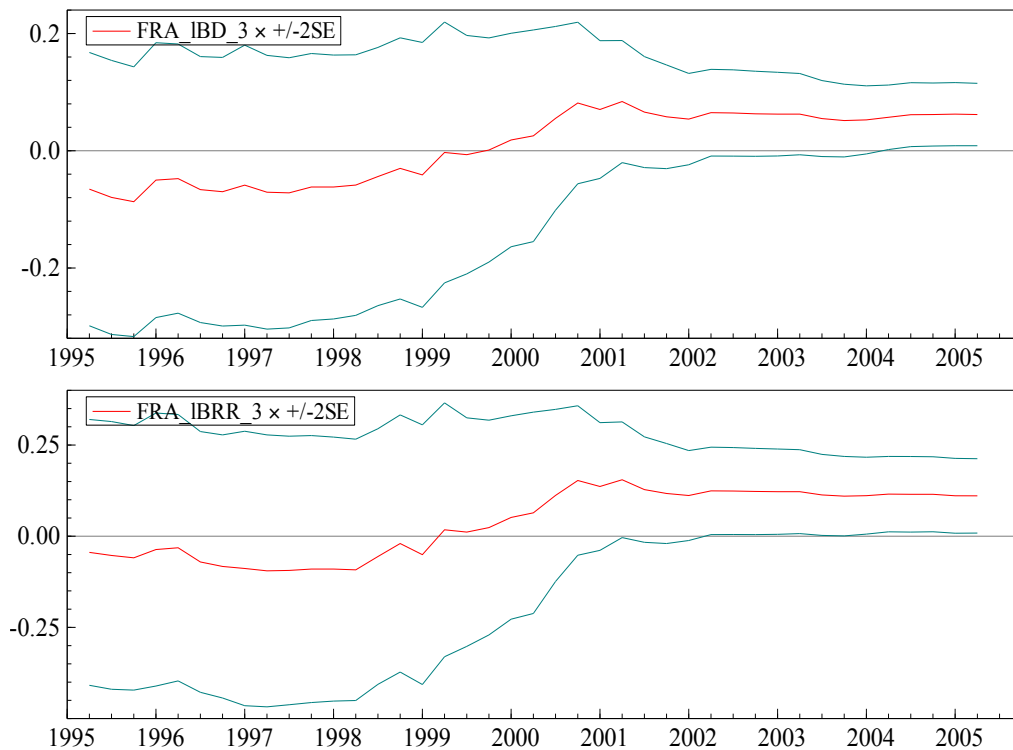
The absence of expectations from the French wage equation evidenced in Chapter 1 (and attributed there to measurement error) allows us to use a somewhat larger sample covering 1982q4-2005q2. The French specification models quarterly changes in the hourly compensation of workers, with the equilibrium correction term appearing at a lag of three quarters. This was suggested by the dynamics of the data expressed through the

equation and is consistent with the specification found in Curto Millet (2004). The speed of adjustment term has a value of 31% and the productivity trend is quantitatively important and statistically strongly significant. It may seem from a comparison of the coefficients for the lagged level of the dependent variable and the productivity trend that a long-run homogeneity restriction can be imposed. However, although the equation remains congruent when this is carried out, a Chi2 test of the restriction implied yields a strong rejection; accordingly, we maintain the model in the shape presented above. Unemployment is a significant dampening influence for real hourly compensation in equilibrium.

Three institutional variables are kept in the analysis. The first of these is union density. The bargaining power of unions in France is well-known. It may seem surprising that this is manifested in our equation through the *density* variable rather than the bargaining *coverage* variable. Indeed, France is the country with the lowest union density in our sample, but the highest bargaining coverage. We might therefore have expected the time series movement of the latter variable to be of greater relevance to the wage bargain. However, as already discussed in Section 3.6.3 above, the data available on the coverage dimension are somewhat “rough”, most of it being interpolated from data points five years apart. This may go some way towards explaining this finding. The other two variables retained represent different facets of the benefit system, namely the duration of benefits and the replacement rate. As can be seen from Figure 35, their coefficients only become significant recently after experiencing growth in the period 2000-2001. This period was characterised by an overhaul of the French jointly-managed unemployment insurance scheme (*UNEDIC*), which was agreed between the government and most of

the social partners. This notably involved an extension of unemployment insurance coverage, and introduced some degree of “active labour market policy” through the *PARE* scheme.

FIGURE 35: BENEFIT DURATION AND REPLACEMENT RATE COEFFICIENTS



The dynamics of the French wage equation are unusual and therefore especially interesting. As already highlighted, the absence of inflation expectations sets this equation apart from those of all the other countries. There is evidence, on the other hand, of highly significant dynamics reflecting compensation for past inflation (although the coefficient on inflation in the present quarter is insignificant at the 10% level). Taking the magnitude of the speed of adjustment coefficient as a benchmark, it can be assessed that the coefficient on these dynamics is important. This is hardly a surprise in the case of

an economy like France, due to the exceptional circumstances created by its minimum wage system.

A statutory minimum wage framework has existed in France since 1950 (SMIG – *Salaire Minimum Interprofessionnel Garanti*); this was replaced in 1970 by the current system (SMIC – *Salaire Minimum Interprofessionnel de Croissance*) (CSERC, 1999). The French minimum wage has a number of important characteristics that set it apart. First, its *level* in absolute terms is among the highest for European countries – in 2004, 1,286 Euros monthly for workers on a 39-hour week (EIRO, 2005a). Second, the “*Kaitz Index*” for the French system (i.e. the minimum wage as a proportion of average monthly gross earnings; see Kaitz, 1970 and Dolado et al., 1996) again places the country in the highest group in the EU (EIRO, 2005a). Third, the *proportion* of workers covered by the system is by far much higher than that in other European countries, as shown by the Eurostat data in Table 30. Earlier estimates show a similarly high level of coverage in France. For instance, Bazen and Martin (1991) quote an overall coverage estimate of 12% of all wage and salary earners for 1987 from the *Ministère du travail, de l'emploi et de la formation professionnelle*.

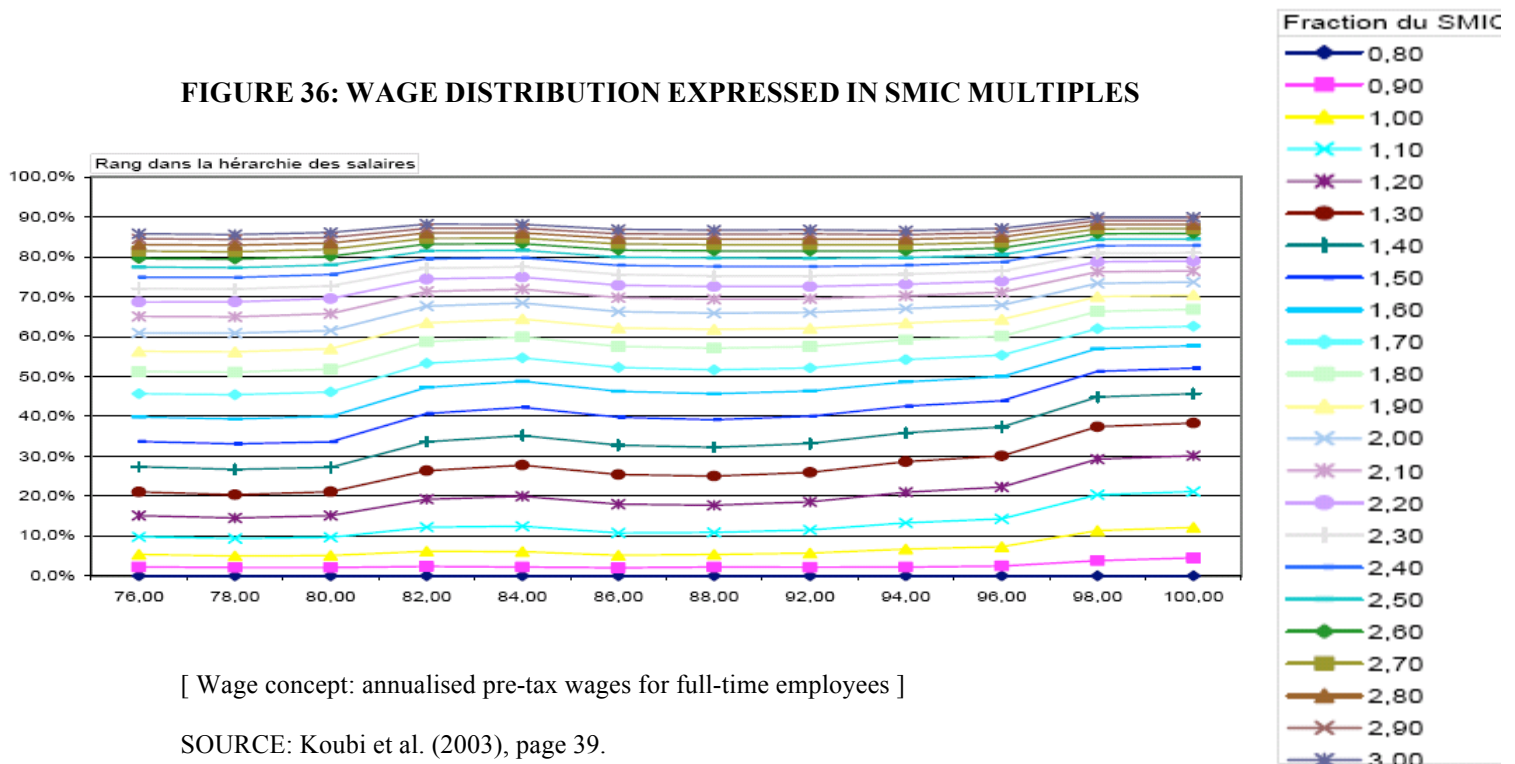
TABLE 30: FULL-TIME EMPLOYEES EARNING THE MINIMUM WAGE (%)

Year	Country			
	Spain	France	Netherlands	UK
1999	2.6	12.8	2.2	2.5
2000	1.4	13.6	2.1	1.4
2001	0.85	13.9	2.2	1
2002	0.78	14	2.28	1.8
2003	0.86	13.4	2.24	1.2
2004	0.77	15.6	2.07	1.4
2005	-	-	-	1.8

SOURCE: Eurostat, Labour Market Indicators

Finally, the numbers in Table 30 probably severely underestimate the *overall* influence of the system in France. Anderton and Barrell (1992, 1995) quote a figure of 40 per cent of the workforce as being covered “*directly or indirectly*” by the minimum wage. The authors provide no evidence of how this number was arrived at or what they meant precisely. Nevertheless, their assessment appears to be broadly correct. It stands to reason that wages above the SMIC but nevertheless close to it will be importantly influenced by SMIC increases. Koubi et al. (2003) provide helpful decompositions of the distributions of wages in terms of fractions of the SMIC for the years 1976 to 2000, which we reproduce in Figure 36. As can be seen, around 40% of workers earned roughly 1,3 times the SMIC or less in 2000. It seems reasonable to suppose that their wage bargains will be heavily influenced by SMIC developments.

FIGURE 36: WAGE DISTRIBUTION EXPRESSED IN SMIC MULTIPLES



Thus, the evidence presented suggests that the SMIC system can only be expected to exert an important influence on French wage bargaining. The consequences of this depend on the adjustment mechanism in place for the minimum wage. The French SMIC is adjusted annually (normally in July) by government decree, following a consultation of the social partners (EIRO, 2005a).

Various factors are taken into account in this decision (e.g. the forecast for the CPI; labour productivity). However, the two most important components are *past inflation* and the evolution of the *basic hourly wages of workmen* (SHBO – *Salaire Horaire de Base des Ouvriers*), yielding the “formula” (CfDT, 2005; EIRO, 2000b; Bazen and Martin, 1991):

$$SMIC\ Increase = Inflation + 1/2 * [SHBO - Inflation]$$

Other factors may be taken into account (e.g. the CPI forecast, labour productivity) (EIRO, 2005a). The government must also provide compensation between SMIC adjustments in case annual inflation moves beyond 2%, and may also increase it at other times at its discretion.

Thus, this suggests that backward-looking adjustments in the wage bargain may play an especially important role, and provides intuition and support for the results of our wage equation. We do not believe, however, that this would be sufficient to justify the complete absence of expectations from the specification and still ascribe that finding to measurement error, as underlined in Chapter 1 (Section 7.3.3). However, a high-degree

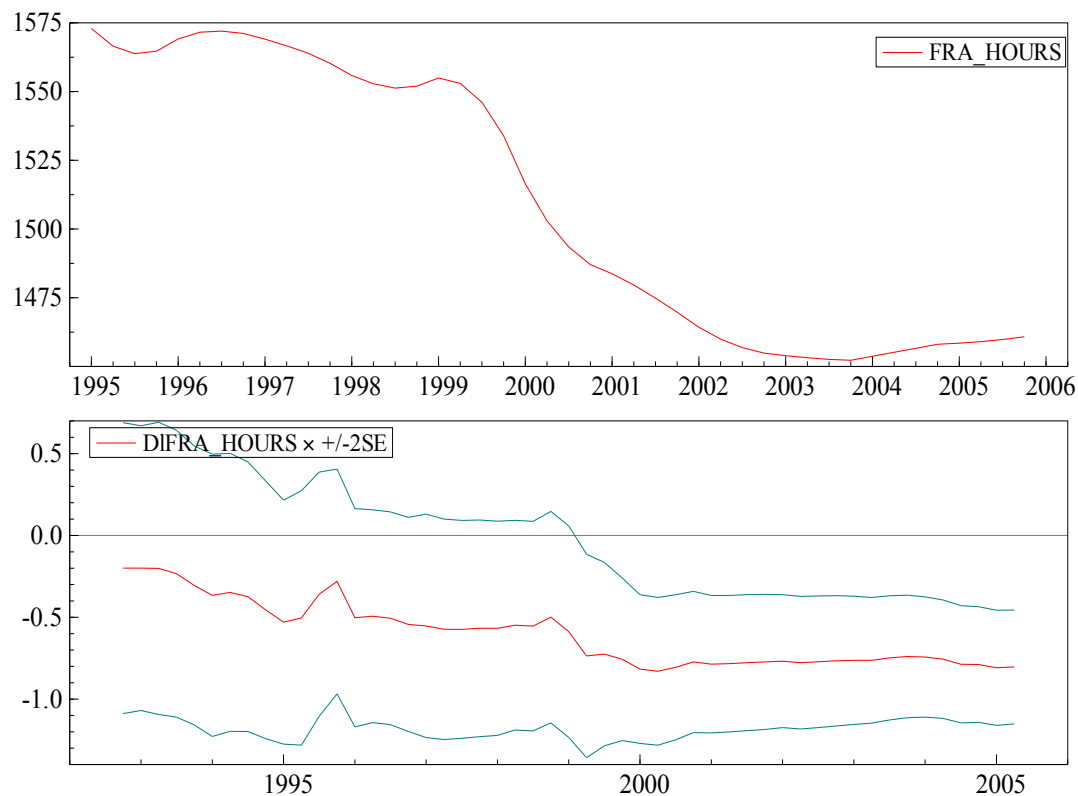
of backward-lookingness in France may have made expectations effects “less obvious”, and therefore made it easier for measurement error to mask their importance.

The dynamics of our wage equation also contain a strong role for unemployment and a short-run productivity effect (examined below). Furthermore, and exceptionally, the dynamics include a variable which measures the change in average hours worked by employees from one quarter to the other, DIFRA_HOURS. The rationale for its presence lies in Curto Millet (2004), who observed substantial under-prediction of growth in hourly compensation by the post-1999 forecasts of the model used there. It was argued that this was due to the introduction of the 35 hour week reforms in France.

Indeed, lower working hours at similar pay levels – as was largely the outcome in France (EIRO, 1999a) – should have the effect of increasing hourly compensation substantially. The timeline of events is supportive of this interpretation. The first “Aubry law” on “guidelines and incentives relating to the reduction of working time” was issued in June 1998 and concerned the reduction of the statutory working week from 39 hours to 35 hours by 1 January 2000 for companies with more than 20 workers. This was completed by the second and definitive “Aubry law” of December 1999 which included more detailed legal provisions. Already in 1999, collective bargaining in France was clearly marked by the implementation of the first law (*Ministère de l’Emploi et de la Solidarité*, 2000). This is transparently reflected in the top panel of Figure 37, which shows a very sharp fall historically in average hours worked starting in 1999.

These observations suggest that the change in working hours should be part of the model. This is confirmed by the data; indeed, the coefficient on the change in hours is strongly significant (at the 1% level) and quantitatively important. The recursive plot of this coefficient is particularly interesting (bottom panel of Figure 37) as it reflects empirically the timing of the reforms known from the institutional information just discussed. It is clear from the figure that the coefficient visibly increases in magnitude and shifts towards significance in 1999.

FIGURE 37: AVERAGE HOURS PER EMPLOYEE AND RECURSIVE COEFFICIENT



Changes in working hours are delicate matters for wage equations. First, it is not always straightforward to find data on this dimension at the right frequency and which is such as to be consistent with the employment data used in deriving total hours. Second, working time changes may also deteriorate the accuracy of the productivity measure used,

depending on whether it is computed as output per head (as the OECD data used here are) or per hour. Nevertheless, the evidence presented here suggests that they can play a significant role and ought to be considered where necessary.

Dummy variables were introduced to control for outliers in the first quarters of 1990, 1993 and 1998. The dummy introduced in early 1998 is almost certainly due to the particularities of the period preceding the introduction of the guidelines and incentives law on the reduction of working time (*loi d'orientation et d'incitation relative à la réduction du temps de travail*), which came into effect on 13 June 1998. Indeed, the important activity in 1998 in company-level bargaining was largely centred around the working time issue; similarly, although sectoral bargaining was active on this issue, it can otherwise be described as slow (EIRO, 1998d). Pay negotiations did not make much progress as employers tried to preserve bargaining flexibility in this area for the negotiations on working time reduction. In fact, employers threatened a termination of sectoral collective agreements to achieve this. In some cases they acted on the threats, an example being the banking sector (EIRO, 1998d). Hence the conclusion that the dummy in question is likely to be related to a 'waiting effect' before the introduction of the 35-hour week law.

We also included a dummy for an outlier in 1984q3. This is situated in the middle of a period of important economic changes for France, marked by the reversal of the expansionary policies of 1981-1982 known as the Mitterand U-Turn. Previous authors have also found instability around this point: Artis and Ormerod (1991) excluded from their regression "*a lengthy period of disturbance*" from 1982q4-1985q4; later, Artis and

Ormerod (1996) split their sample in 1983q2 to allow for the Mitterrand U-Turn. While we detect no such dramatic developments in our wage equation, the specificities of the period should certainly be borne in mind.

The specification tests showed the model to be congruent, with an extremely small sigma (0.31%). The adequacy of the specification is confirmed by Figure 38, suggesting that we turn to the consideration of possible “EMU effects”. Figure 39 presents a post-1999 forecast analysis. As can be seen, these are highly unsatisfactory, with considerable under-prediction of actual developments and must be reflecting certain coefficient changes. Figure 40 shows that the forecasts from 2001 onwards are entirely adequate, thereby locating the relevant coefficient changes in the period 1999-2000. However, the stability tests in Figure 41 show the equation to be stable, suggesting that the changes that took place did so in a smooth way rather than causing a serious break. In fact, three coefficients are found to change. The first and most important is the coefficient on working time, which shifts from -0.49 to -0.80 in little more than two years (Figure 37). As already emphasised, a working time reduction carried out without a proportionate reduction in pay straightforwardly translates into increases in hourly compensation. This is what happened in France, and is the fundamental reason for the misforecasting of our model as the coefficient adapts.

The other two changes are subsidiary in comparison. The coefficient on the lagged hourly compensation of workers shifts down in magnitude post-1999 (from -0.39 to -0.31), a development that is not mirrored in the productivity coefficient; this can be seen in Figure 42 and may contribute to our inability to impose a long-run homogeneity

restriction for these variables. The change in the IRLCOMP_{TH} coefficient matters due to a clear acceleration in the underlying variable (Figure 44). With *ceteris paribus* coefficients this would push the predictions of hourly compensation growth from our wage equation downwards, as observed. This development acts in a direction opposite to what might have been hoped by the theory set out earlier and is not a sign of a positive “EMU effect”. However, this change may in part be reflecting the absence of expectations from the equation. As will be explored in Chapter 3, Section 2.1, these increased dramatically with the advent of EMU. Had measurement error not plagued the EC data, it is possible that this would have kept the forecasts of the model on track. Finally, the coefficient on a short-run productivity effect DIPRODI becomes nearly significant, shifting from 0.097 to 0.161. This may be picking up certain short-term disturbances in post-EMU wage bargaining, but its presence does not materially influence the other coefficients.

Figure 42 also emphasises the stability of the long-run unemployment coefficient, which is not suggestive of any important “EMU effect”.

FIGURE 38: FRENCH MODEL, GRAPHICAL ANALYSIS, 1982Q4-2005Q2

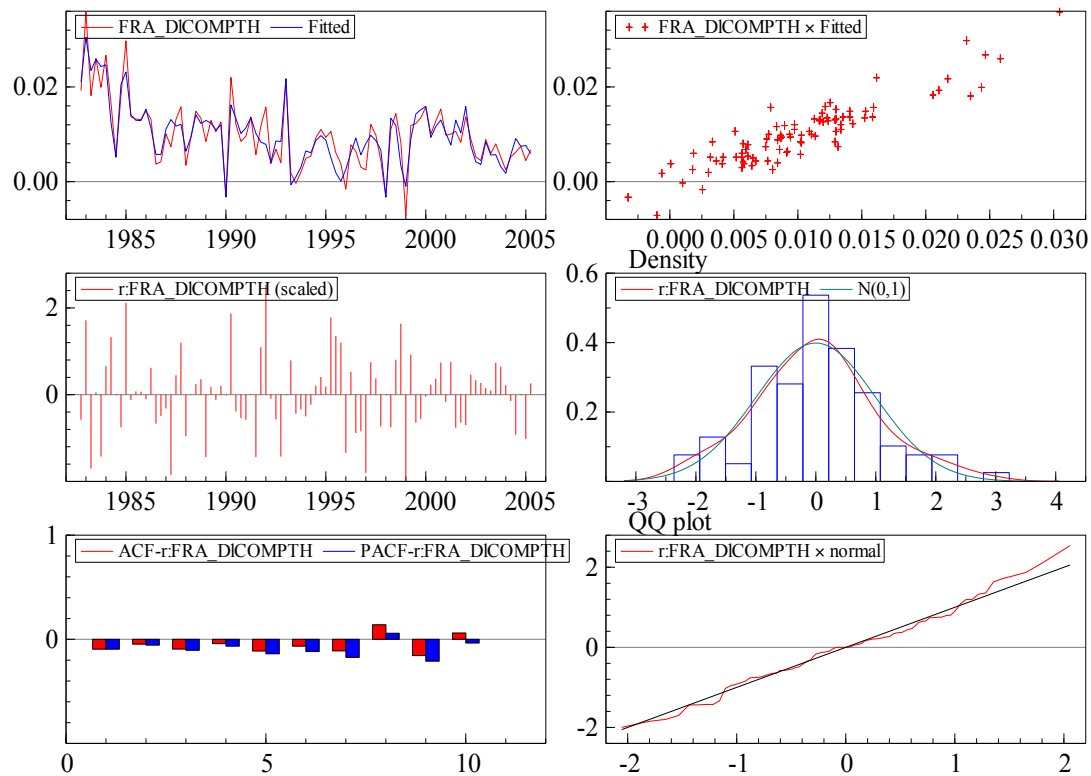


FIGURE 39: FRENCH MODEL, FORECASTS FROM 1999:1

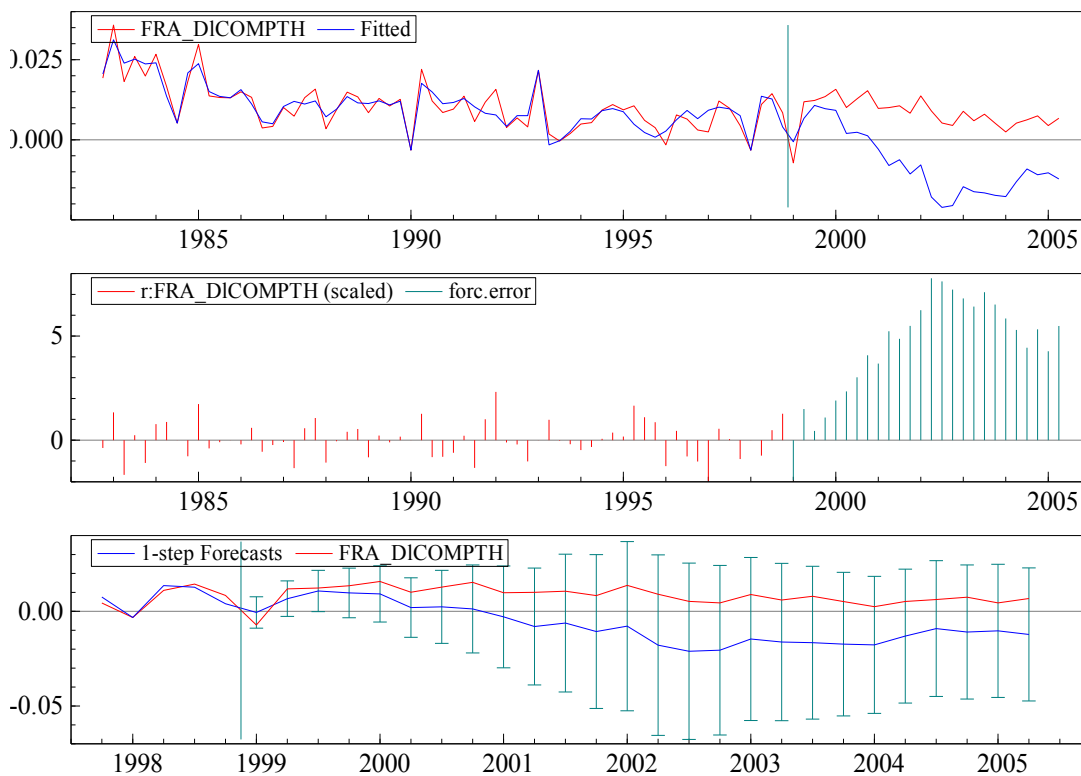


FIGURE 40: FRENCH MODEL, FORECASTS FROM 2001:1

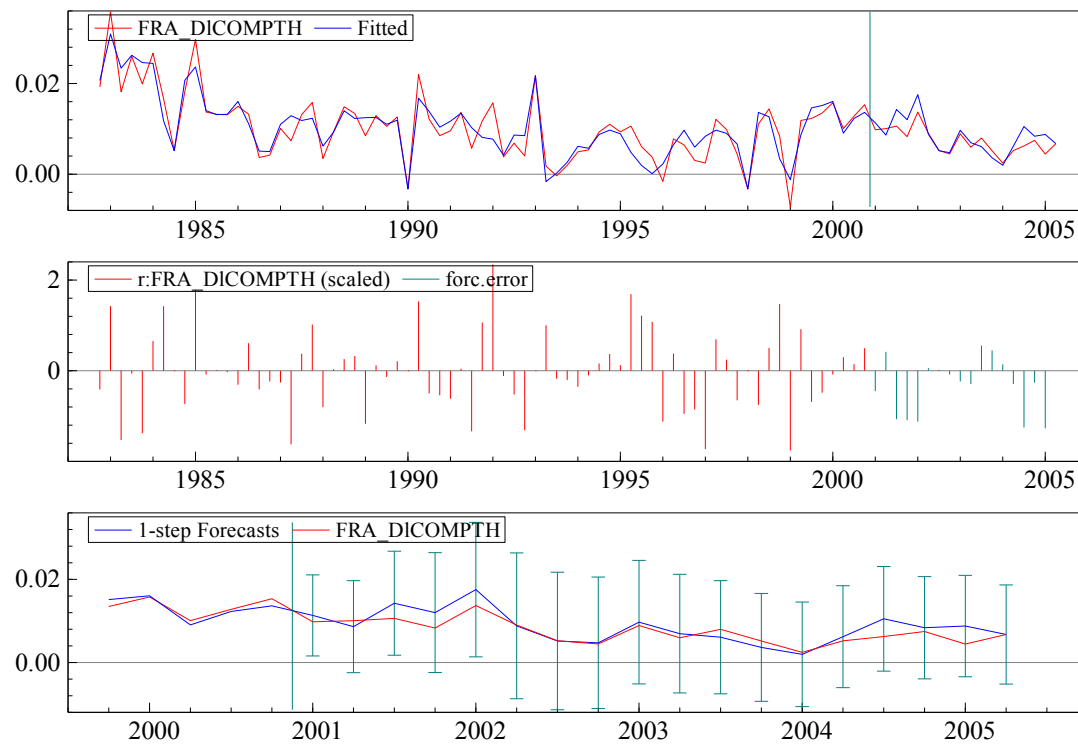


FIGURE 41: FRENCH MODEL, STABILITY TESTS

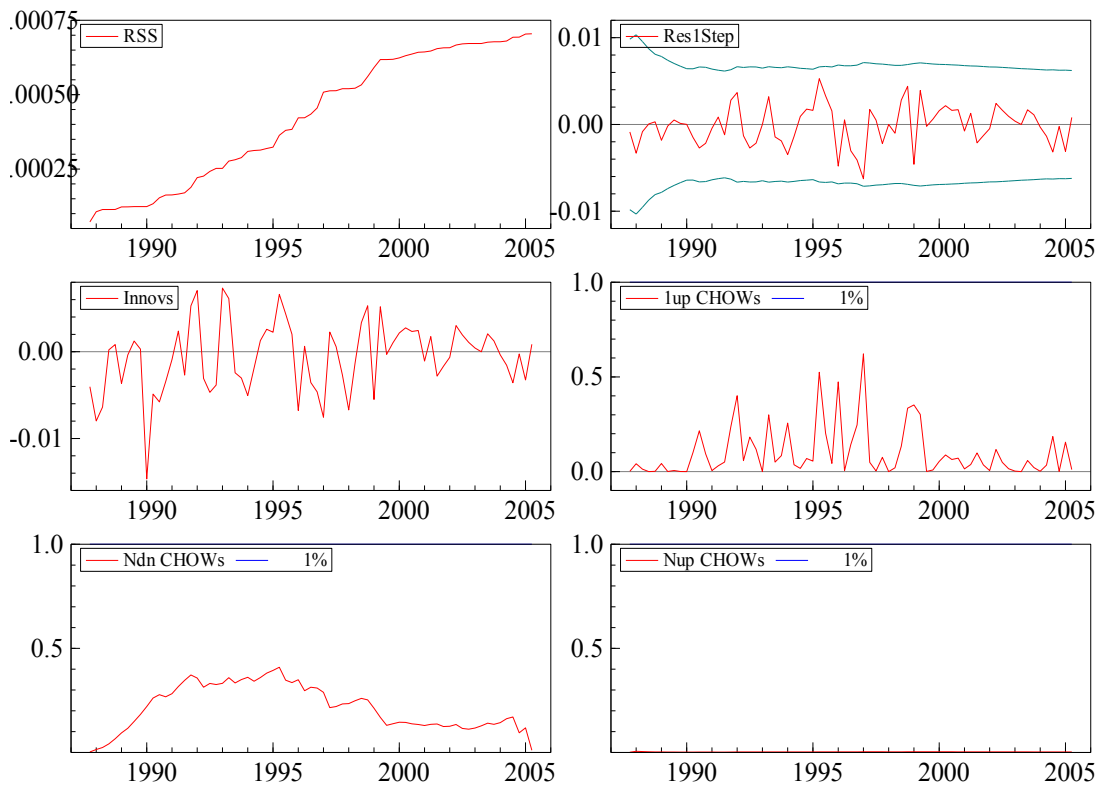


FIGURE 42: FRENCH MODEL, RECURSIVE EqCM COEFFICIENTS

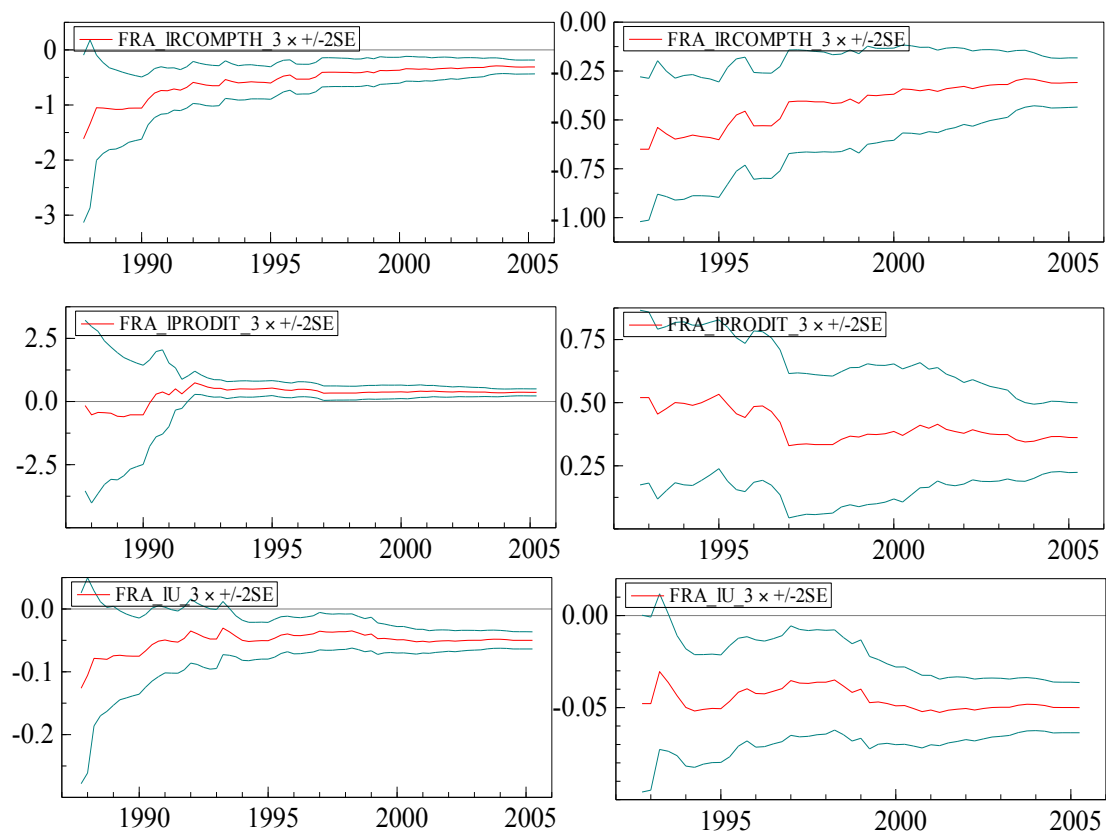


FIGURE 43: RECURSIVE COEFFICIENT, DLPRODI

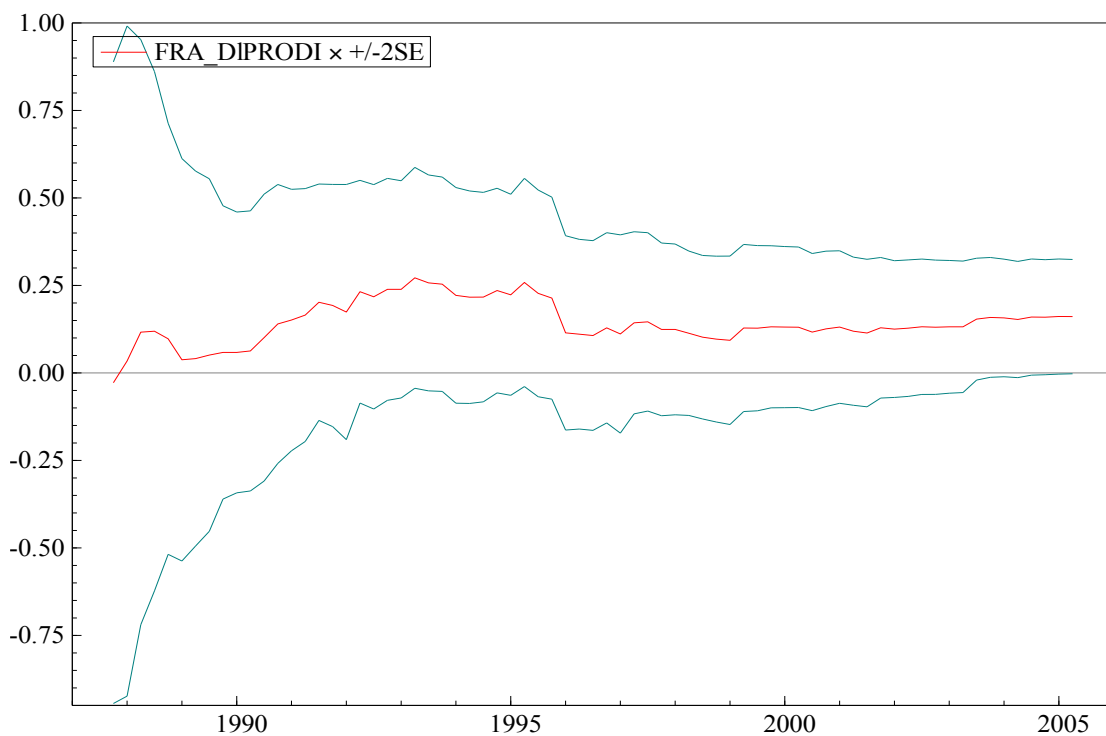
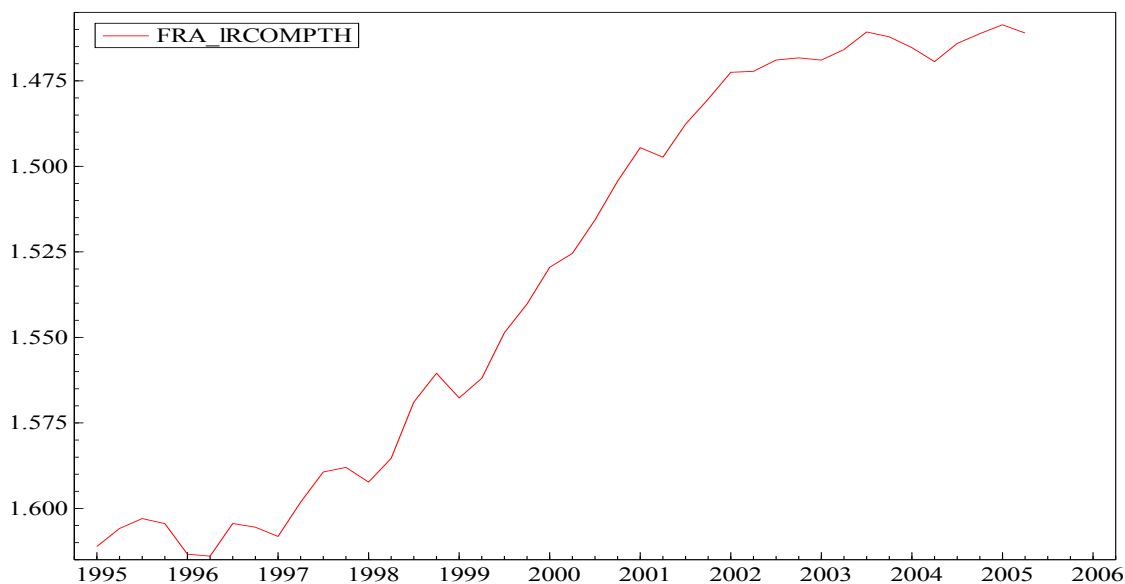


FIGURE 44: LOG HOURLY COMPENSATION OF WORKERS (LRCOMPTH)



4.1.3 Belgium

Modelling BEL_D8lCOMPTH; Estimation Sample: 1985q3-2005q2					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-4.07910	0.8791	-4.64	0.000	0.2432
BEL_D7lCOMPTH_1	0.599318	0.07604	7.88	0.000	0.4811
BEL_EXPINFLPBIS_2	0.947447	0.4020	2.36	0.021	0.0766
BEL_D3lCPI	0.379149	0.1849	2.05	0.044	0.0591
BEL_D3lCPI_5	0.360748	0.1863	1.94	0.057	0.0530
BEL_DlU	-0.178435	0.06422	-2.78	0.007	0.1033
I1995:1	-0.0471252	0.009814	-4.80	0.000	0.2560
I1998	-0.0126998	0.005733	-2.22	0.030	0.0682
BEL_lRCOMPTH_8	-0.395399	0.08654	-4.57	0.000	0.2376
BEL_lPRODIT_8	0.395399*				
BEL_lU_8	-0.0573486	0.01214	-4.72	0.000	0.2499
BEL_lBRR_8	0.390585	0.08739	4.47	0.000	0.2297
BEL_lUDENS_8	0.458514	0.1409	3.26	0.002	0.1366
BEL_lRELPMP_8	-0.0598278	0.02624	-2.28	0.026	0.0720
sigma	0.00897871	RSS		0.00540135045	
R^2	0.964334	F(12,67) =		151 [0.000]**	
log-likelihood	270.61	DW		1.93	
no. of observations	80	no. of parameters		13	
mean(BEL_D8lCOMPTH)	0.0788932	var(BEL_D8lCOMPTH)		0.00189304	

Specification Tests				
AR 1-5 test:	F(5, 62)	=	1.4555	[0.2174]
ARCH 1-4 test:	F(4, 59)	=	1.4549	[0.2275]
Normality test:	Chi ² (2)	=	1.6554	[0.4371]
hetero test:	F(22, 44)	=	1.3550	[0.1922]
RESET test:	F(1, 66)	=	0.019850	[0.8884]

SOURCE: own calculations; starred * coefficients have been restricted.

The Belgian wage equation is notable for requiring the modelling of D8lCOMPTH: the two-year change in hourly compensation of workers. This is rather natural in the Belgian context given that the country has a two-year collective bargaining cycle, with a two-year national inter-sectoral agreement normally concluded at the end of even years (EIRO Annual Reviews; European Commission, 2004). The first years of such cycles are marked by the adoption of (mainly) sectoral collective agreements within the inter-sectoral framework. The second years of the cycle generally see more company agreements.

The dynamics of the equation suggest very strongly that the equilibrium correction term is appropriately located at a lag of 8 quarters, which is consistent with the modelling of D8lCOMPTH. Long-run homogeneity for productivity is accepted, and the speed of adjustment term is found to have a value of around 40%. A strongly significant role is found for unemployment in equilibrium; this contrasts with the opposite finding in Scheremet (2000) but seems more reasonable, notably given the collective bargaining framework in place in Belgium (see below). The institutional variables retained suggest that bargaining power in wage negotiations is significantly influenced by union density and benefit replacement rates. Union density saw a reasonably important increase of around 4% in the first ten years of the period to a figure of 56% (the second highest

density in our sample after Sweden) and fluctuated around that level since then; benefit replacement rates fell discontinuously from 0.5 in 1985 to 0.44 in 2001 before recovering to 0.47 by 2003. A further notable feature of the equilibrium is that consumer prices are the relevant long-run concept in the wage bargain for Belgium, although we do find a significant role for the relative price of imports.

The dynamics of the equation are of particular interest. As can be seen, they include highly significant roles for unemployment and inflation expectations (with a particularly large coefficient that could even be restricted to one), as well as a remarkable influence of compensating claims for inflation over the past two years of the wage bargain. The equation is in fact faithfully reflecting a number of institutional specificities of the Belgian labour market environment.

A significant role for backward-looking elements is only to be expected in Belgium. First, the statutory minimum wage is linked to the “health index” (see below) and adjusted annually by the social partners (EIRO, 2005a; European Commission, 2005b). Second, and more importantly, Belgium is with Luxembourg the only country in Western Europe still operating an automatic indexation system, linking pay and social security benefits to the cost of living (EIRO, 2002b). This is crucial for an understanding of the Belgian situation, yet seems to have generated some confusion in past studies.

Belgium was one of the first countries to introduce a consumer price index (May 1919) and soon afterwards automatic indexation of wages became a feature of the Belgian labour market. In April 1935, the index was placed under the control of the “Index

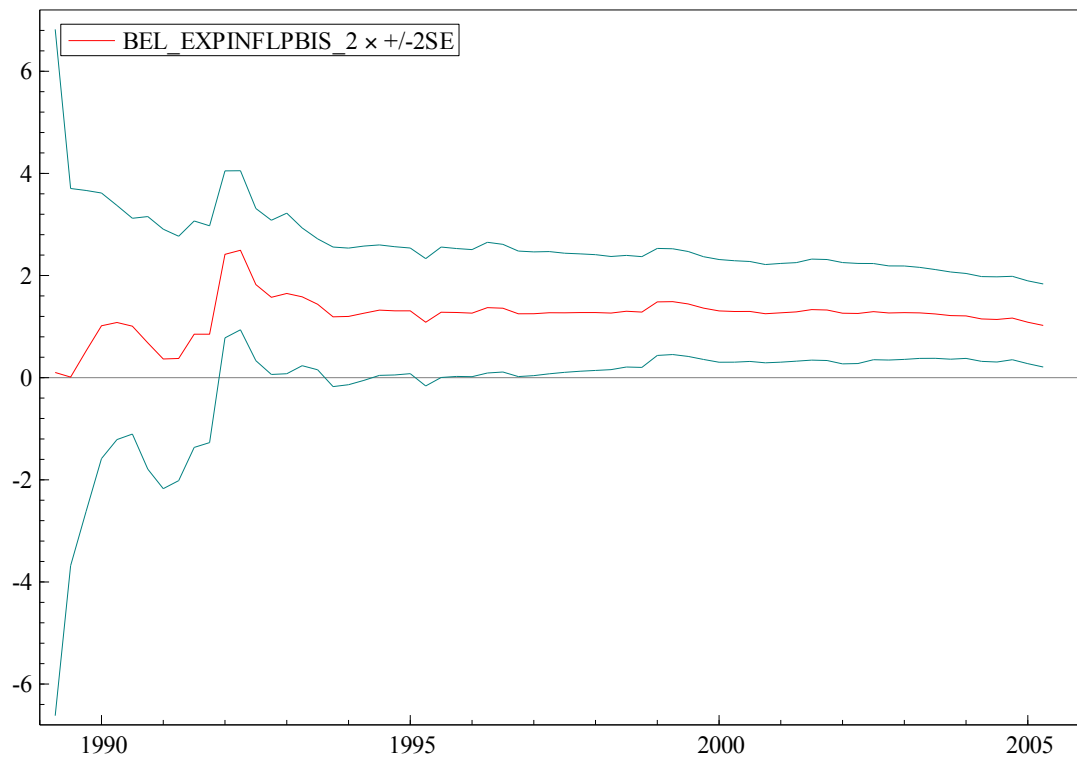
Commission” (*Commission de l’index des prix de détail*), consisting of employer and employee representatives as well as academic personalities, a system still operating to this day. The system was quite stable; indeed, as noted by Vandaele and Boucké (2003), “*From 1951 until March 1982, the wage indexation mechanism functioned without substantial modifications*”. The early 1980s is precisely a period when some authors have found substantial disturbance in wage equations: Artis and Ormerod (1991) found a break in 1981 – attributed to an “EMS effect” – whereas Artis and Ormerod (1996) noted the implementation of some drastic reforms, notably “*the elimination of wage indexation*” and speculated that such reforms may be linked to EMS pressure. Our sample does not allow us to explore these issues. However, it is important to note that following the government’s decision to devalue the currency, wage indexation was merely *suspended* in the first half of 1982 (Vandaele and Boucké, 2003). A flat-rate index was implemented for the other half, and a wage freeze enforced over 1983-1984. Unlimited government control came to an end and automatic full indexation was restored in 1986, roughly coinciding with the start of our sample period. The indexation mechanism was further modified in 1994 with the introduction of the “health index” (excluding products such as tobacco, alcohol and petroleum – which typically experience more inflation). The wage growth in the early 1990s which accompanied economic growth came to a sudden halt as the government introduced a wage freeze from 1995 to 1996 in the context of worsening economic conditions. This change motivates the dummy variable introduced in 1995q1.

Thus, this institutional information helps interpret the backward-looking part of the equation – both in terms of its dynamics found and the finding that it is consumer wages

that are the object of bargaining. We now turn to the institutional motivation of the strongly forward-looking part of the equation.

In 1989, a “Law on the Preservation of the Competitiveness of the Country” was passed (“*Loi de sauvegarde de la compétitivité du pays*”, 6 January 1989). It represented the opening of a new conceptual era for Belgian wage bargaining (Mermet, 2001). The law called for the benchmarking of wage costs and competitiveness with the country’s seven most important trading partners (France, Germany, Italy, Japan, the Netherlands, the United Kingdom and the United States). The law allowed for ex post government intervention after consultation of the Central Economic Council (“*Conseil Central de l’Economie*”, CCE) in the event wage bargaining would compromise competitiveness, and the possibility of overruling collective agreements. The practical implementation of these measures was made possible by an *Arrêté Royal* dated 24 December 1993. It is reasonable to suppose that the threat of government intervention changed the bargaining context and provided an additional incentive for unions and employer organisations to avoid generating excessive inflation. It is precisely in this period that inflation expectations become significant in our equation, as shown in Figure 45.

FIGURE 45: BELGIUM, RECURSIVE COEFFICIENT, INFLATION EXPECTATIONS



Subsequent policy developments moved further in this direction. July 1996 saw the introduction of a law on the “promotion of employment and the *preventive* safeguard of competitiveness”⁷⁸, mandating the CCE to produce a yearly technical report determining the maximum permitted nominal wage cost increase margin. This is based on anticipated wage cost movements in three reference countries – Germany, France, Netherlands – and on wage cost movements in Belgium over the previous two years. The first year of operation of this system (1997) saw a breakdown of central bargaining in Belgium as the social partners were unable to agree on the pay increases proposed by the CCE for the 1997-98 cycle, and the level of increase was imposed by the government. This justifies the 1998 dummy present in the equation.

⁷⁸ Emphasis added; Law dated 26 July 1996.

If there is a staggering structure to bargaining, then it is subsumed in the second lag of inflation expectations rather than appearing explicitly. The lag may also suggest discrete information collection and lengthy bargaining. The magnitude of this second lag is especially notable, and rather large relative to what the dynamics in hourly compensation would suggest, although its standard error is also large. It may be somewhat upwards biased by the omission of positively correlated information from neighbouring countries relevant to the setting of the CCE wage norm – although the coefficient is very stable over time. Rather large effects have tended to be found in the literature (e.g. Anderton and Barrell, 1995; Artis and Ormerod, 1996) though interestingly Barrell and Dury (2001) do not confirm this.

The model is congruent as evidenced by the specification tests and has an excellent fit (sigma of 0.89%). Figure 46 provides some graphical evidence. The complete lack of autocorrelation in the error structure is remarkable, given the fact that we are modelling an 8-quarter difference. The stability tests in Figure 48 confirm that the model is stable. Figure 47 shows that the forecasts from 1999q1 are fully satisfactory for the first two years, but then significantly over-predict actual developments, especially for 2001 and 2002. The explanation of this fact reveals a number of interesting issues.

The period of 2001-02 is covered by the inter-sectoral agreement signed in December 2000. This agreement covered two main themes (EIRO, 2001a): (1) The renewal of the 'pay norm' (calibrated in view of the labour costs in neighbouring countries) that guides subsequent sector and company-level bargaining; and (2), the balance between work and family life.

The fixing of a maximum margin for pay rises in the inter-sectoral agreement proved to be a highly controversial issue, with trade unions arguing that periods of recession had been left behind and employers highlighting the need to remain competitive in an open economy. In the end, wage moderation prevailed in both the setting of the pay norm (which was merely indicative in this bargaining round) and in the subsequent sectoral negotiations.

Regarding the second issue, the inter-sectoral agreement reduced the maximum normal working week to 38 hours from the previous level of 39, starting in January 2003. However, data from the OECD's Productivity Database show that average hours worked in the economy increased in both 2001 and 2002, against an overall decreasing trend before that. Working time fell again in 2003. The temporary increase in working time against a background of wage moderation may explain part of the deviation between our forecasts and the actual outcomes for hourly compensation.

However, this can hardly be a full explanation. 2001 appears to be a year of change for several coefficients, although without there being a break in the series to properly speak of. This is notably the case for the union density and benefit replacement rate coefficients, which become significant in this year. The most remarkable change observed, however, concerns the unemployment coefficient which decreases in magnitude that year (c.f. Figure 49). This outcome may not seem encouraging at first sight (it acts in the direction opposite to the "EMU effect" being sought theoretically). However, it is important to understand what may have originated it. The quarters

between 1998q1 and 2002q1 saw a sharp decrease in the unemployment rate in Belgium (from 9.43% to 6.55%, almost three percentage points). Most of this fall (around 2.5%) took place in 1999 and 2000. This could have been a significant factor in promoting a relaxation of the policy of wage moderation. We know from the outcomes of hourly compensation that wage moderation did prevail. This must mean that, in a sense, wage moderation became more independent of the cyclical state of the economy, and it is this observation that the change in the coefficient may be reflecting.

Thus, although we find no evidence of an “EMU effect” having increased wage flexibility in Belgium, there is evidence of a willingness to reach moderate wage bargaining outcomes. The underlying question, of course, remains the extent to which EMU might be responsible for such a development. An argument against this view is that the Belgian government itself *crafted* wage moderation into the system, with the implementation of the 1989 law and the willingness to impose wage freezes (1995-1996) or particular wage increases (1997-1998) when the social partners proved unable to provide satisfactory solutions; indeed, the importance of expectations in the equation seems to date from the early 1990s, as noted early. However, there is no doubt that many of these reforms were taken *in view of monetary union*. The wage freeze of 1995-1996 was arguably partly motivated by the *Maastricht Criteria* (Vandaele and Boucké, 2003). Furthermore, an important revealed intention behind the design of the 1996 Law was to achieve stabilisation in the context of EMU (e.g. CCE, 1998). As a matter of political economy, EMU may well have made wage moderation easier to justify and enforce in a small open economy such as Belgium.

FIGURE 46: BELGIAN MODEL, GRAPHICAL ANALYSIS, 1985Q3-2005Q2

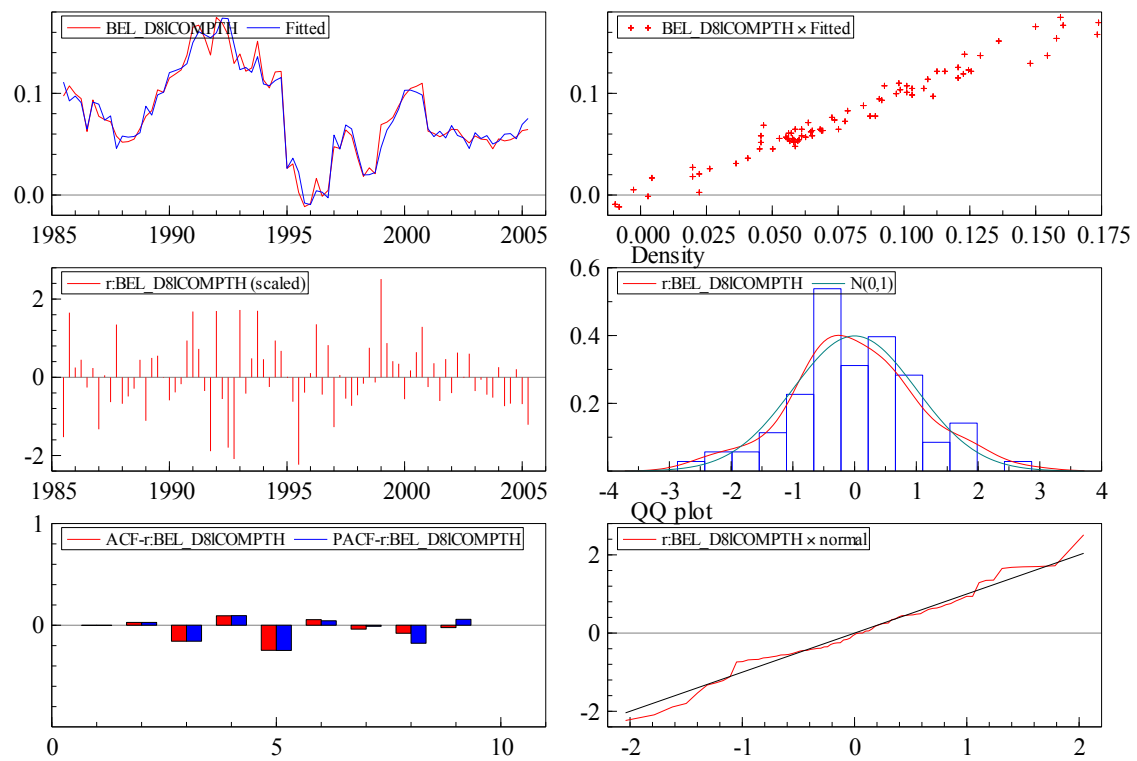


FIGURE 47: BELGIAN MODEL, FORECASTS FROM 1999:1

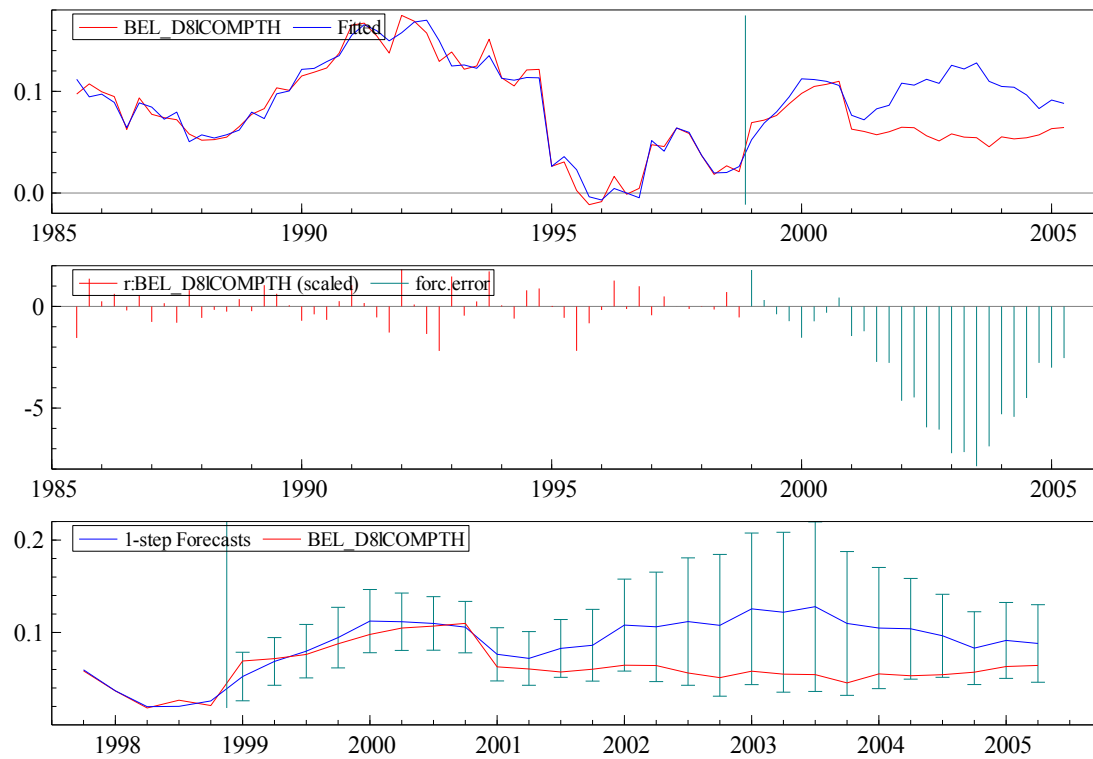


FIGURE 48: BELGIAN MODEL, STABILITY TESTS

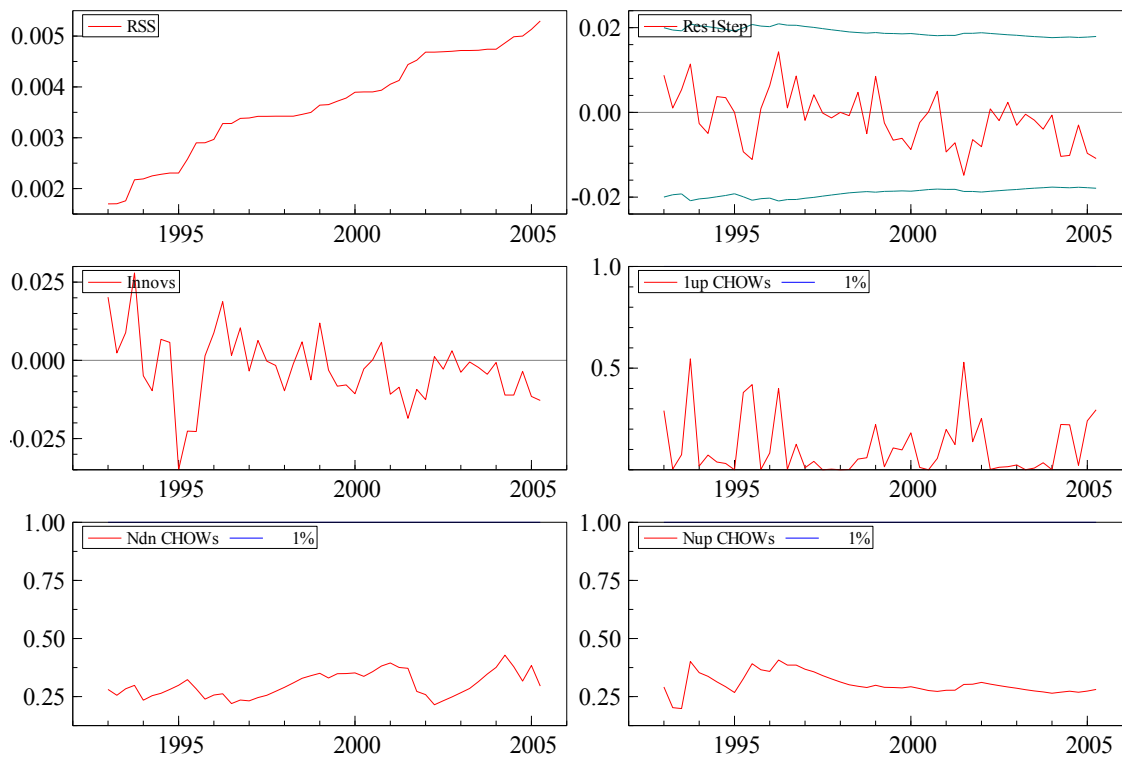
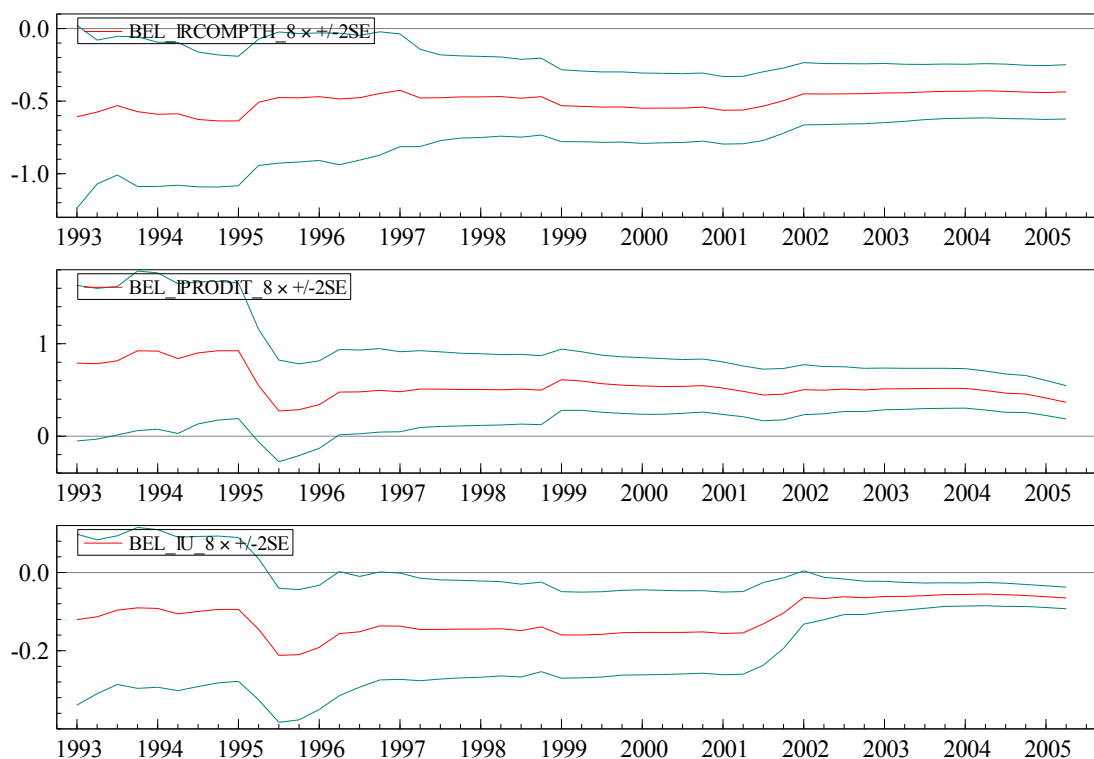


FIGURE 49: BELGIAN MODEL, RECURSIVE EqCM COEFFICIENTS



4.1.4 Spain

Modelling ESP_D4lCOMP_4; Estimation Sample: 1987q2-2005q2					
	Coefficient	Std.Error	t-value	t-prob	Part.R ²
Constant	-4.50576	0.4930	-9.14	0.000	0.5820
ESP_DlCOMP_1	0.173255	0.05577	3.11	0.003	0.1386
ESP_DlCOMP_4	0.158889	0.06594	2.41	0.019	0.0882
ESP_EXPINFLPBIS_3	0.498795	0.1195	4.17	0.000	0.2251
ESP_D3lCPI_1	0.611006	0.1151	5.31	0.000	0.3195
I2005h1	0.0244255	0.005946	4.11	0.000	0.2195
I1992:1	0.0399082	0.007932	5.03	0.000	0.2967
I1997:1	-0.0219670	0.007747	-2.84	0.006	0.1182
ESP_lRCOMP_4	-0.640408	0.07191	-8.91	0.000	0.5693
ESP_lPRODIT_4	0.640408*				
ESP_lU_4	-0.0220836	0.008392	-2.63	0.011	0.1035
ESP_lEPL_4	0.136718	0.01896	7.21	0.000	0.4643
ESP_lUDENS_4	0.0309446	0.01483	2.09	0.041	0.0676
ESP_lRELPMP_4	-0.103802	0.02773	-3.74	0.000	0.1893
sigma	0.00743714	RSS		0.0033186613	
R ²	0.936848	F(12, 60) =	74.17	[0.000]**	
log-likelihood	261.368	DW		2.25	
no. of observations	73	no. of parameters		13	
mean(ESP_D4lCOMP_4)	0.05799	var(ESP_D4lCOMP_4)		0.000719871	

Specification Tests				
AR 1-5 test:	F(5, 55)	=	1.8904	[0.1109]
ARCH 1-4 test:	F(4, 52)	=	0.27360	[0.8937]
Normality test:	Chi ² (2)	=	0.78535	[0.6752]
hetero test:	F(21, 38)	=	1.1994	[0.3052]
RESET test:	F(1, 59)	=	0.17483	[0.6774]

SOURCE: own calculations; starred * coefficients have been restricted.

The Spanish specification is concerned with the modelling of the yearly change in hourly compensation, D4LCOMP_{TH}. We find from the dynamics in the equation that the equilibrium correction term is optimally modelled with a lag of one year, and displays a speed of adjustment coefficient of 64%. This is a reasonably high value for this coefficient and seems consistent with data from the European Commission (2004), which estimate the usual duration of collective agreements in Spain to be between two and three years.

The reason why the rhythm of bargaining might be less transparently reflected for Spain than in the Italian or Belgian equations, for instance, might be the highly fragmented nature of Spanish wage bargaining. Data from the European Commission (2004) shed light on the importance of different bargaining levels – defined as “national”, “sector” and “company” bargaining. In Spain, the national sector is catalogued as “existing” whereas both other levels are ranked as “important, but not dominant”. Spain is the only country in our sample with such a situation (all others display a “dominant” level of bargaining, even when others coexist). In fact, a finer judgement would describe the key level of bargaining in the country as *provincial*, but existing alongside other varieties (e.g. national and company) (EIRO 1999b, 2001b).

We are able to impose long-run homogeneity for the influence of trend productivity on real hourly compensation. Unemployment is found to be a significant influence in dampening compensation increases in equilibrium, as are relative import prices. On the other hand, two institutional variables are found to act as significant motives for real increases in hourly compensation: employment protection legislation and union density. Employment protection legislation has been reduced in Spain over the sample period overall, with sharp decreases in the 1990s. In 1994, for instance, procedural requirements for economic dismissals were reduced and notice periods shortened; temporary work agencies were also allowed. In 1997 the maximum compensation for unfair dismissal was reduced and the relevant definitions were adjusted. On the other hand, the indicator shows a slight increase more recently linked to the adoption of Law 12/2001, which tightened the legislation applicable to fixed-term contracts somewhat (OECD, 2004b). Union density increased sharply until 1993 (doubling to 18% from a starting value of around 9% in 1987q2) before gradually coming down to 14%.

Significant dynamics in the dependent variable suggest the importance of earlier settlements in bargaining. Furthermore, significant roles are also attributed to both inflation expectations and compensation for inflation experienced in the recent past. Inflation expectations appear lagged three periods, which may be reflective of the time spent on bargaining negotiations and/or of an implicit structure of staggered contracting. The latter would be expected to prove especially difficult to detect explicitly in the Spanish case, given the highly fragmented structure of Spanish bargaining already noted. The important backward-looking dynamics for inflation were only to be expected, and may be reflecting the commonality of wage indexation clauses in collective agreements

and the fact that inflation is used as one of the criteria to adjust statutory or collectively agreed minimum wages (European Commission, 2005b). Remarkably, short-run dynamics in unemployment could not be evidenced in Spain.

The literature having considered Spanish wage equations in the past has generated a number of findings regarding their stability. Scheremet (2000) found no signs of breaks. However, Anderton and Barrell (1995) found evidence of structural change in 1986q1, which they attribute to the growing significance of fixed-term contracts in the labour market⁷⁹; Lauer (1999) found evidence of a break in 1988 (unexplained), whilst Barrell and Dury (2001) found one for 1987, noting that it followed Spain's entry into the ERM. Another plausible candidate for breaks around those dates would of course be Spain's entry into the EEC (1 January 1986; see Curto Millet, 2004). Unfortunately, our sample prevents us from analysing such questions here.

As can be seen, we found it necessary to introduce a dummy to control for the significant increases in hourly compensation that took place in the first two quarters of 2005. It is difficult to pinpoint *exactly* the source of this behaviour in the series, as the relevant data have not yet been made available by the OECD. Two hypotheses stand out, nonetheless. First, there was an extraordinary increase of 6.6% for the minimum wage (*Salario Mínimo Interprofesional, SMI*) in July 2004, bringing the annual variation in 2004 to 8.6% - a change not matched since 1983, in a very different inflationary environment. The SMI concerns about a million workers directly (out of 18-19 million in employment). Furthermore, on 27 December 2004 the government, employer

⁷⁹ A substantial relaxation on restrictions for these types of contracts took place in 1984 (OECD, 2004b).

organisations and trade unions reached an agreement for the SMI to be raised by 4.5% in 2005 (EIRO, 2005b). Although some compensating social security contributions were planned to mitigate the effect on employers' labour costs, it seems likely that these have nonetheless increased. Furthermore, these developments may have affected the bargaining of non-SMI wages in an upward fashion. Second, the government decided to increase minimum pensions by between 4.5% and 8.07% in January 2005 (EIRO, 2005c), which may have further increased employers' social contributions.

The dummy introduced in the first quarter of 1997 reflects important tensions in Spanish wage bargaining. As can be seen in Figure 50, wage growth falls importantly that quarter. This coincides with a wage freeze imposed by the newly elected *Partido Popular* on civil servants and workers in public administration (*funcionarios*). Indeed, the government decided not to implement a three-year wage agreement (1995-7) for these workers concluded by the previous PSOE (Socialist) government and the trade unions (CC.OO, UGT and CSI-CSIF, among others). This deal involved annual pay increases equivalent to the increase in the retail price index (EIRO, 1997b; 2001e). The government's actions led to a series of strikes in the sector, starting in late 1996.

The specification tests show the model to be congruent, with an excellent fit (sigma of 0.74%). The autocorrelation result is especially encouraging given the smoothness of the series being modelled (Figure 50). The model appears to be stable overall, and in particular in the post-EMU period. The forecasts from the point of introduction of EMU (Figure 51) are broadly satisfactory, only becoming questionable in late 2004 and 2005 given the absence of the appropriate dummy. The battery of statistical tests in Figure 52

fails to detect any signs of instability in the EMU period. The recursive estimates of the coefficients otherwise appear to be stable. In particular, this behaviour is displayed by the key EqCM coefficients in Figure 53.

Thus, overall, there seems to be no indication of an “EMU effect” with respect to Spain. As in the case of Italy, this may be seen as particularly damaging to the case for “EMU effects” in flexibility given the country’s history of competitive devaluations (Lauer, 1999).

FIGURE 50: SPANISH MODEL, GRAPHICAL ANALYSIS, 1987Q2-2005Q2

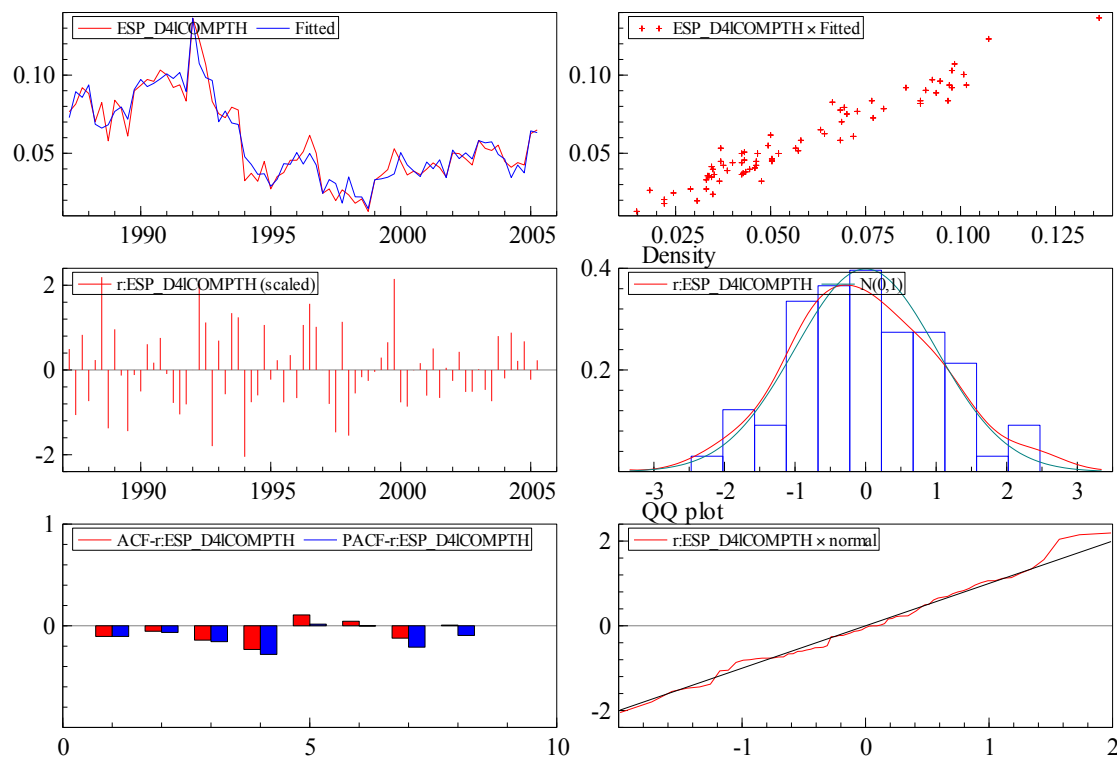


FIGURE 51: SPANISH MODEL, FORECASTS FROM 1999:1

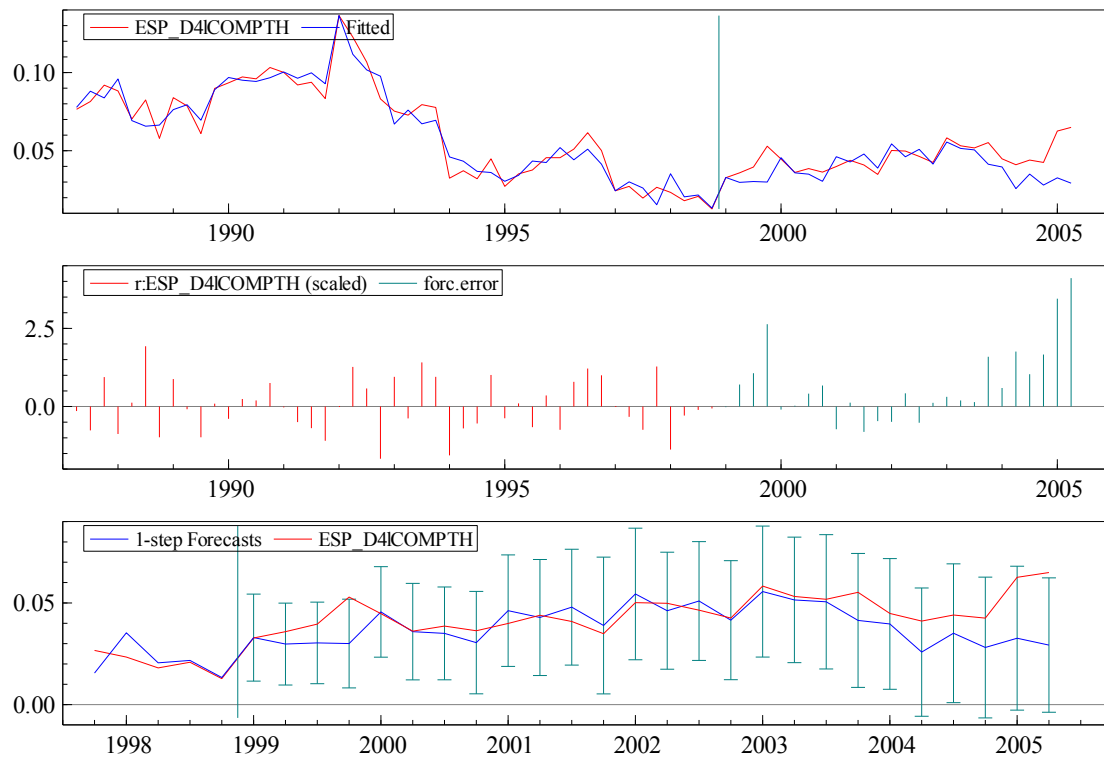


FIGURE 52: SPANISH MODEL, STABILITY TESTS

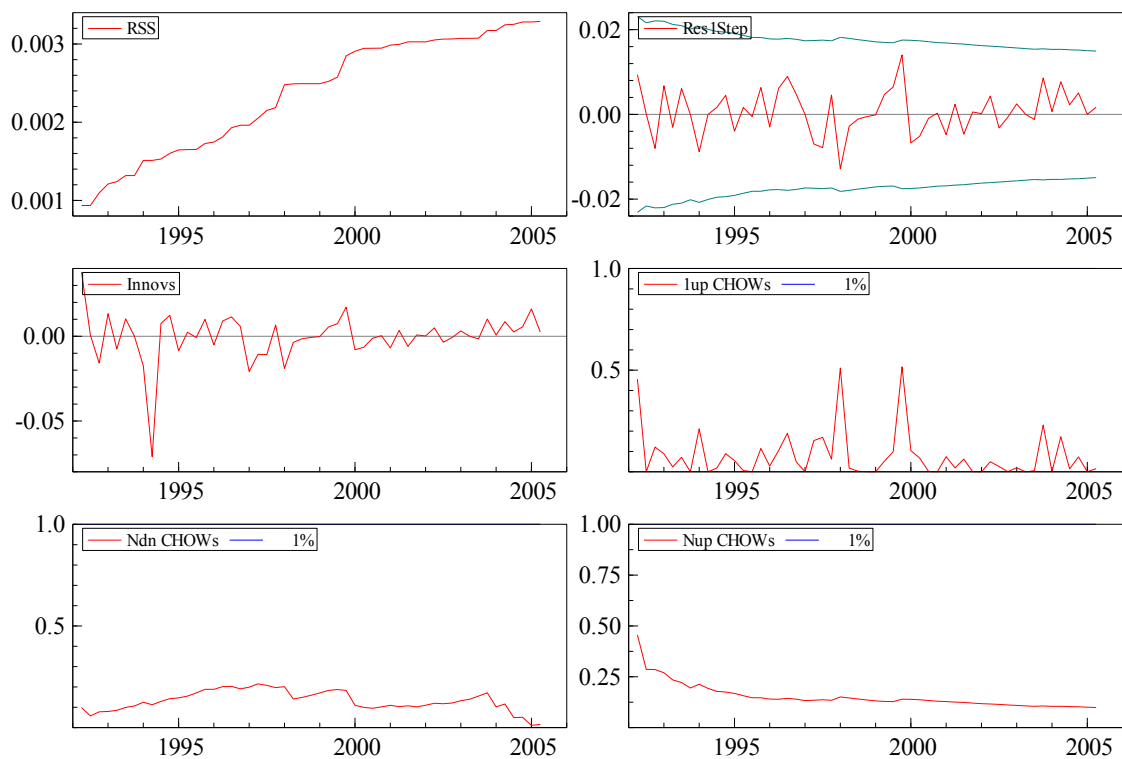
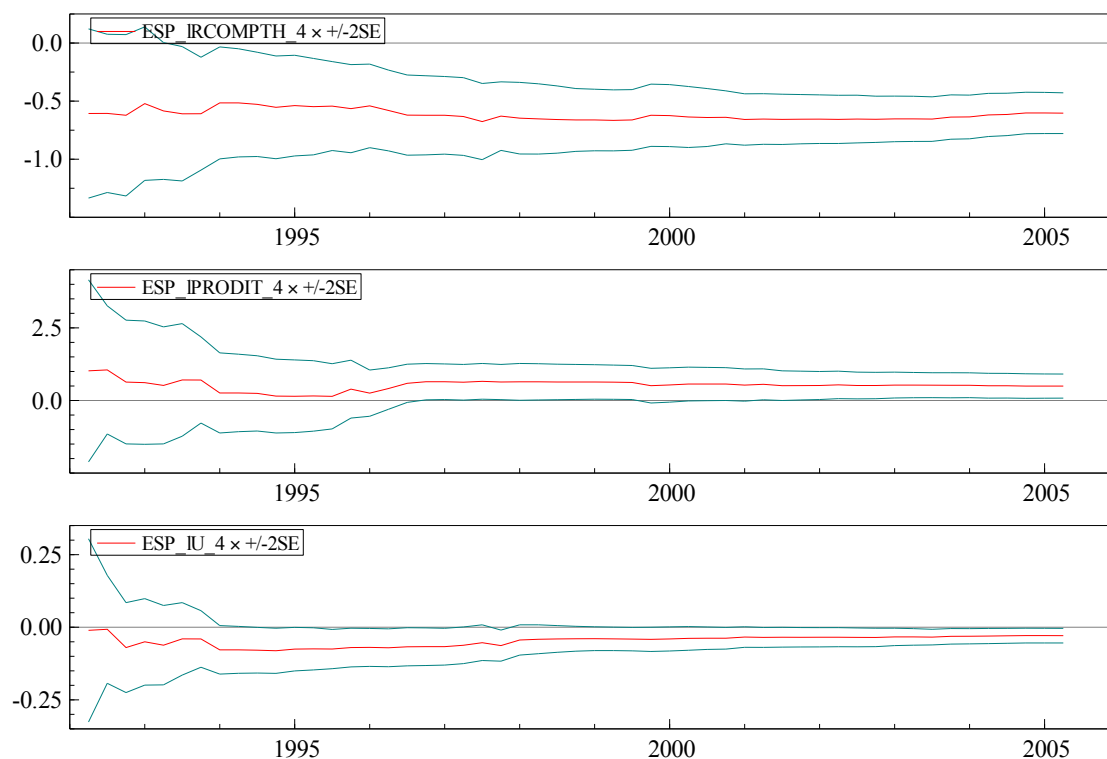


FIGURE 53: SPANISH MODEL, RECURSIVE EqCM COEFFICIENTS



4.1.5 Germany

Modelling DEU_D21COMP_2; Estimation Sample: 1991q3-2005q2					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-11.5290	2.967	-3.89	0.000	0.2554
DEU_D1COMP_1	0.447762	0.09998	4.48	0.000	0.3131
DEU_EXPINFLPBIS_4	0.309560	0.1111	2.79	0.008	0.1500
DEU_D21CPI	0.546021	0.1357	4.02	0.000	0.2689
DEU_D1U_3	-0.0210874	0.01127	-1.87	0.068	0.0737
I1994:2	-0.0140562	0.004556	-3.09	0.004	0.1778
DEU_1RCOMP_2	-0.559194	0.09855	-5.67	0.000	0.4226
DEU_1PRODIT_2	0.273009	0.08714	3.13	0.003	0.1824
DEU_1ITAX_2	-0.995611	0.1800	-5.53	0.000	0.4102
DEU_1CTAX_2	-0.387168	0.2255	-1.72	0.093	0.0628
DEU_1BRR_2	0.194528	0.05929	3.28	0.002	0.1965
DEU_1ACOV_2	2.10984	0.6381	3.31	0.002	0.1990
sigma	0.00423326	RSS	0.000788500237		
R^2	0.919762	F(11,44) =	45.85	[0.000]**	
log-likelihood	233.32	DW	2.22		
no. of observations	56	no. of parameters	12		
mean(DEU_D21COMP_2)	0.0143062	var(DEU_D21COMP_2)	0.000175483		

Specification Tests				
AR 1-4 test:	$F(4, 40)$	=	0.37551	[0.8247]
ARCH 1-4 test:	$F(4, 36)$	=	1.1189	[0.3628]
Normality test:	$\chi^2(2)$	=	2.9091	[0.2335]
hetero test:	$F(21, 22)$	=	0.17760	[0.9999]
RESET test:	$F(1, 43)$	=	1.2207	[0.2754]

SOURCE: own calculations; starred * coefficients have been restricted.

We model the half-yearly change in the hourly compensation of workers for Germany over the period 1991q3-2005q2. The equilibrium correction term is accordingly lagged by two quarters, with a speed of adjustment coefficient of 56%. The size of this coefficient appears consistent with a usual duration of collective agreements in Germany estimated at 1-2 years by the European Commission (2004).

The influence of trend productivity on the equilibrium is highly significant. However, the coefficient is only half that on lagged real hourly compensation, and a homogeneity restriction cannot be imposed (more on this below). Remarkably, both direct and indirect taxation seem to matter in the long-run solution, although the latter is only significant at the 10% level and the former appears to be somewhat large in magnitude. A longer sample would clearly be beneficial to estimate their effects more precisely. In any event, it seems possible on the present evidence to make the claim that bargaining in Germany largely concerns post-tax real producer wages. The importance of these variables, and especially of direct taxation, is not surprising in the German context. Indeed, it is well-known that in order to meet the costs of re-unification, social security contribution rates were increased as well as the burden of direct taxation (e.g. the “solidarity surcharge” in West Germany) (DG ECFIN, 2002). Unemployment is completely insignificant in the

equilibrium solution, possibly due to the very limited variation it displays over our sample period.

Two institutional variables are retained by our selection process. The first is adjusted bargaining coverage, which experienced an increase in Germany in the early 1990s. Its coefficient is rather large, although the variable is also estimated imprecisely. A highly significant effect is also found for the benefit replacement rate variable, which fell in the early 1990s, reaching its lowest value in 1997, and rose again subsequently. A dummy variable was inserted to take account of an outlier in the second quarter of 1994.

Our data on inflation expectations switch from concerning West Germany to the whole of Germany in January 1991⁸⁰. This means that, given the four-quarter lag on the inflation expectations term for the specification above, two points of West German expectations data are used above. We checked the equation for robustness by limiting the sample to 1992q1 onwards and obtained the same results.

The dynamics include a role for unemployment, although the coefficient is only significant at the 10% level. On the other hand, highly significant roles are found for both inflation expectations and compensating claims for past inflation. These findings deviate substantially from some made previously in the literature. Indeed, several studies have highlighted the *insignificance* of the German inflation expectations variable (Artis and Ormerod, 1996; Anderton et al., 1992; Anderton and Barrell, 1995; Barrell and Dury, 2001). This was generally attributed to the low-inflation history of Germany, as well as

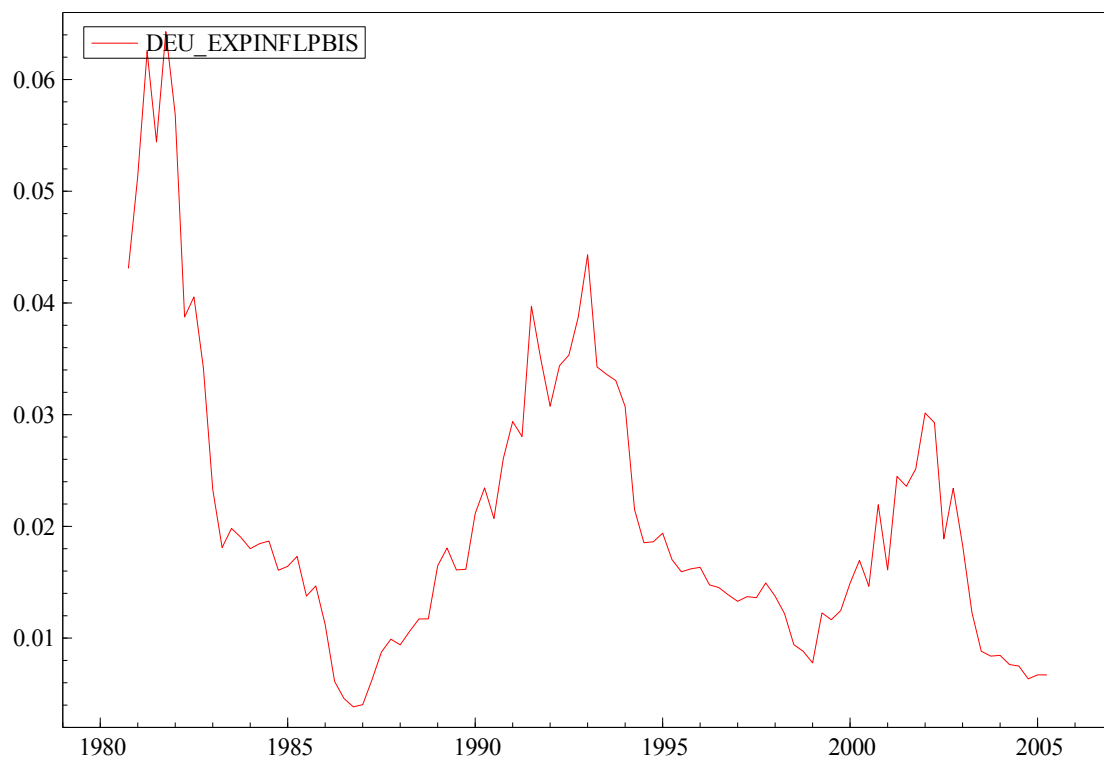
⁸⁰ We thank Rolf Bürkl for clarifying this issue in a private communication (21/03/2005).

the low-variability of inflation; it was then argued that such low-variability would also be a characteristic of expectations, so that their effects would be absorbed in the constant term. Three possible explanations for the diverging findings reached here come to mind.

1. The EMU years add data points with higher variability for inflation expectations than that over the previous 20 years, allowing identification.

This hypothesis is contradicted by the data used in this study, as can be seen from a consideration of Figure 54.

FIGURE 54: INFLATION EXPECTATIONS IN GERMANY, 1980Q1-2005Q2



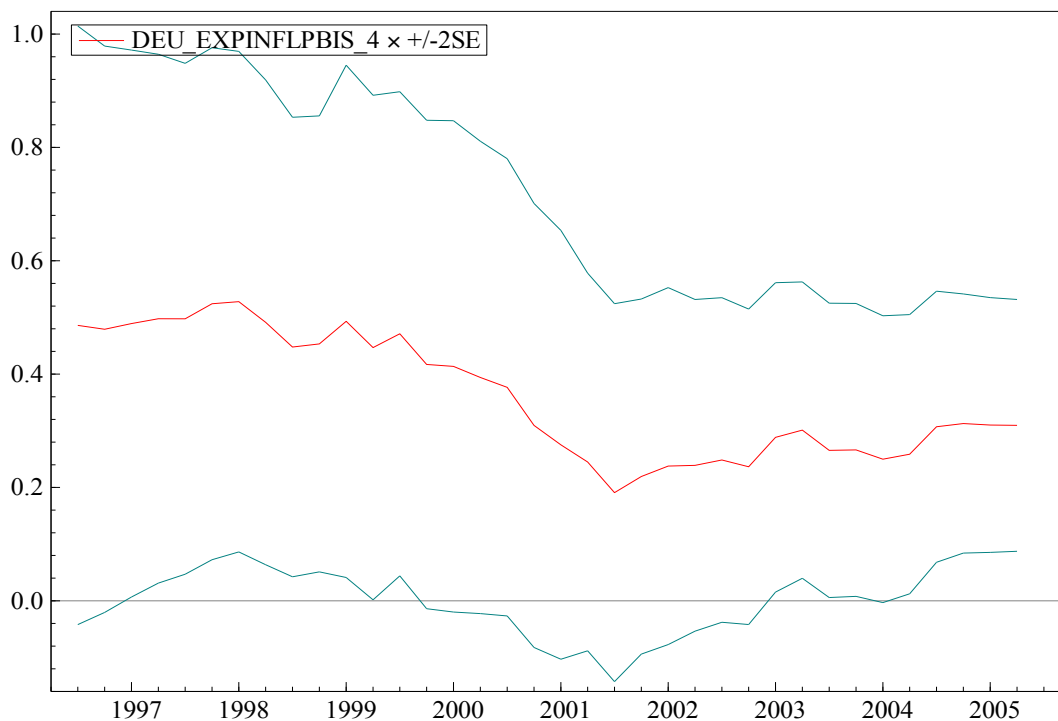
2. The divergence arises from the fact that the present study uses an *actual* measure of expectations as an input into the wage equation.

This is a plausible contributory explanation to the finding here. It must be noted however that Curto Millet (2004) using the same methodological approach as was deployed here found expectations to be insignificant. Nevertheless, it now appears that this earlier finding was probably due to the very limited sample available for consideration then (1992q1-2002q4) as will be suggested in the following point.

3. The importance of expectations is a recent phenomenon in German wage bargaining, possibly a post-EMU one.

The time sample available here is obviously very limited; in particular, no judgement can be expressed over wage bargaining in the 1980s. However, the data available suggest that, although expectations only became durably *significant* in 2003, their coefficient was already important earlier in the period, suggesting that the phenomenon is not new (Figure 56).

FIGURE 55: RECURSIVE COEFFICIENT, GERMAN INFLATION EXPECTATIONS



Thus, the temptation of drawing more general policy conclusions from the apparent novelty of the present finding of significance relative to previous research should be resisted.

The presence of an expectations term lagged four quarters does not have a unique and obvious explanation. It may reflect in part negotiation lags and discrete information collection, as in Mankiw et al. (2003). It is also possible that other expectations terms with less distant lags had to be dropped due to collinearity with the past inflation terms.

Another important difference with some previous studies in the literature concerns the lagged inflation terms present in the wage equation. Anderton et al. (1992) and Anderton and Barrell (1995) surprisingly find an absence of lagged inflation terms in their German

wage equations; Anderton et al. (1992) attribute this to the “*low degree of persistence of German inflation*”. In contrast, the equation proposed here does evidence a significant dynamic response of wages to past inflation. This is more in tune with institutional information. For instance, it is well-known that in Germany inflation is one of the criteria used to adjust statutory or collectively agreed minimum wages (European Commission, 2005b). Moreover, German wage bargaining has traditionally “*always oriented itself to the "margin of distribution" (Verteilungsspielraum)*” (EIRO, 1997a), i.e. to the sum of the increases in inflation and productivity.

The equation is clearly congruent according to the specification tests, with an excellent fit (sigma of 0.42%) (Figure 56). However, the forecasts from the first quarter of 1999 yield a very substantial overprediction (Figure 57) whereas those from the first quarter of 2000 onwards are on track (Figure 58). The adequacy of the stability tests in Figure 59 points to a reasonably smooth change in coefficients as the cause of this. Indeed, only two are found to change importantly in 1999. First, there is a decrease in the benefit replacement rate coefficient from 0.36 to 0.25. This probably reflects German wage moderation in 1999 in the face of an increase in the underlying variable that year. Second, and most importantly, the productivity trend’s influence drops from 0.47 to 0.35 in 1999; by 2005, the coefficient had further decreased to 0.27 (see Figure 60). This follows a number of years of very moderate wage increases. For instance, in 1998 wages again rose “*above the rate of inflation but below the increase in labour productivity*” (EIRO, 1998a). This was reflected by a particularly sharp drop in the labour share⁸¹.

⁸¹ The variable is computed as the compensation of workers (employed and self-employed) as a proportion of gross value added at basic prices (excluding taxes less subsidies on products). We use quarterly Eurostat data for the latter series. See Batini et al. (2000).

Indeed, Figure 61 shows that our sample period is characterised by an important decrease in this variable, which also displays a downward trend for most of the period. As such, the weakness of the productivity coefficient is therefore not surprising.

Overall, evidence supportive of an “EMU effect” in terms of flexibility seems rather limited, although there is a noticeable increase in the magnitude of the speed of adjustment coefficient in 1998. Instead, as was the case in Belgium, we find the period surrounding the advent of EMU to be associated with important wage moderation. Unlike in the case of Belgium, however, the cause here seems to be more related to domestic conditions, namely weak internal demand and persistent unemployment. Indeed, these seem to be the principal concerns which motivated unions to agree to the policy at the time (EIRO, 1998b). What is more, the observed wage moderation predates the introduction of EMU by many years, having characterised practically every year in the 1990s.

FIGURE 56: GERMAN MODEL, GRAPHICAL ANALYSIS, 1991Q3-2005Q2

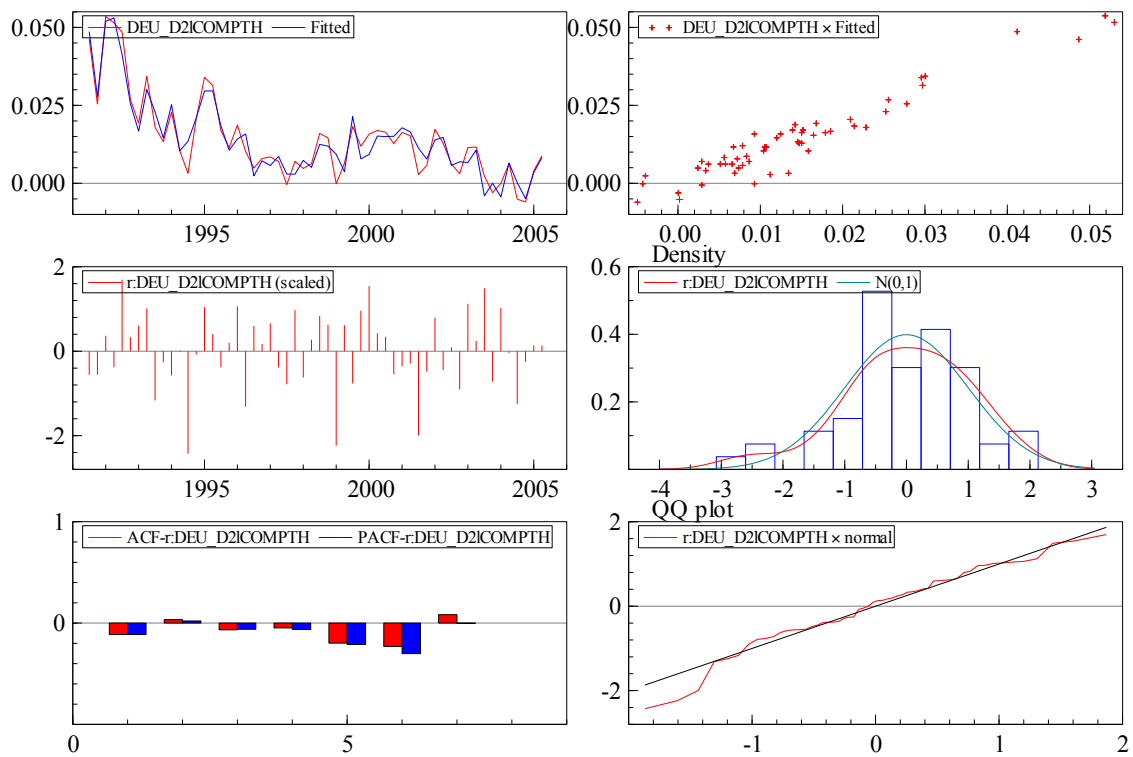


FIGURE 57: GERMAN MODEL, FORECASTS FROM 1999:1

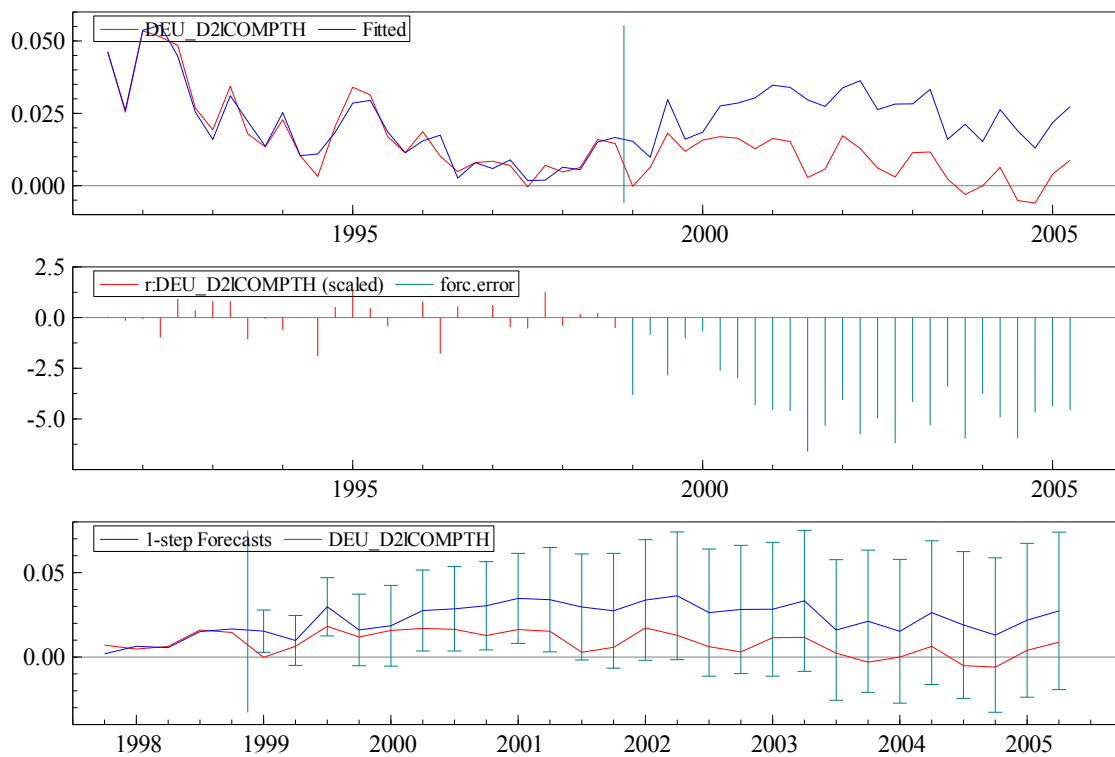


FIGURE 58: GERMAN MODEL, FORECASTS FROM 2000:1

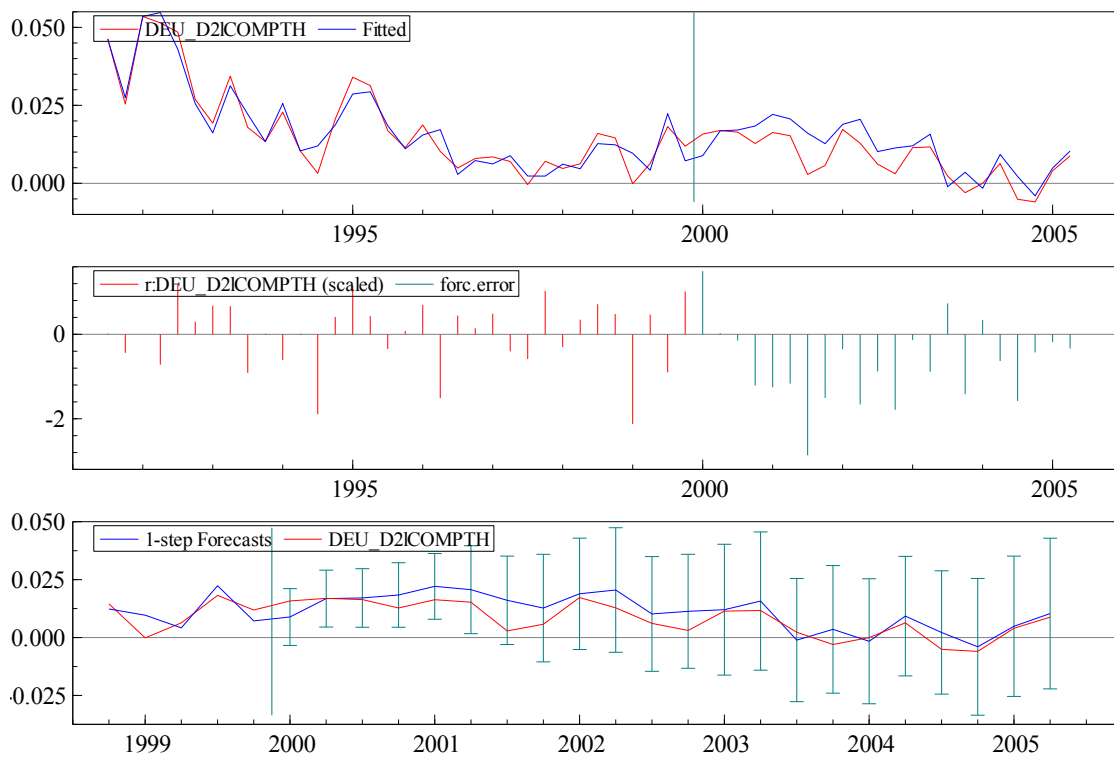


FIGURE 59: GERMAN MODEL, STABILITY TESTS

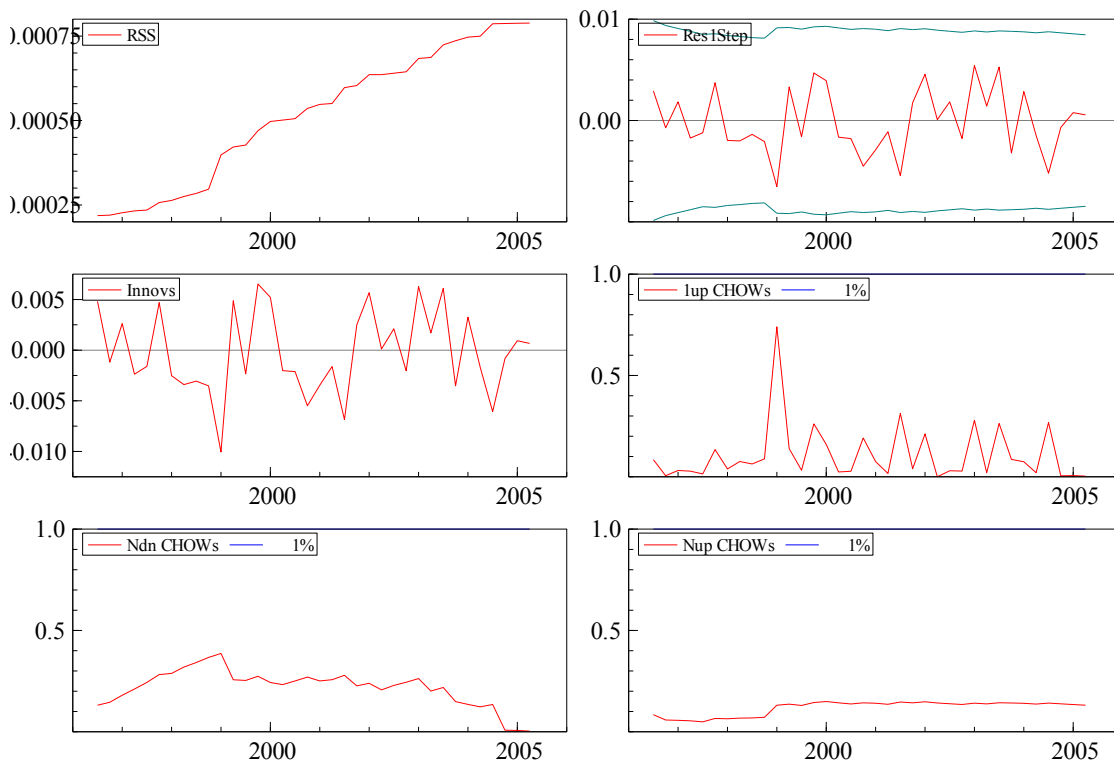


FIGURE 60: GERMAN MODEL, RECURSIVE EQCM COEFFICIENTS

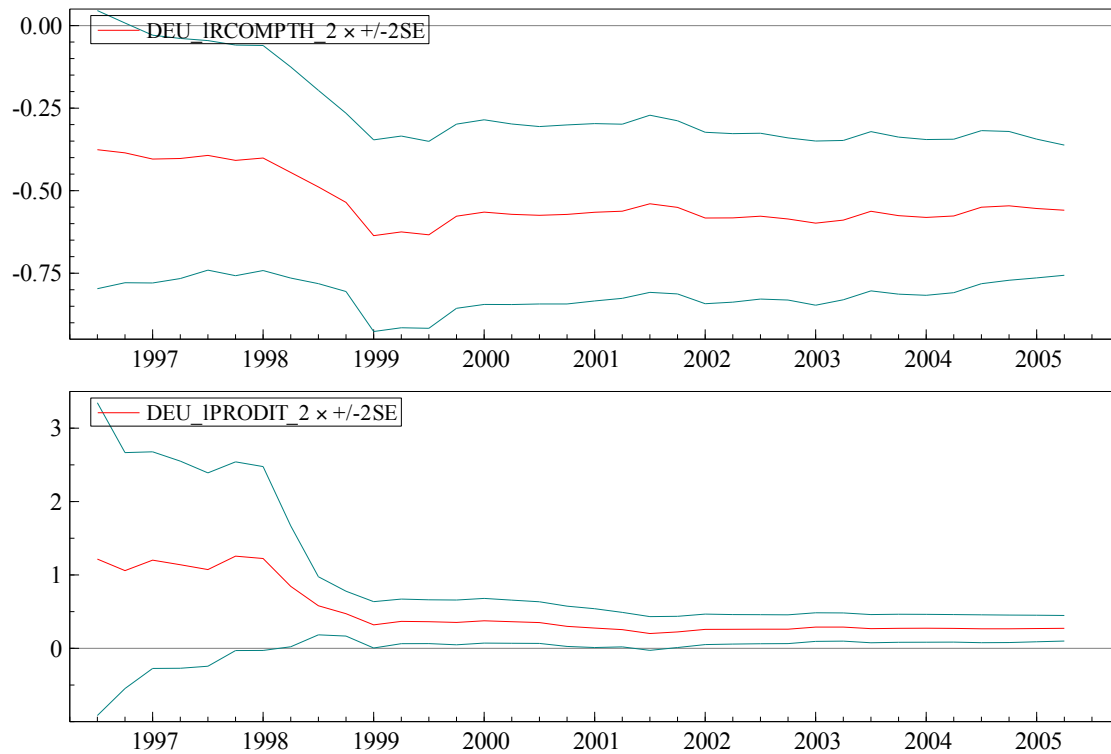
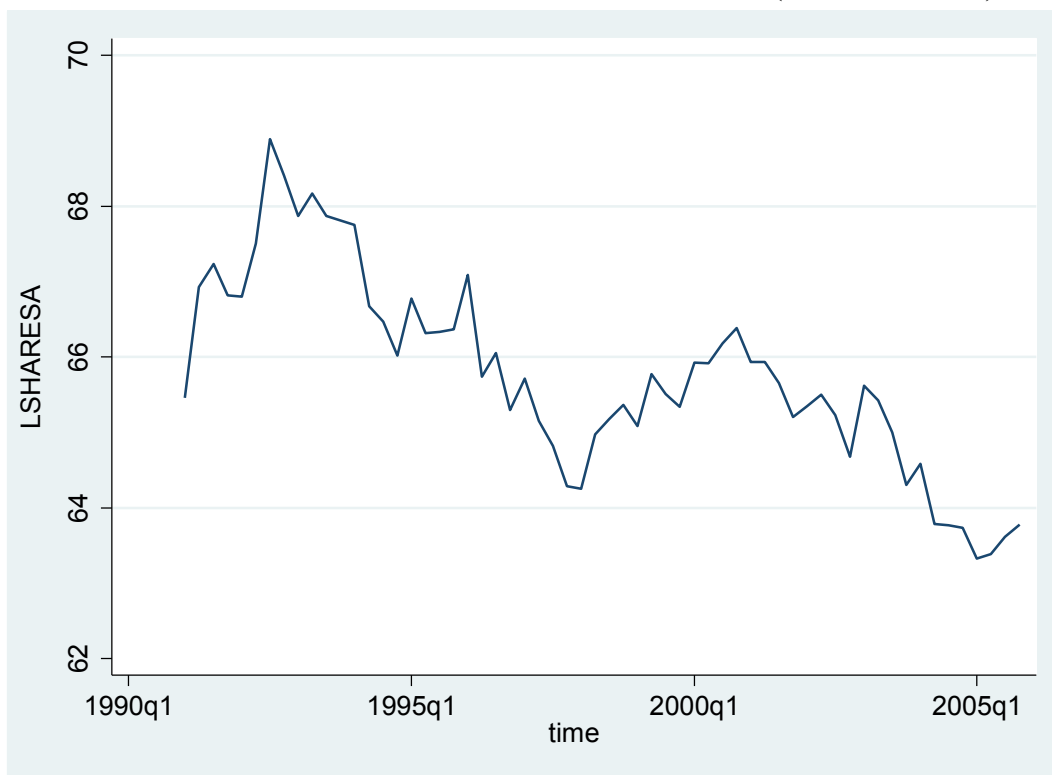


FIGURE 61: LABOUR SHARE OF GROSS VALUE ADDED (BASIC PRICES), SA



4.1.6 The Netherlands

Modelling NLD_D4lCOMPTh; Estimation Sample: 1974q3-2005q2					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-3.16784	0.2307	-13.7	0.000	0.6358
NLD_DlCOMPTh_1	0.230347	0.06937	3.32	0.001	0.0926
NLD_D3lCOMPTh_2	0.341442	0.04980	6.86	0.000	0.3033
NLD_DlCOMPTh_5	0.243376	0.06118	3.98	0.000	0.1278
NLD_EXPINFLPBIS_3	0.0941051	0.02623	3.59	0.001	0.1065
NLD_DlU	-0.0200008	0.01204	-1.66	0.099	0.0249
NLD_DVOLAT	0.773417	0.3207	2.41	0.018	0.0511
I1982:1	0.0299674	0.008985	3.34	0.001	0.0934
I1986:1987	-0.0372009	0.004148	-8.97	0.000	0.4268
I2001	0.0285016	0.005134	5.55	0.000	0.2220
I2004:1on	0.0206416	0.004313	4.79	0.000	0.1750
NLD_lRCOMPTh_4	-0.527844	0.03883	-13.6	0.000	0.6311
NLD_lPRODIT_4	0.527844*				
NLD_lCTAX_4	-0.527844*				
NLD_lITAX_4	-0.236830	0.05535	-4.28	0.000	0.1449
NLD_lU_4	-0.0423189	0.003958	-10.7	0.000	0.5142
NLD_lCOORD_4	-0.0889601	0.009792	-9.09	0.000	0.4332
NLD_VOLAT_4	0.415654	0.1790	2.32	0.022	0.0476
sigma	0.00864561	RSS	0.00807263797		
R^2	0.966726	F(15,108) =	209.2	[0.000]**	
log-likelihood	421.704	DW	2.07		
no. of observations	124	no. of parameters	16		
mean(NLD_D4lCOMPTh)	0.0446043	var(NLD_D4lCOMPTh)	0.00195655		
Specification Tests					
AR 1-5 test:	F(5,103) =	0.91811	[0.4724]		
ARCH 1-4 test:	F(4,100) =	1.4265	[0.2307]		
Normality test:	Chi^2(2) =	4.1285	[0.1269]		
hetero test:	Chi(26,81) =	2.0969	[0.0063]**		
RESET test:	F(1,107) =	0.034141	[0.8538]		

SOURCE: own calculations; starred * coefficients have been restricted.

Greater availability of key data in the Netherlands has allowed us to build a model for a considerably longer period than is possible in all other countries in our sample except Italy, namely: 1974q3-2005q2. As already noted in the case of Italy, the sample period available here presents us with an opportunity of examining the potential link between *both* EMU *and* the EMS on the one hand, and labour market flexibility on the other.

The model concerns the yearly change in the hourly compensation of workers D4ICOMPTH. Consistent with this level of differencing, the data suggest that the equilibrium correction term should be lagged by one year, and has an estimated speed of adjustment coefficient of 53%. This value and the structure of our specification appear reasonable in light of prior institutional information, namely that the usual duration of collective agreements in the Netherlands is of two years (European Commission, 2004).

Long-run homogeneity holds for the productivity trend, and the relevant equilibrium prices are *producer* prices as shown by the restriction on the indirect taxation variable CTAX. Similar evidence could only be found for Germany in our sample of countries. Rather than representing a uniquely Dutch or German characteristic, it may well be the case that for many countries only a large time series sample such as the one employed here can give us enough variation to allow an identification of the effects of this variable. Indeed, not all the variation in indirect taxation may be relevant to wage bargaining (see discussion in Section 4.1.7 in the UK context). Note however that the only other country studied here with a comparable sample size (Italy) failed to display a significant CTAX variable, contrary to what is the case here. This situation is not unexpected over the time period considered given an *a priori* view of greater product market competition in the Netherlands and its small open economy status. Surprisingly, however, the ratio of import to consumer prices does not appear to be significant, unlike in Belgium. Direct taxation ITAX also plays a significant and important role in the Dutch bargaining process but is not so large that it could be sensibly restricted.

The unemployment level is strongly significant, whilst the corresponding differenced term is just significant at the 10% level. This finding is broadly in line with those in the literature. Indeed, Lauer (1999) and Scheremet (2000) both found the *level* of unemployment to be highly significant, whereas its difference was not. An exception is however provided by Anderton and Barrell (1995), who reported a “*particularly low*” and insignificant long-run coefficient on unemployment for the Netherlands.

The long sample period available for this estimation poses particular challenges given the economic volatility present in the 1970s. We aim to control for this by introducing an index of inflation volatility in our equation. The rationale for its presence is that in a volatile environment, inflation surprises can hurt workers’ real compensation significantly over the contract period; as a result, they may seek to insure against that possibility *ex ante* in wage bargaining. We capture the concept of inflation volatility by the following moving average measure:

$$MA[abs(D4/CPI - D4/CPI_{-4})]$$

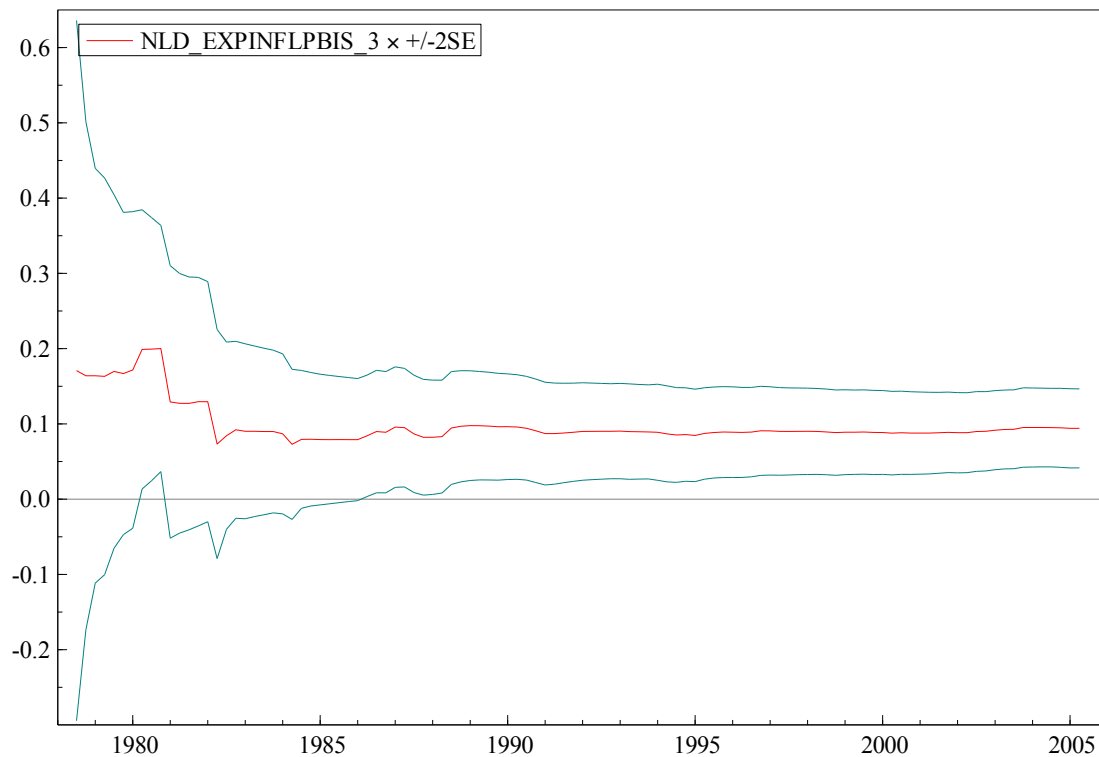
In contrast to findings for the Italian wage equation, inflation volatility is found to be significant both as a differenced variable and as part of the long-run solution of the equation. This is the first time this issue is considered in the literature, and the results here suggest that allowance should be made for the possibility of such effects when considering samples of this type.

In terms of institutional variables, a significant role is identified for bargaining coordination. Our index of coordination shows that a significant contribution is made by this variable to wage moderation. This result is consistent with multiple findings in the

literature pointing to a relationship between good employment outcomes and coordination – whether in a monotonic or non-linear fashion (Boeri et al., 2001). Finding evidence of it in the context of the Dutch labour market seems logical, given its operation as a highly coordinated system for decades and its characteristic *Overlegeconomie* (“social dialogue economy”) features. In particular, over our sample, the coordination index rises in the period following 1982, a year characterised by the introduction of the landmark *Wassenaar* agreement. This agreement amounted to a commitment to wage moderation reached by the social partners (ILO, 2005b).

The dynamics of the equation warrant further consideration. As can be seen, inflation dynamics were found to be insignificant in our specification, whilst the inflation expectations term is highly significant. This finding differs importantly from the results of earlier studies. Both Anderton and Barrell (1995) and Barrell and Dury (2001) found no role for expectations in the Dutch wage equation. Curto Millet (2004) also reached a finding of insignificance. Artis and Ormerod (1996) did include expectations in their specification, but the coefficient found was rather low. The difference in findings relative to our earlier study may well be due to the difference in sample period (1982q1-2002q4 in the earlier study), as one would expect inflation expectations to be of particular relevance in the higher inflation environment of the 1970s. Indeed, the coefficient appears to have had a large size from early on and has remained stable throughout our sample (Figure 62). The lag found in the expectations term may reflect both negotiation lags and staggering.

FIGURE 62: DUTCH INFLATION EXPECTATIONS, RECURSIVE COEFFICIENT



The specification includes a number of dummy variables which deserve comment. We have introduced a dummy for the period 1986-1987, which is characterised by a trough in the series of hourly compensation growth, as can be seen in Figure 63. These two years are in fact the years of strongest positive growth in average working hours in our sample (2.33% and 2.28%, respectively), representing a break in an otherwise steady trend towards decreasing working time. Part of this development may be explained by changes in working time regulations. Indeed, Nunziata (2001) reduced the value of his indicator for working time regulations for the Netherlands in 1986, noting both the abolition of the prohibition for women night shifts and a trend towards decentralisation on the issue of working time. Such developments, in the context of an environment of wage moderation, help account for the particular developments of this period.

The dummy for 2001 seeks to characterise a year of tense bargaining conditions. This was already noted in Curto Millet (2004). For instance, some employers affiliated to the VNO-NCW employers' confederation pushed for reductions in wage costs (notably in the area of sick pay), meeting frontal opposition from the Dutch Trade Union Federation (FNV). Tensions were also present inside the trade union movement, as a number of unions (e.g. FNV Bouw; FNV unions represented at the Akzo-Nobel chemicals group and in the healthcare sector) sought pay increases well above the 4% recommended by FNV – successfully, in the cases of some (EIRO, 2001d, 2001f).

Finally, we have included a dummy variable for the quarters from 2004q1 onwards. This variable broadly covers two exceptional developments in Dutch wage bargaining. First, the period is influenced by a central tripartite agreement signed in November 2003 (the ‘autumn agreement’), which imposes a wage freeze in 2004 and 2005 (EIRO 2004a). However, as noted in EIRO (2004b), trade unions were keen to “*test the limits of the wage freeze*” in negotiations during 2004. Second, the period is marked by a crisis of occupational pension funds. These are jointly managed by the social partners and constitute the “second pillar” of the Dutch pensions system, covering around 90% of the workforce (EIRO, 2004c). However, in recent years the pension fund reserves have been substantially depleted. The shortfall already amounted to EUR 10 billion by the end of 2002 and concerned 60% of all the 284 funds (EIRO, 2003a); it worsened subsequently. The logical consequences of this situation have been increases in employer and employee contributions. The scale of the problem is considerable; data from the Central Statistical Office (*Centraal Bureau voor de Statistiek*, CBS) reveal that social contributions by employers rose by EUR 7.5 billion in 2003, of which EUR 5.6 billion were related to

occupational pension contributions. A large part of the reason behind this lies in stricter requirements imposed by the *Pensioen- en Verzekeringskamer* (Pension and Insurance Supervisory Board, PVK), and it is expected that a further EUR 70 billion will be required for compliance in the coming years (EIRO, 2004d). The crisis is associated with a number of issues, including an ageing population, the shrinking capital reserves held by pension funds due to the stock market slump and new EU book-keeping rules (EIRO, 2003a). As a result, the ‘pensions problem’ has become a prominent issue in Dutch bargaining (EIRO, 2003b).

The specification tests show the model to be congruent in all aspects but heteroscedasticity, with an excellent fit (sigma of 0.86%). This residual amount of heteroscedasticity could not be eliminated and is without doubt linked to the larger sample used here (it is absent over the shorter period considered in Curto Millet, 2004). The autocorrelation result is especially encouraging given the smoothness of the series being modelled and the time period considered.

Stability tests presented in Figure 65 confirm the stability of the model over the entire period examined. In particular, neither EMS nor EMU effects are apparent here. Probing the possibility of an EMU effect further with the forecast analysis leads to the same conclusion (Figure 64) – the forecasts fail, but merely due to the absence of the special “dummied” effects of 2001 and 2004 described earlier.

Individual recursive estimates of the coefficients show no break. Figure 66 plots recursive estimates of the key EqCM coefficients, once again making clear the absence of any ‘EMU effect’.

FIGURE 63: DUTCH MODEL, GRAPHICAL ANALYSIS, 1974Q3-2005Q2

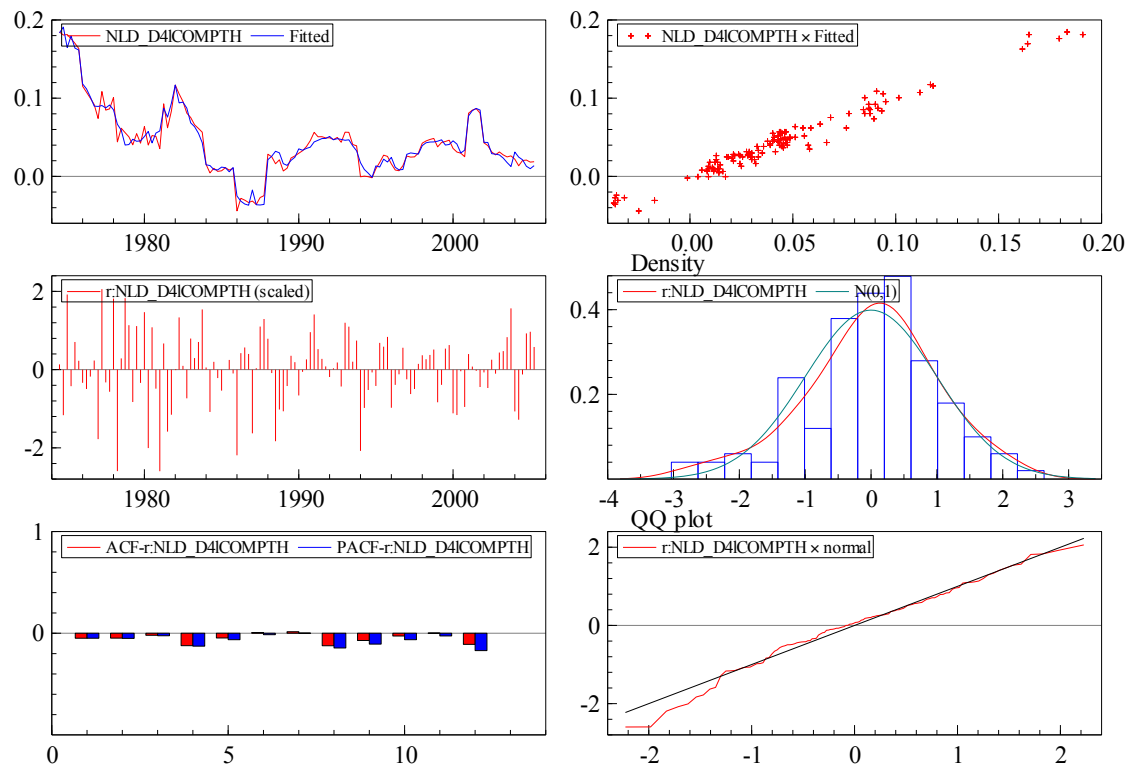


FIGURE 64: DUTCH MODEL, FORECASTS FROM 1999Q1

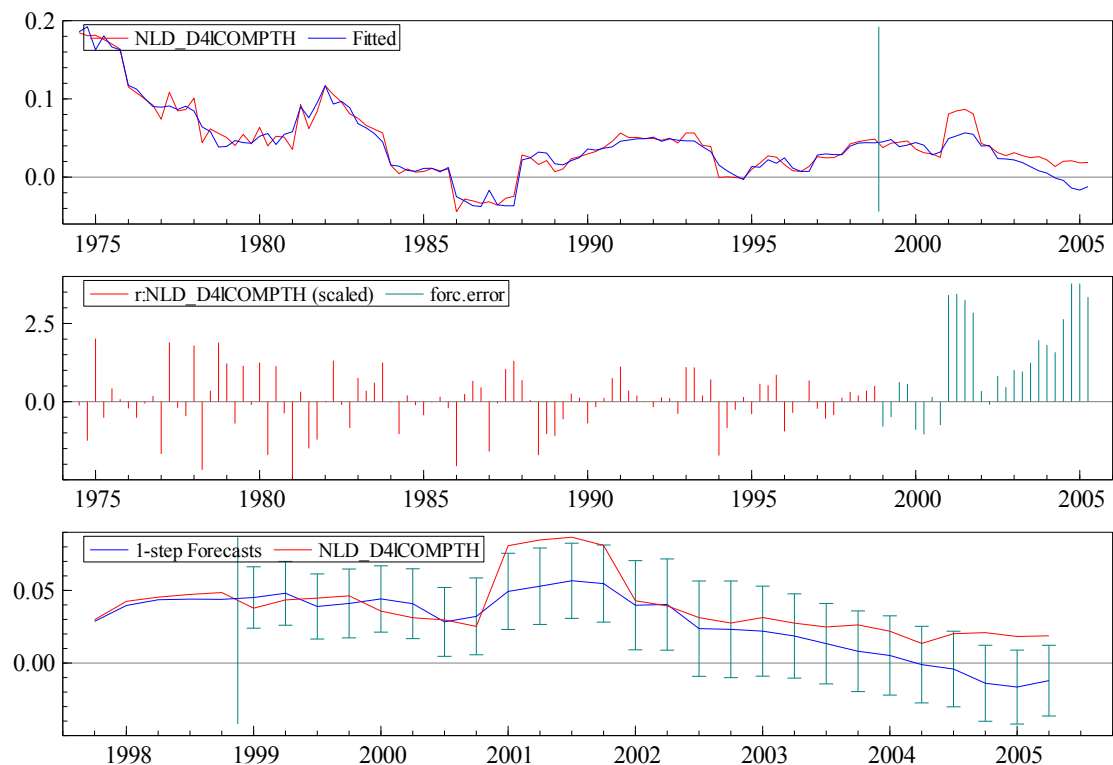


FIGURE 65: DUTCH MODEL, STABILITY TESTS

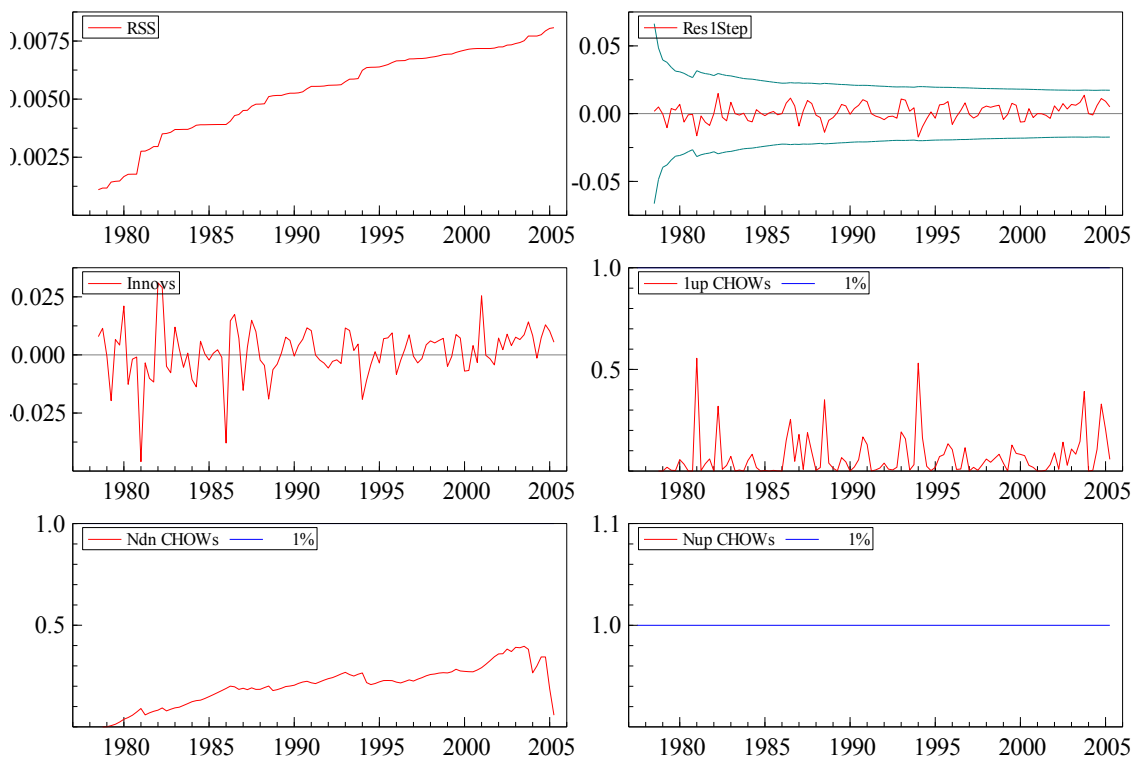
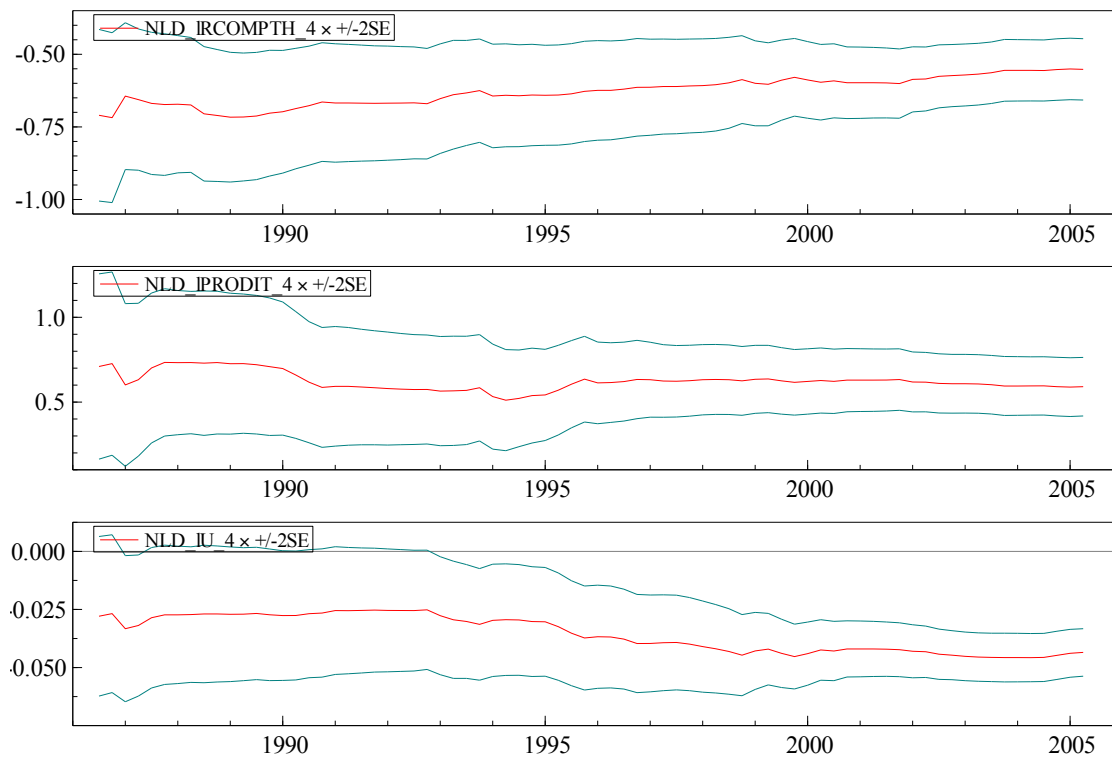


FIGURE 66: DUTCH MODEL, RECURSIVE EQCM TERMS



4.1.7 United Kingdom

Modelling GBR_D4lCOMPTH; Estimation Sample: 1983q1-2005q1					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-1.56834	0.2749	-5.70	0.000	0.3055
GBR_DlCOMPTH_1	0.608526	0.07396	8.23	0.000	0.4777
GBR_DlCOMPTH_2	0.510956	0.07608	6.72	0.000	0.3787
GBR_DlCOMPTH_3	0.393911	0.08400	4.69	0.000	0.2291
GBR_EXPINFLPBIS_3	0.546320	0.1054	5.18	0.000	0.2663
GBR_DlCPI_1	0.245233	0.1041	2.36	0.021	0.0697
GBR_DlCPI_2	0.328881	0.1020	3.22	0.002	0.1231
GBR_D4lU	-0.0615586	0.01036	-5.94	0.000	0.3232
I1992:1	0.0413339	0.007049	5.86	0.000	0.3173
I1988:1	-0.0360666	0.007442	-4.85	0.000	0.2409
I1991:1	0.0336455	0.007290	4.62	0.000	0.2235
GBR_lRCOMPTH_4	-0.422542	0.05338	-7.92	0.000	0.4585
GBR_lPRODIT_4	0.166931	0.03965	4.21	0.000	0.1933
GBR_lEPL_4	0.116566	0.01858	6.27	0.000	0.3472
GBR_lRELPMP_4	-0.152075	0.02017	-7.54	0.000	0.4345
sigma	0.00661348	RSS	0.00323662446		
R^2	0.933913	F(14,74) =	74.7	[0.000]**	
log-likelihood	328.587	DW	1.9		
no. of observations	89	no. of parameters	15		
mean(GBR_D4lCOMPTH)	0.0582597	var(GBR_D4lCOMPTH)	0.000550286		
Specification Tests					
AR 1-5 test:	F(5,69)	=	0.27186	[0.9270]	
ARCH 1-4 test:	F(4,66)	=	0.39278	[0.8131]	
Normality test:	Chi^2(2)	=	2.0088	[0.3663]	
hetero test:	F(25,48)	=	1.2677	[0.2359]	
RESET test:	F(1,73)	=	1.2746	[0.2626]	

SOURCE: own calculations; starred * coefficients have been restricted.

For comparison with the Eurozone results presented above, we have also implemented our wage equation in the context of the United Kingdom over the period 1983q1-2005q2.

The model for the UK concerns the yearly change in the hourly compensation of workers, D4lCOMPTH. Consistent with this level of differencing, the data suggest an

equilibrium correction term lagged by one year, which displays a speed of adjustment coefficient of 42%.

The long-run solution implied by the above regression is surprising in many ways. The absence of a long-run unemployment coefficient is perhaps its most visible feature⁸² and it seems that the role of unemployment is fully subsumed within the dynamics of the equation, with the *yearly* change in unemployment. The insignificance of the level of unemployment reproduces earlier results in the literature by Anderton and Barrell (1995) and Anderton et al. (1992), whilst Scheremet (2000) noted the weakness of the variable's effect.

Furthermore, the equilibrium surprises by a double failure to accept long-run homogeneity – a finding already highlighted in Curto Millet (2004).

First, the absence of the consumer tax wedge term CTAX is remarkable, as one would have expected the UK to have a long-run solution based on producer prices, given its competition framework and the extreme decentralisation of its bargaining process; indeed, company-level bargaining overwhelmingly dominates the system, as emphasised by the European Commission (2004) or the data in Ochel (2001). This absence is also surprising given the significance of the coefficient corresponding to the relative price of imports to CPI (RELPMPC). A plausible explanation for these findings could be that our CTAX variable displays relatively little variation over the sample period, making an identification of its effect difficult. Indeed, the main rate of VAT in the UK has been constant at 17.5% since 1991, and was constant at 15% in our sample before that

⁸² When estimated, the coefficient has a value of -0.00995, with a p-value of 0.31.

(European Commission, 2005a). Considerably more variation has been exhibited by excise duties, such as ‘sin taxes’, which can be hypothesised not to affect the wage bargain. Second, we fail to impose long-run homogeneity for the productivity variable. Thus, the equilibrium in the UK appears to indicate a situation in which wages fail to rise in line with productivity developments. This finding casts doubt on a strategy of systematically imposing long-run homogeneity in this respect, such as that followed in Anderton and Barrell (1995) and Barrell and Dury (2001).

The only institutional variable significantly present is employment protection legislation, which saw a modest increase over the sample period recently, between 1999 and 2002. These changes notably correspond to a reform introduced in 2000 halving the trial period in regular contracts, and a reduction to four years of the maximum total duration of fixed-term contracts (heretofore unlimited) in 2002.

The dynamics of the equation feature strong roles for both past inflation and inflation expectations. Finding a role for inflation expectations has been common for the UK in the literature. On the other hand, the results here contrast with those in Barrell and Dury (2001), whose resulting equation exhibited neither dynamics in past inflation nor in the dependent variable. A number of dummies were also added to control for statistical outliers (with standardised residuals close to three standard errors).

The specification tests show the model to be congruent, with excellent fit (sigma of 0.66%). The autocorrelation result is especially encouraging given the smoothness of the series being modelled. The equation appears to be stable according to the tests presented

in Figure 69. There does seem to be a certain amount of coefficient shifting in the post-1999 period, as evidenced by the forecasts in Figure 68. A detailed examination of the recursive coefficients confirms the absence of breaks, but evidences some smooth changes, notably:

- A decrease of the wage dampening effect of unemployment dynamics
- A decreased weight on inflation expectations
- An increased weight on the relative price of imports
- A decrease in the importance of the speed of adjustment coefficient

For instance, there has been a trend decrease in UK inflation expectations in recent years. Failure to adjust the coefficient on this variable would lead to an exaggeration of its dampening effect on wage claims over this period, which may contribute to the situation observed in Figure 68.

Recursive estimates of the EqCM coefficients show no break, as displayed in Figure 70. The figure also includes the evolution of the (insignificant) long-run unemployment term for information.

FIGURE 67: UK MODEL, GRAPHICAL ANALYSIS, 1983Q1-2005Q1

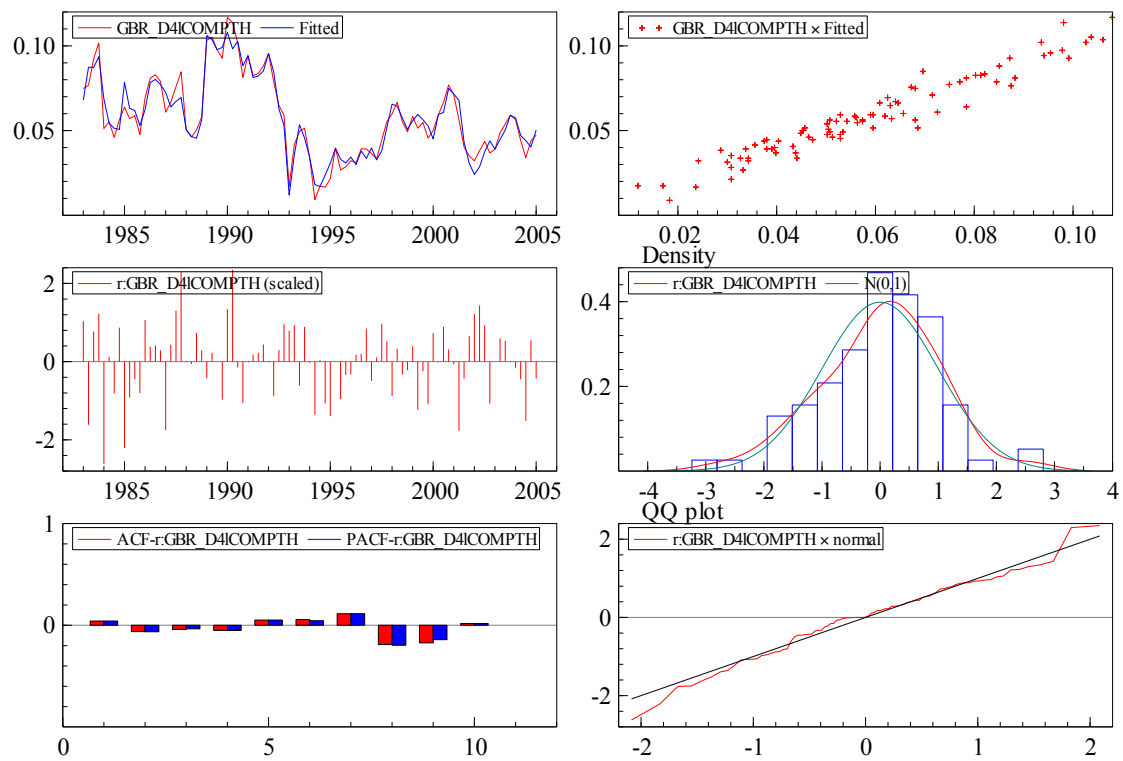


FIGURE 68: UK MODEL, FORECASTS FROM 1999:1

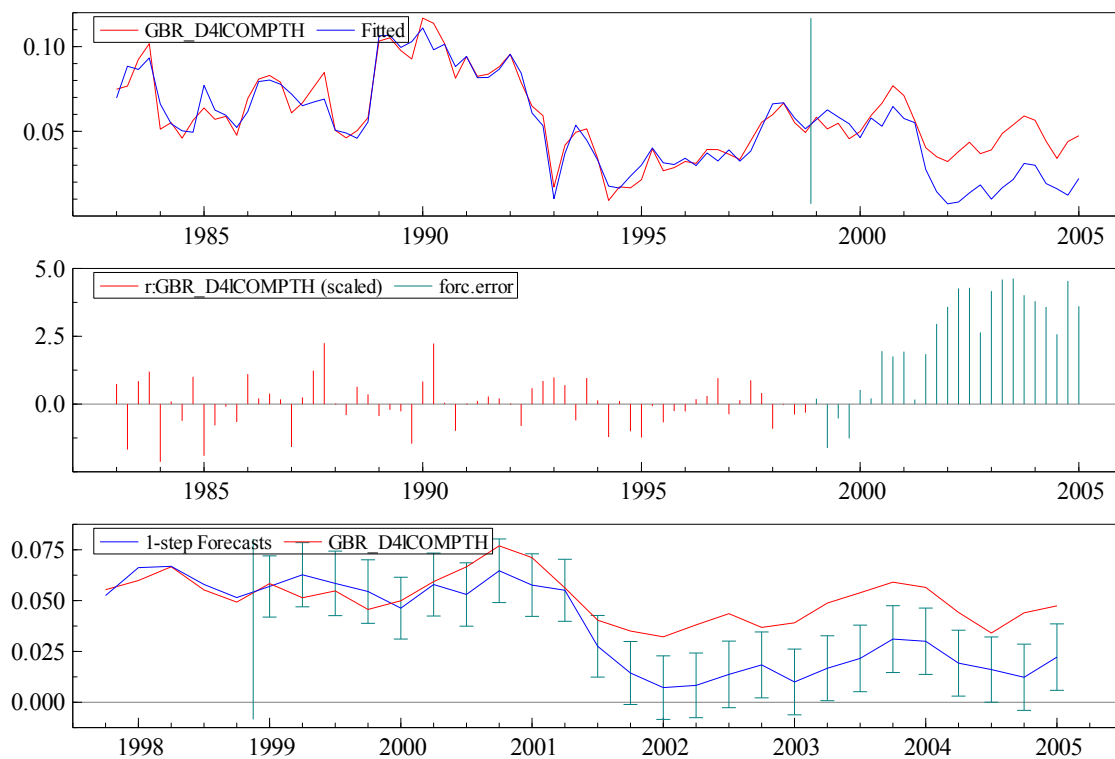


FIGURE 69: UK MODEL, STABILITY TESTS

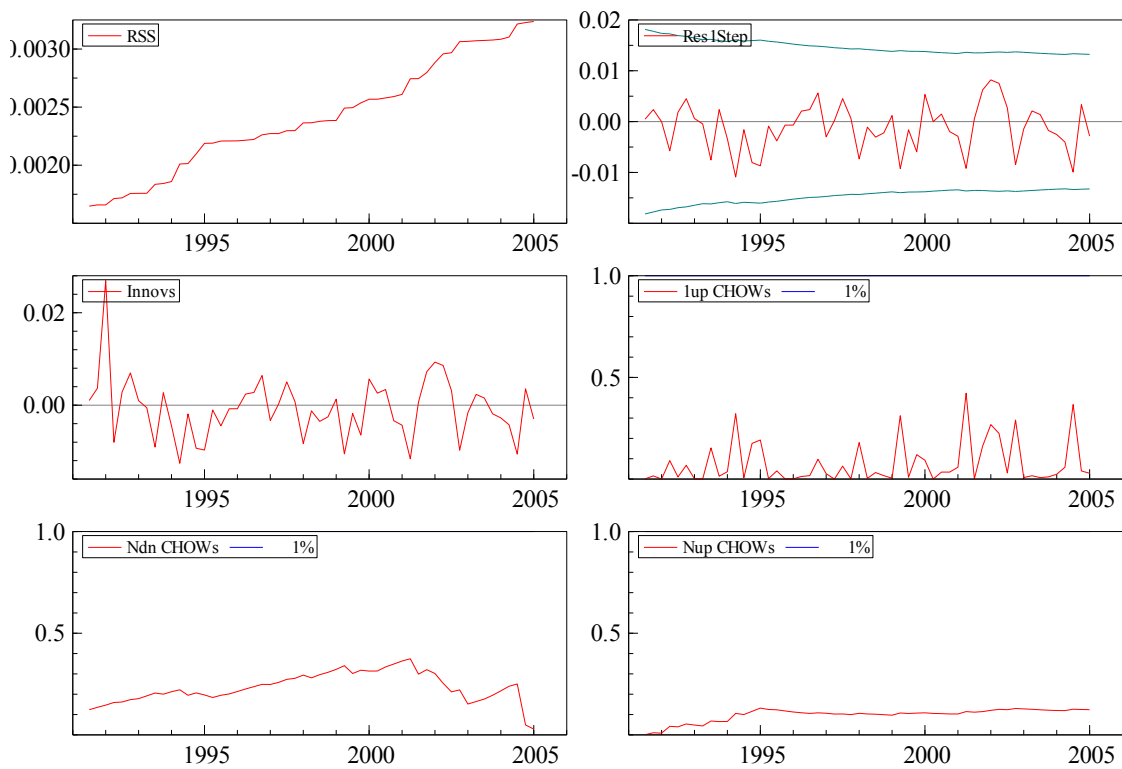
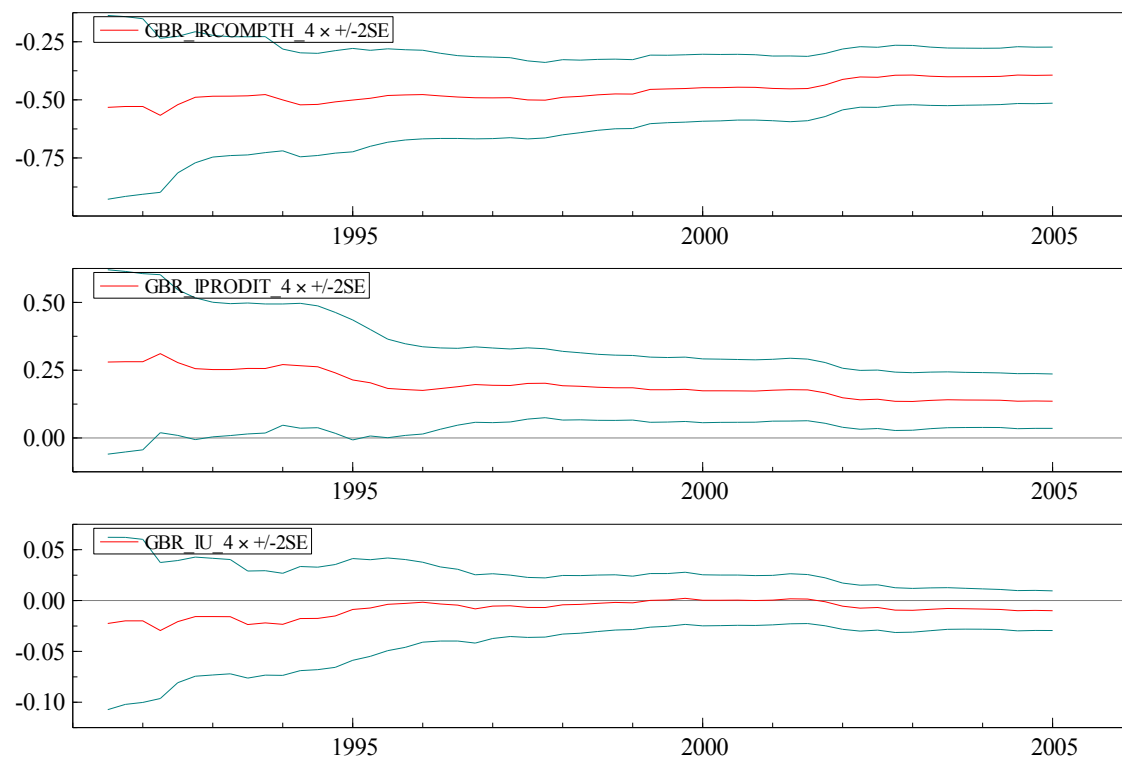


FIGURE 70: UK MODEL, RECURSIVE EQCM COEFFICIENTS



4.1.8 Sweden

For a further comparison with the earlier Eurozone results, we also deployed our wage equation in the context of Sweden. In this case, though, the extreme sample limitations which would be imposed by the qualitative survey data have led us to exclusively use the quantitative survey data collected by SCB and, more recently, by GfK (c.f. Chapter 1, Section 5.2).

These results were already presented as an example in Chapter 1. The discussion provided below is however much more complete than the summary offered earlier.

Modelling SWE_D12ICOMPTH; Estimation Sample: 1983q1-2005q1					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-13.4651	2.200	-6.12	0.000	0.3391
SWE_D41COMPTH_1	0.301122	0.06065	4.97	0.000	0.2525
SWE_EXPSCBEV	0.288568	0.1191	2.42	0.018	0.0744
SWE_D41CPI	0.466969	0.09037	5.17	0.000	0.2678
SWE_D41CPI_4	0.830359	0.07022	11.8	0.000	0.6570
SWE_D41CPI_8	0.802234	0.09408	8.53	0.000	0.4990
SWE_D41U	-0.0157295	0.009014	-1.74	0.085	0.0400
SWE_D41U_4	-0.0281434	0.01090	-2.58	0.012	0.0837
SWE_D41U_8	-0.0737631	0.007703	-9.58	0.000	0.5568
I1999	-0.0221632	0.004721	-4.69	0.000	0.2319
I1995:1	-0.0225431	0.008142	-2.77	0.007	0.0950
I2002:1	0.0235681	0.007949	2.96	0.004	0.1075
SWE_1RCOMPTH_12	-0.701336	0.04742	-14.8	0.000	0.7497
SWE_1PRODIT_12	0.514141	0.07215	7.13	0.000	0.4102
SWE_1U_12	-0.0710404	0.009310	-7.63	0.000	0.4437
SWE_1ACOV_12	2.54092	0.5211	4.88	0.000	0.2457
sigma	0.0075778	RSS	0.00419188561		
R^2	0.987712	F(15, 73) =	391.2 [0.000]**		
log-likelihood	317.079	DW	1.8		
no. of observations	89	no. of parameters	16		
mean(SWE_D121COMPTH)	0.166093	var(SWE_D121COMPTH)	0.00383307		

Specification Tests				
AR 1-5 test:	F(5, 68)	=	1.6675	[0.1543]
ARCH 1-4 test:	F(4, 65)	=	0.66158	[0.6209]
Normality test:	Chi ² (2)	=	3.2054	[0.2014]
hetero test:	F(27, 45)	=	0.75711	[0.7774]
RESET test:	F(1, 72)	=	0.14568	[0.7038]

SOURCE: own calculations; starred * coefficients have been restricted.

The selected wage equation for Sweden has an especially interesting dynamic structure. The dependent variable modelled here is the three-year change in the hourly compensation of workers, D121COMPTH. The equilibrium term is accordingly lagged by three years. This is unique among the countries in our sample and is reflective of very interesting idiosyncratic features of the Swedish labour market.

Sweden is presently with Denmark (and possibly Cyprus) the country with the longest usual duration of collective agreements in the European Union (three years) (European Commission, 2004). Fregert and Jonung (1998) examine historical data on the average length of blue-collar worker collective agreements in Sweden. There were a number of major two-year agreements in the 1980s, although it would appear that three-year agreements were already in existence in the labour market (Swedish Institute, 2005). Fregert and Jonung (1998) locate the shift to a system mainly based on three-year agreements in 1995, although as noted in EIRO (2004e) the current formalised system of three-year pay agreements in both the private and public sectors was initiated with the 1998 bargaining round.

The equation exhibits a sensible long-run solution, with strongly significant productivity and unemployment effects. The speed of adjustment is found to be 70%. Long-run

homogeneity for productivity cannot quite be imposed, however. In terms of institutional variables, only bargaining coverage could be retained, although it is estimated imprecisely given the limited variation that is typical of this series (in this case, a single continuous increase over the years 1990-1995).

Unemployment dynamics over the past three years are found to be strongly significant⁸³. The inflation dynamics are equally strong in this equation and the coefficients for inflation in the previous two years could (almost) be restricted to one, and remain high for the more recent changes within the past year. This could have been consistent with a bargaining framework in which expectations do not matter *ex ante* – because, for instance, unions are strong enough or otherwise confident to obtain compensation *ex post*. Nevertheless, the equation also displays a forward-looking component as evidenced by the significance of contemporaneous inflation expectations.

The dummy variable introduced for 1999 aims to control for an important trough in the series for the growth of hourly compensation of workers. This was a relatively quiet year between bargaining periods, with most parts of the public and private sectors covered by the results of the 1998 negotiations. The year was characterised by controlled wage increases with only marginal examples of wage drift. The four new collective pay agreements negotiated at the sectoral level in 1999 all followed the 3% pay increase norm for the year that was set in the 1998 bargaining round (EIRO, 1999c). The moderation in hourly compensation growth observed in 1999 was further importantly influenced by a decrease in employer social contributions, which fell by close to 10%

⁸³ We note that the results here are in stark contrast with those of Scheremet (2000), who found no significant role for unemployment, whether in terms of changes or levels.

according to the OECD Revenue Statistics Database, mainly on account of reductions in employer contributions to certain pension categories. This is set against the background of the introduction of a new pension system in Sweden in 1999 (Swedish Institute, 2004).

The specification tests show the model to be congruent in all aspects, with an excellent fit (sigma of 0.76%). That there is no significant autocorrelation when modelling a three-year change in compensation over the period of the last twenty years is rather remarkable.

Stability tests presented in Figure 73 confirm the stability of the model over the entire period examined. The forecast analysis is quite satisfactory, although there is some marginal evidence of forecast failure early on (Figure 72). However, this is solely due to the fact that the moderating effect of the 1999 dummy variable, which is outside the model, is missing in prior forecasting. Indeed, forecasting for the full post-1999 period gives entirely satisfactory results.

Individual recursive estimates of the coefficients show no break. Figure 74 plots recursive estimates of the key EqCM coefficients which appear to be stable.

FIGURE 71: SWEDISH MODEL, GRAPHICAL ANALYSIS, 1983Q1-2005Q1

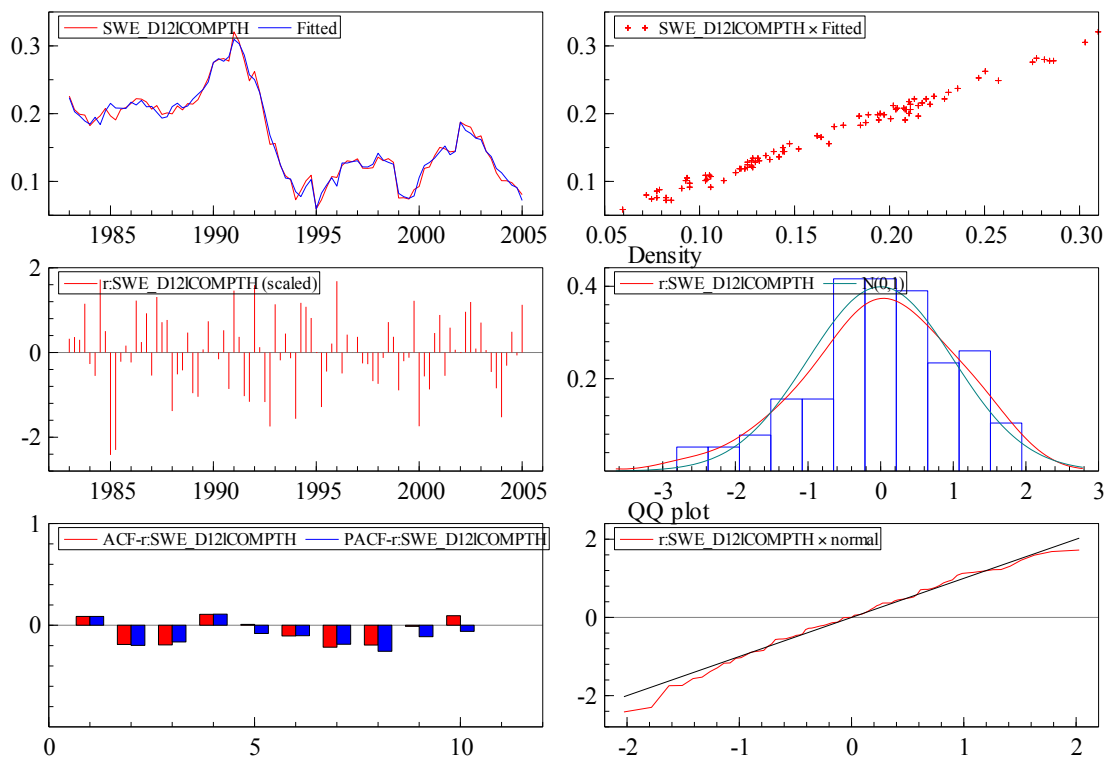


FIGURE 72: SWEDISH MODEL, FORECASTS FROM 1999:1

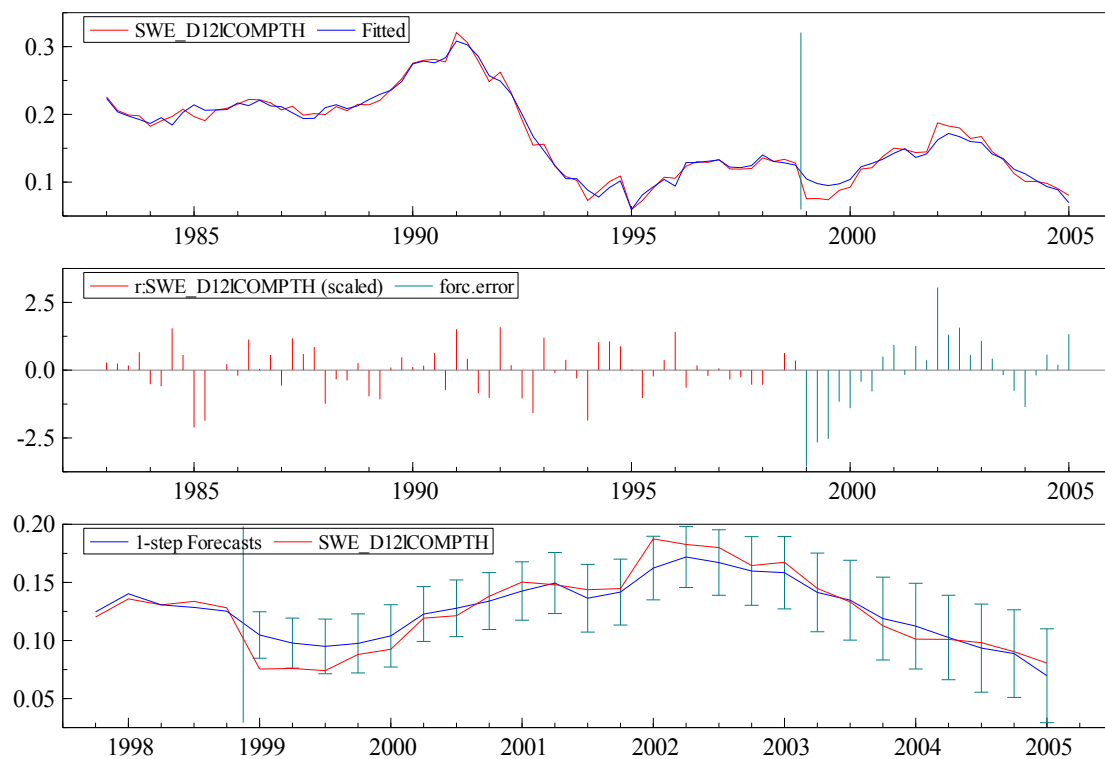


FIGURE 73: SWEDISH MODEL, STABILITY TESTS

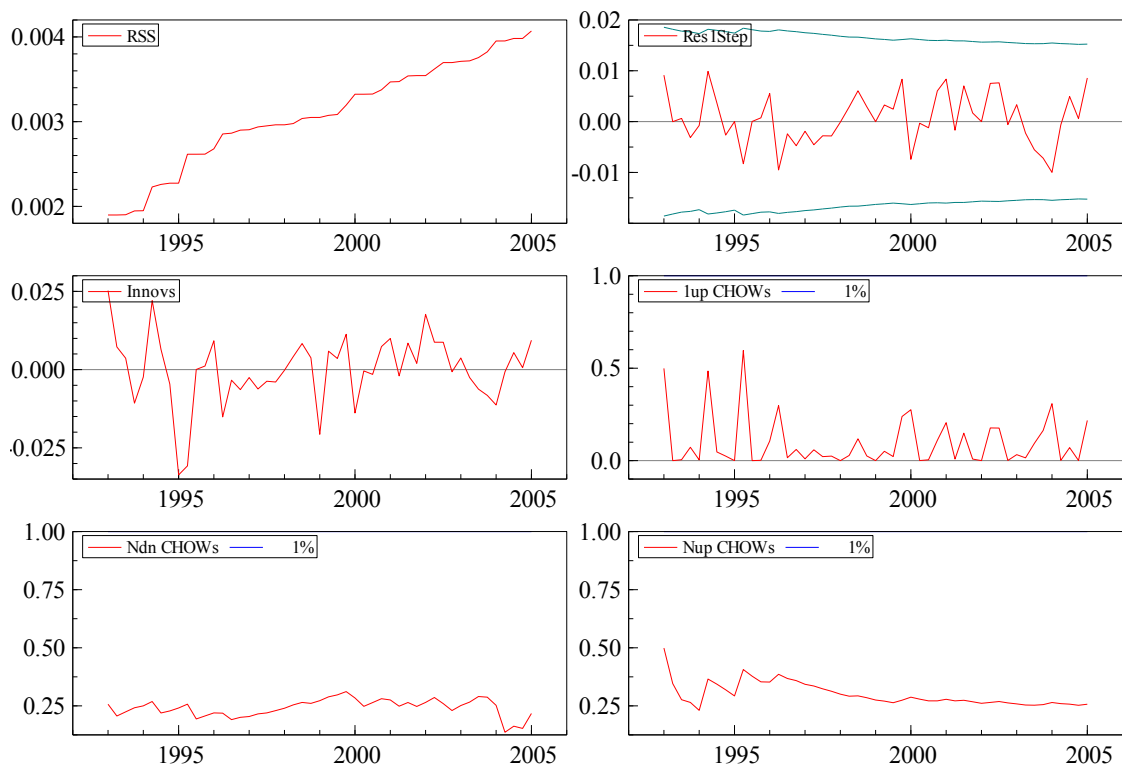
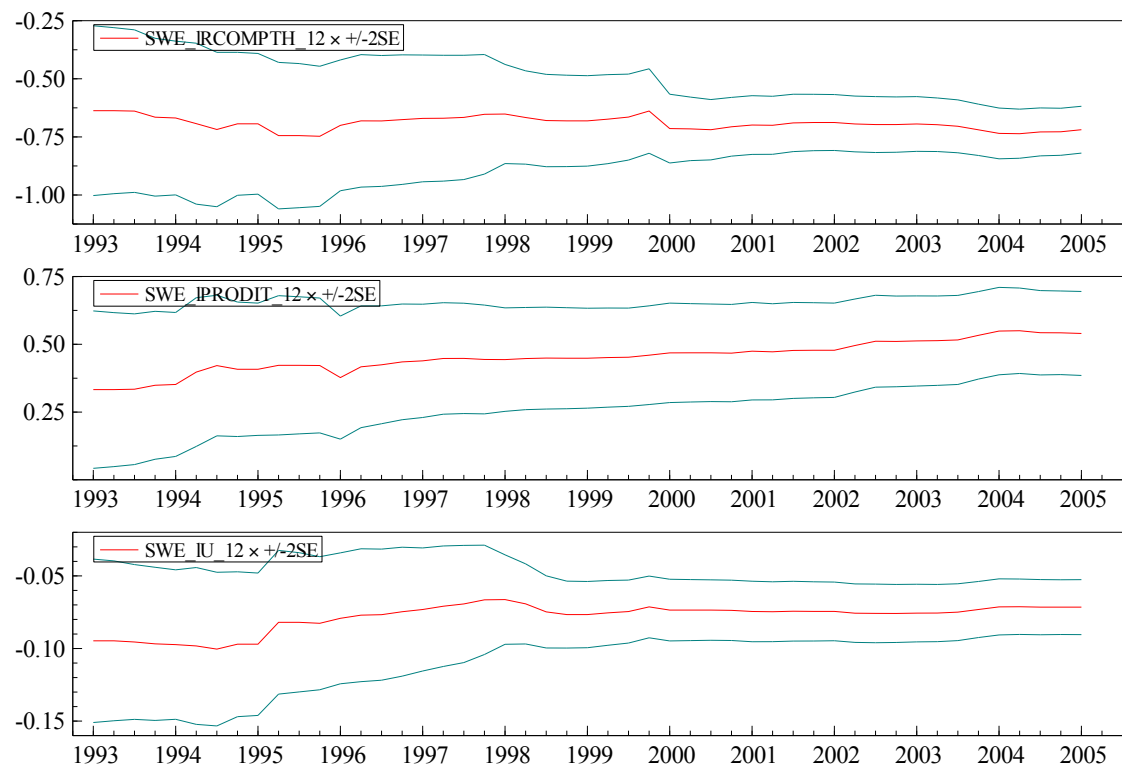


FIGURE 74: SWEDISH MODEL, RECURSIVE EQCM COEFFICIENTS



4.2 Comparative overview of the results

Taken together, the above results yield a number of striking observations.

First, we were able to provide an unambiguous answer to the central question of this study, namely the hypothesis of an “EMU effect” on wage flexibility. The speed of adjustment, the long-run elasticity of real compensation with respect to unemployment and the long-run productivity effect all displayed remarkably stable coefficients with the introduction of the euro, as judged by their recursive estimates. Nor was the overall stability of the models in question; the short-run coefficients were similarly stable overall. Furthermore, no convincing “EMS effects” on flexibility were found for the two countries whose sample size allowed us to explore the issue, namely Italy and the Netherlands. Thus, we must reject the null hypotheses of a visible linkage between EMU/EMS and increased wage flexibility. There was, however, some evidence of increased wage moderation in Belgium, which could have been supported by the entry into EMU. Further discussion is provided in Section 5.

Second, we were able to impose long-run homogeneity for the productivity trend in all Eurozone countries in our sample except France and Germany. Homogeneity could not be imposed for the non-Eurozone countries considered (the UK and Sweden). In all cases however, the variable was found to be highly statistically significant. The strength of these productivity results suggests that earlier findings of non-homogeneity – and in some instances even the absence of a long-run effect – may in many cases have been caused by misspecification, possibly due to omitted variable bias (Scheremet, 2000; Lauer, 1999).

This finding of homogeneity is an encouraging feature of equilibrium, implying a policy of keeping unit labour costs unchanged. A productivity-oriented wage policy is an important requirement for a successful monetary union, given that exaggerated wage increases can no longer be offloaded onto monetary or exchange rate policy. Our recursive estimates show that this coefficient has proved to be stable in the years preceding EMU, demonstrating that this condition was met before monetary unification. This is notably the case for the Southern countries in our sample, Italy and Spain, whose nominal unit labour cost developments in excess of the EU average had tended to be compensated by devaluation in the past (Lauer, 1999). It is also apparent from the recursive productivity coefficients for both Italy and Spain that these exhibited some growth in the late 1990s. Importantly, such growth was compatible with the homogeneity restriction; this may perhaps reflect the influence of the prospect of EMU on wage bargaining, with social partners realising that a strong productivity orientation would be required.

The strength of the homogeneity results reached here is striking given that we are considering *whole economy* data. Indeed, the absence of ‘hard’ budget constraints in sectors such as public employment could lead one to expect a looser relationship with productivity, which would be reflected in the aggregate data to some extent. It may well be the case that the fiscal tightening imposed by the Maastricht Convergence Criteria in view of EMU kept the public sector settlements more in line with private ones in most Eurozone countries considered over recent years.

As already noted, it was not possible to impose the homogeneity restriction in either France, Germany, Sweden or the UK. However, only in the cases of Germany and the UK is the long-run productivity coefficient substantially smaller in magnitude than the speed of adjustment coefficient. The UK in particular has the smallest such coefficient relatively. This finding was already noted in Curto Millet (2004). The UK is certainly no outlier in terms of the size of its public sector (only Germany and the Netherlands have lower shares of public sector employment; Cadiou et al, 1999a). The finding translates a tendency for wages to grow by less than productivity on average in the UK, and thus in a declining labour share over the period overall (see Cameron and Wallace, 2002).

Third, inflation expectations have been found to be an important element of the dynamics in all countries (except France). Often, these entered the wage equation with some lags, possibly suggesting discrete information collection – much in the spirit of the Mankiw and Reis (2002) theory of inflation stickiness – and bargaining lags. In no equation were multiple lags of expectations retained. These might perhaps have been expected as a reflection of staggering in the economy (Section 3.1), or of differences across groups in bargaining complexity⁸⁴. However, the detection of such a structure may have been complicated by issues of multicollinearity to some extent, and may be implicit in the lags of expectations evidenced. The finding of significant inflation expectations for both Germany and the Netherlands is essentially new in the literature, and may be reflecting the use of better measures of inflation expectations in this study, as well as longer samples than were heretofore available. On the other hand, the absence of expectations in

⁸⁴ For instance, one could imagine a situation in which all groups reach an agreement at the same time, but the complexity of bargaining for a particular sector requires it to start negotiations and collect information a few quarters ahead of the others.

the case of France seems most likely to be related to measurement error (see Chapter 1), although institutional reasons discussed earlier would have led us to expect a relatively moderate impact of the variable in any event. The exceptional forecasting failure exhibited by the French equation upon the introduction of the euro – a pattern replicated by no other country in our sample – may in part be a testimony to the importance of including expectations in wage equations.

Fourth, the tax wedge terms were found to be mostly insignificant in the equilibrium of the equations. However, they featured very prominently in both the Dutch and German wage equations, notably suggesting in both countries that the “producer wage” was the relevant equilibrium concept. Obviously, the relatively limited variation offered by such variables over the sample period makes them difficult to identify, especially when taking into account the fact that not all such variation may be of relevance to the wage bargain (see the discussion in the UK context, in Section 4.1.7). Nevertheless, the findings here suggest that it might be optimal to err on the side of caution when considering these variables. Notably, this would mean refraining from imposing a long-run solution in producer prices (Barrell and Dury, 2001) or from imposing equal effects in bargaining from direct and indirect taxation on grounds of “rationality” (Anderton et al., 1995).

Fifth, it can be said that institutional variables have had a relatively limited presence in the above equations. This could be predicted from Section 3.6, as many such variables lacked enough variation over the sample period to ensure identification. Nonetheless, a number of significant findings were made. These are summarised in Table 31.

TABLE 31: SIGNIFICANT INSTITUTIONAL VARIABLES, SUMMARY

	BEL	FRA	NLD	ESP	DEU	GBR	SWE	ITA
BD		X						
BRR	X	X			X			
ACOV	+				X		X	
UDENS	X	X		X				
COORD			X		+	+		X
EPL				X		X		

+: variable constant in sample period

Sixth, the country results allow us to finally carry out a cross-country comparison in terms of wage flexibility dimensions (Table 32).

TABLE 32: WAGE FLEXIBILITY MEASURES, ALL COUNTRIES

Concept	DEU	FRA	ITA	ESP	BEL	NLD	UK	SWE
Equilibrium Correction Half-life (quarters)	1.7	5.6	65.5	2.7	11.0	3.70	5.1	6.9
Long-run unemployment effect ⁸⁵	0	0.162	0.279	0.035	0.145	0.080	0.024	0.101
Long-run productivity	0.488	1.17	1	1	1	1	0.396	0.733

The “equilibrium correction half-life” results presented above show the number of quarters required to remove half the effect of a shock *on the basis of the equilibrium correction process alone* (i.e. ignoring any dynamics). This is measured *from the point at which the EqCM mechanism starts acting*, which therefore excludes the “waiting period” implied by the EqCM lag in the wage equations⁸⁶.

As can be seen, the Italian wage equation is distinguished by an exceptionally slow rate of convergence to the equilibrium. This is notably due to what essentially amounted to a

⁸⁵ The UK coefficient was insignificant, but the point estimate is included here for purposes of comparison. The German (insignificant) point estimate is negative, and therefore restricted to zero in the table above.

⁸⁶ The only effect of accounting for the “waiting period” in terms of ranking is a swap of places between France and the UK relative to the ranking in Table 33. If we further take the autoregressive dynamics into account (following a fully exogenous shift in the equilibrium), Sweden and Belgium would be the only countries trading places relative to the ranking displayed.

disappearance of equilibrium-correcting behaviour in the 1980s, as discussed earlier in the text. The weakness of equilibrium correction may also be reflecting the well-known regional specificities of the Italian labour market (e.g. Obstfeld and Peri, 1998). We note that the results for the long-run unemployment coefficient in Table 32 confirm a “traditional” finding in the wage equation literature, namely that real compensation in Britain is weakly influenced by the unemployment level in equilibrium, whilst that in Italy is strongly so (Layard et al. 1991, Cadiou et al. 1999b).

We now present the country rankings obtained in this study and in previous research regarding the speed of equilibrium correction (Table 33). As can be seen, there is a large amount of variation in this respect in the literature. The interpretation of the table requires some care, as the sample periods differ substantially across studies (see Table 25 and Table 26). The only “stylised fact” that emerges relatively clearly is the consistent position of Germany and Spain at the top of the table.

TABLE 33: COUNTRY RANKINGS, SPEED OF EQUILIBRIUM CORRECTION

Study	BEL	NLD	ESP	FRA	ITA	DEU	UK	SWE
This Thesis	7	3	2	5	8	1	4	6
Barrell and Dury, 2001	1	7	3	6	4	2	5	N/A
Anderton and Barrell, 1995	7	5	1	6	4	2	3	N/A
Anderton et al., 1992	N/A	N/A	N/A	4	3	1	2	N/A
Scheremet, 2000	3	4	6 ⁺	7	1	1	5	2
Lauer, 1999	N/A	3	2	5	4	1	N/A	N/A
Cadiou et al., 1999b	N/A	5 ⁺⁺	N/A	2	3	1	4	5 ⁺⁺

Notes: + Failure to find cointegration leads to estimation of a distributed lag model

++ The term is insignificant

It must be emphasised that the rankings presented above solely reflect the strength of the equilibrium correction process. A full account of adjustment to shocks in the economy would necessarily have to consider dynamic effects – from the dependent variable, and also from the shocked variables, should they be included in the dynamics. The actual speed of adjustment will therefore vary according to which dimension of the model is shocked. This is clearly an important issue. For instance, we simulated the effect of a 2% permanent step increase in the unemployment rate for France and Sweden. The French half-life was roughly a quarter of that displayed by Sweden⁸⁷, despite the two countries ranking similarly in the table above.

The next section reconsiders our long-run results by turning to the issue of institutional heterogeneity across countries.

⁸⁷ These results are available from the author in Excel format.

4.3 Pooled Mean Group (PMG) estimation

It is tempting when carrying out a multi-country study to pool the data in order to increase the efficiency of the estimates. This must be done with care however, as ignoring parameter heterogeneity in dynamic models would make the pooled estimator inconsistent even as the sample size T tends to infinity. The problem essentially arises because when the regressors are serially correlated, incorrectly ignoring coefficient heterogeneity results in serial correlation in the disturbance and thus inconsistent estimates in models with lagged dependent variables (Pesaran and Smith, 1995).

In light of this result, Pesaran et al. (1999) suggested a pooled mean group estimator, involving both pooling and averaging. The estimator allows the intercepts, short-run coefficients and error variances to differ freely across groups but constrains the long-run coefficients to be the same, as there are often good reasons to expect long-run equilibrium relationships to be the same across groups. The PMG estimator is consistent under its assumptions and seems to be a more robust choice than the more traditional dynamic fixed effects approach.

For our purposes, allowing heterogeneity in short-run slope coefficients has the advantage that the dynamic specification can differ across countries, which was found to be important in the earlier regressions. Moreover, we further relax the PMG assumptions by enforcing pooling only on a subset of long-run coefficients, namely those corresponding to *institutional* variables (union density, bargaining coordination and coverage, benefit duration, benefit replacement rates and employment protection legislation). As a rule, these variables were found to vary relatively little over the sample,

making identification of their effects difficult on single-country regressions, and it is here that the efficiency brought by pooling is likely to make the most difference.

We choose to implement the PMG idea by estimating our model as a nonlinear system of equations using TSP 5.0. Our aim will be to impose and test the nonlinear cross-equation restrictions implied by the hypothesis of similar institutional effects across countries. We make use of a multivariate least-squares method based on the GLS principle: the disturbances of the model are assumed to be independent across observations but to have free covariance across equations. A consistent estimate of the covariance matrix is then used to weigh the observations when the equations are re-estimated. Taking account of the correlation across equation disturbances has the potential of yielding more efficient estimates, which could already be beneficial for the identification of institutional effects. The downside of this approach is that specification errors anywhere in the system can impact upon estimators of parameters elsewhere.

We consider the following set of nested hypotheses:

1. We start from the most general specification, that is, the set of equations implied by our model augmented by the totality of the institutional variables.
2. We impose the sign priors detailed in Section 3.2. This model then becomes our maintained hypothesis, and we refer to it as the “unrestricted model” H_U .
3. We then test the actual specification implied by the single country results against this encompassing unrestricted model, H_U .

4. Using H_U as our alternative hypothesis, we test the following possible homogeneity restrictions:

- a. Homogeneity across *all* remaining institutional dimensions.
- b. Homogeneity for *each* remaining institutional dimension in turn.

In each case, the restricted model embodying the null hypothesis will be denoted by H_R .

The tests are based on the likelihood ratio statistic, which has the asymptotic distribution:

$$LR = 2\{\log L_{Unrestricted} - \log L_{Restricted}\} \xrightarrow{D_{H_0}} \chi^2(df) \quad (61)$$

The model is estimated over the period 1987:2 – 2005:2, the sample size being restricted by the inclusion of Spain in the pooling. Five Eurozone countries are included in total: France, Spain, the Netherlands, Belgium and Italy⁸⁸. Following imposition of our priors, we obtained the following model (step 2):

⁸⁸ Germany was excluded, as it would have forced a considerably smaller and different sample, from 1991q3 onwards only.

TABLE 34: MAINTAINED HYPOTHESIS, H_U

$$\begin{aligned} \text{NLD_D4I} \text{COMPTH} &= \dots [\text{Dynamics}] \dots + \text{NLDECM} * \{ \text{NLD_ECMPP}(-4) \\ &+ \text{NLDU} * \text{NLD_IU}(-4) + \text{NLDIT} * \text{NLD_IITAX}(-4) \\ &+ \text{NLDVOL} * \text{NLD_VOLAT}(-4) + \text{NLDCOORD} * \text{NLD_ICOORD}(-4) \\ &+ \text{NLDUDENS} * \text{NLD_IUDENS}(-4) + \text{NLDBRR} * \text{NLD_IBRR}(-4) \\ &+ \text{NLDA} \text{COV} * \text{NLD_IACOV}(-4) \} \\ \\ \text{ESP_D4I} \text{COMPTH} &= \dots [\text{Dynamics}] \dots + \text{ESPECM} * \{ \text{ESP_ECM}(-4) \\ &+ \text{ESPU} * \text{ESP_IU}(-4) + \text{ESPRL} * \text{ESP_IRELPMPC}(-4) \\ &+ \text{ESPU} \text{DENS} * \text{ESP_IUDENS}(-4) + \text{ESPEPL} * \text{ESP_IEPL}(-4) \\ &+ \text{ESP} \text{BRR} * \text{ESP_IBRR}(-4) \} \\ \\ \text{ITA_D8I} \text{COMPTH} &= \dots [\text{Dynamics}] \dots + \text{ITAECM} * \{ \text{ITA_ECM}(-8) \\ &+ \text{ITAU} * \text{ITA_IUNEMP}(-8) + \text{ITACORD} * \text{ITA_ICOORD}(-8) \} \\ \\ \text{BEL_D8I} \text{COMPTH} &= \dots [\text{Dynamics}] \dots + \text{BELECM} * \{ \text{BEL_ECM}(-8) \\ &+ \text{BELU} * \text{BEL_IU}(-8) + \text{BELRL} * \text{BEL_IRELPMPC}(-8) \\ &+ \text{BELUDENS} * \text{BEL_IUDENS}(-8) + \text{BELBRR} * \text{BEL_IBRR}(-8) \\ &+ \text{BELCOORD} * \text{BEL_ICOORD}(-8) \} \\ \\ \text{FRA_DI} \text{COMPTH} &= \dots [\text{Dynamics}] \dots + \text{FRAECM} * \{ \text{FRA_IRCOMPTH}(-3) \\ &+ \text{FRAPRODIT} * \text{FRA_IPRODIT}(-3) + \text{FRAU} * \text{FRA_IU}(-3) \\ &+ \text{FRAUDENS} * \text{FRA_IUDENS}(-3) + \text{FRABD} * \text{FRA_IBD}(-3) \\ &+ \text{FRABRR} * \text{FRA_IBRR}(-3) + \text{FRAACOV} * \text{FRA_IACOV}(-3) \} \end{aligned} \tag{62}$$

The H_U model has equation standard errors and coefficients which are overall well in line with the results obtained earlier in the chapter. Moreover, it fully nests the specification implied by our country results.

The following table reports the results of our nested tests:

TABLE 35: NESTED TEST RESULTS⁸⁹

STEP 3: H_U vs. specification adapted from country regressions (H_R)

$$LR = 23.59 > \chi^2(6)_{critical,0.1\%} = 22.46 \Rightarrow \text{Strongly reject } H_R$$

STEP 3i: H_U vs. specification adapted from country regressions (H_R), including BRR in the Dutch wage equation.

$$LR = 4.16 < \chi^2(5)_{critical,5\%} = 11.07 \Rightarrow \text{Do not reject } H_R$$

STEP 4a: H_U vs. all homogeneity restrictions (H_R)

$$LR = 691 >> \chi^2(9)_{critical,0.1\%} = 27.88 \Rightarrow \text{Strongly reject } H_R$$

STEP 4bi: H_U vs. homogeneity for COORD (H_R)

$$LR = 0.06 < \chi^2(2)_{critical,5\%} = 5.99 \Rightarrow \text{Do not reject } H_R$$

STEP 4biii: H_U vs. homogeneity for UDENS (H_R)

$$LR = 822 >> \chi^2(3)_{critical,0.1\%} = 16.27 \Rightarrow \text{Strongly reject } H_R$$

STEP 4biii: H_U vs. homogeneity for BRR (H_R)

$$LR = 878 >> \chi^2(3)_{critical,0.1\%} = 16.27 \Rightarrow \text{Strongly reject } H_R$$

STEP 4biv: H_U vs. homogeneity for ACOV (H_R)

$$LR = 309 >> \chi^2(1)_{critical,0.1\%} = 10.83 \Rightarrow \text{Strongly reject } H_R$$

STEP 4bv: H_U vs. “Benelux” homogeneity for COORD, BRR, UDENS (H_R)

$$LR = 598 >> \chi^2(3)_{critical,0.1\%} = 16.27 \Rightarrow \text{Strongly reject } H_R$$

⁸⁹ Chi square critical values are taken from tables compiled by Christopher Dougherty, LSE.

Note that since the employment protection legislation and benefit duration indicators only concerned individual countries once the sign priors were imposed, they were naturally excluded from the tests in Step 4.

As can be seen from the results of the test in Step 3, the specification derived from our single country equations is rejected against the encompassing model H_U . This, however, is solely due to the absence of the benefit replacement rate indicator (BRR) from the Dutch model presented in Section 4.1.6. Step 3i shows that when it is included in the restricted model H_R , the null hypothesis can be accepted, thereby demonstrating that it represents the only important difference relative to our single country results. The rejection in Step 3 is not surprising, given the importance of the difference between the sample used here (from 1987q2 onwards) and the sample used earlier (from 1974q3 onwards) for the Netherlands. On the larger sample, the BRR indicator appears to be insignificant; a detailed examination of its recursive coefficient shows it to be reliably insignificant, although it does appear to rise following 2001 and this may be reflected in the smaller sample by a finding of significance. The efficiency added by the approach taken here could also be credited with this finding. Overall however, the results here are surprisingly well-aligned with those obtained earlier in light of the sample differences at hand, thus providing added support for our selected specifications.

The most notable conclusion to be drawn from Table 35 is that almost all of the restrictions considered are rejected. Note that even an *a priori* relatively mild constraint across the equations for the two Benelux countries represented here (Step 4bv) was strongly rejected. In fact, the *only* acceptable homogeneity restriction is the one corresponding to the bargaining coordination indicator. This provides added support for

the identical result reached in Curto Millet (2004). The finding suggests that in the European countries where bargaining coordination has a significant effect on wage outcomes (Belgium, the Netherlands and Italy), the influence exerted is similar. All three countries display highly coordinated labour markets, as reflected by our index. Their *pattern* of coordination is also similar, and can be characterised as “intra-associational coordination” and/or “state-sponsored coordination” over the sample period we consider (see Section 3.6.3 and Ochel, 2000, for details). This finding may have implications for the feasibility of European wage coordination (e.g. Borghijs et al., 2003).

We therefore reach a clear conclusion. First, the restrictions we imposed in the country equations are vindicated by the tests from the generalised model in this subsample. Second, the PMG principle is overall clearly invalid here in terms of bias, and our results suggest care must be taken before its application to empirical problems. It may be more readily justified in cases where budget, solvency constraints or common technologies influence all groups in the same way, for instance. Clearly, it is a problematic concept with respect to institutions, and our results confirm the importance of national idiosyncrasies.

Two approaches (alone or in combination) could be followed at this point for a better identification of institutional effects. One would be to assemble a larger database in the time series dimension. The main restriction we faced here was due to the length of our inflation expectations data. Extending this rigorously would require further research into realistic expectations formation mechanisms. This issue is touched upon in Chapter 3.

The other approach involves making use of the cross-sectional variation as well, but justifying the restrictions on a weaker MSE basis, de facto trading-off some bias against the gain in efficiency. A principled way to implement this is the poolability test suggested by McElroy (1977) and Baltagi (2001), as deployed by Nunziata (2001a).

Detailed research into institutional effects is a field in its own right. Given the results reached in this section we will not push this any further, but are satisfied that our specification was appropriate for our main purpose.

5 CONCLUSION

5.1 Summary of key points

This chapter has investigated the hypothesis that the formation of the European Monetary Union might have led to an increase in the ‘flexibility’ of its participating labour markets. The theoretical basis for such a claim was established in Section 2, which also surveyed the related empirical literature concerning the EMS, as well as more recent studies. Neither the theory nor past empirical evidence were strong enough to allow much confidence in an ex-ante null hypothesis. This was partly due to the shortcomings of previous empirical work, especially with regards to the treatment of expectations.

The model deployed here had an equilibrium correction form, with the long-run solution provided by a bargaining model à la Layard et al. (1991) and dynamics consistent with a staggered contracts approach. Inflation expectations were taken into account by using the measure developed in Chapter 1, and institutional effects were carefully controlled for through a database on relevant dimensions constructed for this study.

The resulting wage equations proved notably successful at reflecting the institutional knowledge available for our sample of countries – for instance, the three-year collective bargaining dynamics of Sweden or the effects of the dismantling of Italy’s *Scala Mobile*. Overall, no convincing evidence of EMU or EMS flexibility effects relating to the long-run unemployment coefficient of our wage equations could be found. However, the prospect of EMU membership may have been helpful in other respects; these include recent wage moderation in Belgium and the overall strong productivity orientation of the Eurozone bargaining equilibrium, notably for Italy and Spain. Nevertheless, the equations also point to a cost of EMU, albeit possibly a short-term one: the significant role of inflation expectations, combined with their build-up in recent years *associated with monetary union* (see Chapter 3) suggest that this variable has to some extent contributed to generating wage inflation relative to the counterfactual situation.

Finally, our attempt to better identify institutional effects by imposing pooled mean group restrictions made clear that this would have to be at the cost of bias. Institutional heterogeneity across countries is thus too important to ignore.

We next provide some additional interpretation to understand the implications of the wage flexibility findings reached here.

5.2 Looking ahead: implications for EMU

Explaining the absence of an ‘EMU effect’ on flexibility

The key finding of this study – namely the absence of an ‘EMU effect’ in the Eurozone labour markets considered – can be rather disquieting. There are essentially three possible explanations for this finding:

1. An unchanged level of macroeconomic variability

Essentially, all of the theoretical mechanisms discussed in Section 2 that could lead one to expect higher flexibility in the face of monetary union are set in motion by increased macroeconomic variability⁹⁰. This is the case, for instance, of the endogenous selection of contract length. An interpretation of our findings, then, is that macroeconomic variability in Europe has been unchanged since the introduction of EMU. This might appear a plausible explanation in the historical context of the ‘Great Moderation’, i.e. the decrease in macroeconomic volatility in recent years, that has reduced the need for either interest rate or wage flexibility in order to eliminate fluctuations in unemployment.

A variation on this theme is that the current Eurozone countries have been so closely linked over the past couple of decades (e.g. EMS, Maastricht Convergence Criteria) that any flexibility improvements that could be achieved through labour markets have already been so. This additional point is however more applicable to countries such as the Netherlands than Spain or Italy, due to the stronger DM and general economic ties with Germany of the former.

⁹⁰ Or a perception by agents of increased volatility, absent wage adjustment.

2. Weak endogenous response of the wage process to macroeconomic variability

In contrast to the first explanation, this view holds that macroeconomic variability has increased, but that the theoretical mechanisms posited for a flexibility effect are too weak to have much of an impact – at least in the time frame considered.

3. Flexibility has increased along unmeasured dimensions?

This study has focused on wage flexibility. Other labour market dimensions can however also be associated with greater ‘flexibility’, including labour mobility and working time flexibility. The argument here would have to be that these dimensions have shifted in a way that goes unmeasured through our coefficients. A working time reorganisation could for instance improve productivity and help maintain competitiveness in the face of a demand shock.

There are several problems with this last story, however. First, in most instances such shifts would have had some effect on our equations. For example, increased labour mobility would tend to reduce mismatch and increase the magnitude of the coefficient on the unemployment level. No such changes have been observed. Second, increases in flexibility through wages have more plausibility than other possibilities as a response to EMU. Wages are the one component of the labour relationship that is always present in regular bargaining, arguably making such changes easier to implement. In any event, for practical purposes, other flexibilisation paths are likely to have limits that would eventually require improvements in wage flexibility in the face of shocks. Nevertheless – and subject to the above caveats – it is conceptually possible that flexibility has increased

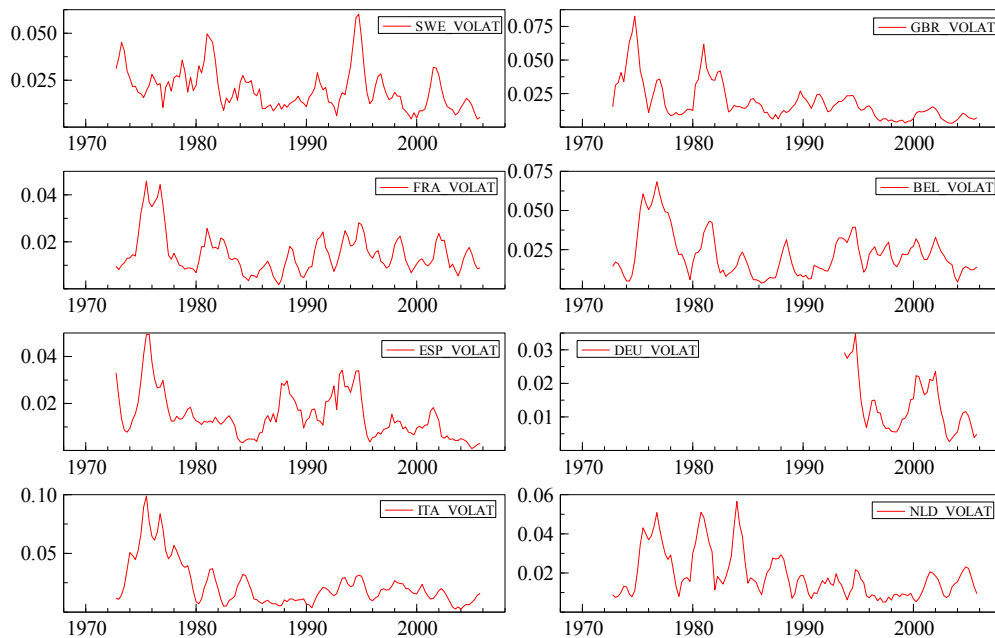
through variables such as non-wage costs of labour. The issue is deserving of further investigation.

Discriminating between explanations 1 and 2 is not straightforward and depends on the assessment one makes of recent changes in macroeconomic volatility. The literature typically measures ‘macroeconomic volatility’ in terms of the standard deviation of real GDP growth, possibly measured per capita (e.g. Kose et al., 2003; Roskelley, 2003; Breen and García-Peñalosa, 2005). We choose a related measure of volatility here, namely the four-quarter moving average of the difference between GDP growth this year and the growth observed a year ago:

$$MA[abs(D4IGDP - D4IGDP_{-4})]$$

Taking the absolute value avoids putting as large a weight on outliers as a rolling standard deviation measure would. The results for all the countries under consideration are shown in Figure 75.

FIGURE 75: MACROECONOMIC VOLATILITY INDEX, ALL COUNTRIES



As can be seen, there appears to be little evidence of an increase in macroeconomic volatility on this count. The overall impression is instead one of stability or even decreasing volatility for this measure, which would tend to favour hypothesis 1 over 2. This is in tune with arguments according to which the ECB has performed better than it is often credited with in the stabilisation of the Eurozone (e.g. Alesina et al., 2001). In any event, it is certainly true that EMU conditions had been relatively well ‘simulated’ under the EMS and pre-EMU convergence process.

Such a conclusion would however rely rather heavily on the particular measure of volatility used in Figure 75. Indeed, it is doubtful to what extent hypothesis 1 can constitute the whole story given the consistency of our results across countries and the variety of macroeconomic conditions Eurozone members experienced – such as Germany’s recession or the strong growth experienced by Spain and Ireland, which could

have been good candidates for a national monetary response. The evidence presented here therefore also reinforces earlier empirical results in the literature by pointing out the practical weakness of the theoretical linkages between labour market *flexibility* and monetary commitments.

Overall, the full explanation is likely to be somewhere in between hypotheses 1 & 2 (with perhaps a touch of 3), and possibly differ across countries. Looking ahead, the uncomfortable conclusion is however that unless there is an (undesirable) increase in volatility in the forthcoming years, little additional support for stabilisation can be expected from the labour market.

Implications for fiscal policy

Our findings underline the importance of fiscal stabilisation under EMU as the other remaining adjustment mechanism with some relevance. The economic literature has emphasised many difficulties here, including arguments of Ricardian equivalence, political economy considerations and empirically, the size of the tax multipliers (Blanchard and Perotti, 1999). Policy discussions also routinely highlight the constraints set by the Stability and Growth Pact.

These issues are well-known and will not be expanded upon here. There is however much scarcer discussion of the possible use of fiscal policy to affect relative prices, which could in a sense achieve the increase in flexibility that has been found missing here. One possibility would be temporary changes in the VAT to encourage households to reallocate spending inter-temporally (Commission on Stabilisation Policy, 2002). This

however runs the risk of cost-push inflationary complications when trying to dampen a boom. The results reached here were not especially encouraging in this respect, as only the long-run component of our Dutch and German wage equations could potentially be expressed in terms of producer prices – although this observation must be tempered with the identification caveat made in Section 4.2. In any event, these are equilibrium results, and one ought to be even more cautious about the implications for short-run dynamics of such policies.

We note however that this is by no means the only possibility. For instance, Calmfors (2002) discusses the possible use of payroll taxes as a stabilisation tool. The Finnish system of ‘EMU buffer funds’ mentioned earlier is close to the spirit of this proposal.

CHAPTER III:

Properties and Applications of Data on Inflation Perceptions and Expectations

1 INTRODUCTION

Chapter 1 dealt with the selection of the optimal method for quantifying qualitative data on inflation expectations. Chapter 2, in turn, used the inflation expectations data derived from the recommended method to contribute to the modelling of wages in a set of European countries. The present chapter now aims to go further by considering the properties of inflation expectations and perceptions *themselves*.

Countless theoretical and empirical macroeconomic studies rely on some posited process of expectations formation. The examples are literally too numerous to count: a list including credibility theory (Kydland and Prescott, 1977; Barro and Gordon, 1983), Keynesian and new Keynesian Phillips curves (e.g. Galí and Gertler, 1999), political business cycle theory (Nordhaus, 1975) or rational partisan theory (Alesina, 1987) would only constitute the tip of the iceberg, at best. Preciously little work, however, has been invested in in-depth empirical studies to ascertain what robust knowledge of the properties of inflation expectations can be gleaned from the data. Given the frequently crucial sensitivity of the outcomes of interest to the form of expectations assumed, this state of affairs appears somewhat disquieting, to say the least. The only exception to this statement is perhaps the truly industrial levels of output generated over the past 30 years by papers aiming to test the rational expectations hypothesis on survey data. Some useful conclusions have been reached (mainly, the rejection of strong versions of the hypothesis) despite the widespread use of inappropriate econometrics (see Curto Millet (2006) for a critique). Notwithstanding this, that research project has long entered its area of diminishing returns.

In contrast, the surface of most of the interesting empirical work has barely been scratched. Yet, the present availability of large samples of high-quality data on inflation expectations truly offers a “gold mine” of research opportunities. Accordingly, the object of this chapter is to check the robustness of selected pioneering contributions in this area and explore new avenues that have as yet been left untouched. The structure of the chapter reflects this exploratory philosophy by being divided into six related – but reasonably self-contained – sections.

Section 2 analyses in detail the data on inflation expectations and perceptions for the eight European countries in our sample. Particular attention is paid to the impact of the introduction of the euro on these variables and on their relationship. Inflation perceptions are studied to derive an understanding of the price index that is considered relevant by consumers, and a test of the “rational perceptions hypothesis” is carried out for the first time in the literature. The implications of these findings for the rational expectations hypothesis and the quantification of qualitative survey data are also explored.

Section 3 provides an empirical test of the aggregate hypothesis of “near-rationality” formulated by Akerlof et al. (2000). These authors posited that the population is divided between “near rational” and rational individuals, the proportions of which depend on the level of inflation. This hypothesis has momentous implications, as it provides an argument for maintaining a certain positive level of inflation in order to minimise unemployment in the economy.

Section 4 presents an overview of the development of the foundational expectations processes that have constituted key building blocks for models across the field of macroeconomics. These are the adaptive, rational and rule-of-thumb expectations models. A test of adaptive expectations is provided as a benchmark, before investing particular care in checking the robustness of past empirical work on the now-ubiquitous rule-of-thumb hypothesis for expectations.

Section 5 considers the relationship between consumer expectations and the forecasts of professional forecasters empirically. The forecasting ability of experts and consumers is contrasted, and Granger causality between the two sets of expectations is explored. A test of Carroll's (2003) version of the "sticky information" model of expectations is provided, and the role of mass media in the diffusion of expert forecasts to the general public is analysed in the context of the United Kingdom.

Section 6 then turns to the implications of the *distributional* aspects of data on inflation expectations. Mankiw and Reis' (2002) "sticky information" model is presented and its predictions for the evolution of the expectations distribution are tested in the context of the UK's ERM disinflation and the US's Volcker disinflation. Section 6.3 then considers the merits of an argument made by van Lelyveld (2000), claiming that the formulation of the EC Consumer Survey questions *directly implies* bi-modality in the distribution of expectations. Section 6.4 examines the explanatory power of a number of *a priori* sensible variables for the interquantile range of the expectations distribution, while Section 6.5 checks whether the level of "disagreement" about expectations present in qualitative survey data adequately reflects that in quantitative survey data. Finally,

Section 6.6 studies in some detail the qualitative features of the expectations distribution and reaches an assessment regarding the comparative advantages of quantitative and qualitative surveys.

Section 7 provides some concluding observations.

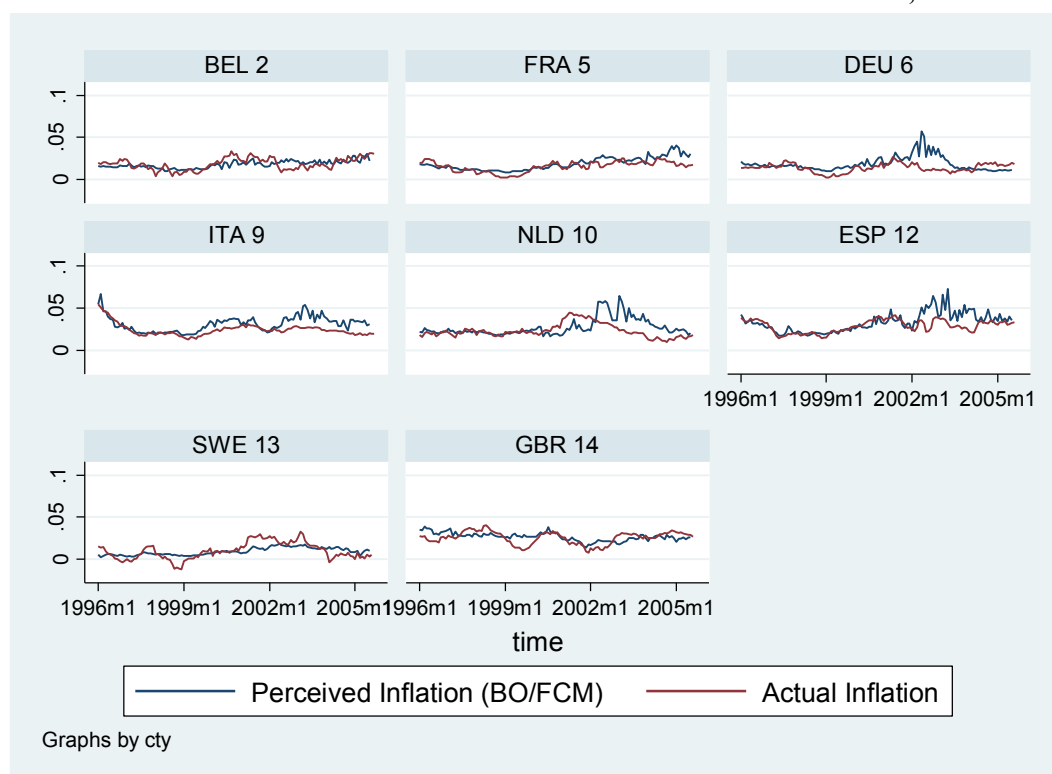
2 UNDERSTANDING INFLATION EXPECTATIONS AND PERCEPTIONS

2.1 Descriptive analysis of our inflation expectations and perceptions series

2.1.1 Inflation perceptions

Figure 76 plots monthly data on inflation perceptions against the actual CPI inflation for all the countries in our sample post-1996.

FIGURE 76: ACTUAL INFLATION & PERCEPTIONS BY COUNTRY, POST-1996

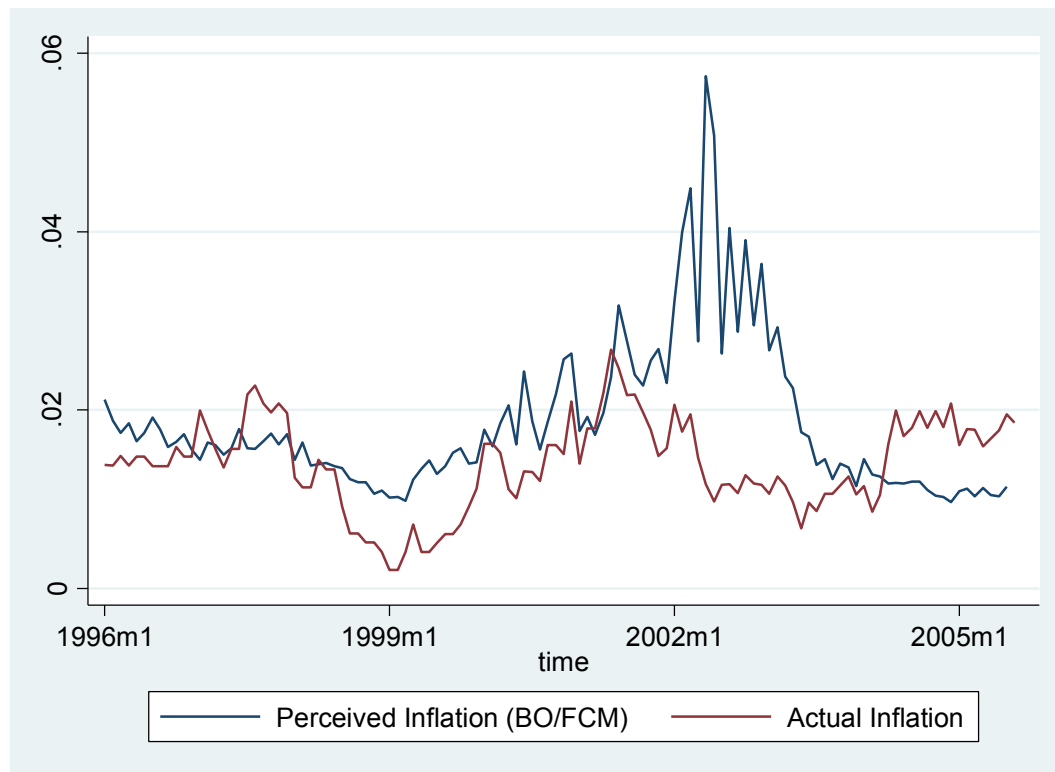


A remarkable feature of the data is the large deviations observed – even at this scale – between inflation perceptions and actual inflation in the post-euro period for the Eurozone countries. These are very large and visible for Germany, the Netherlands, Italy and Spain; a smaller version of the phenomenon can arguably be observed for France and Belgium. The series for the UK and Sweden, on the other hand, do not display such features. Thus, it appears that in recent years inflation perceptions in Eurozone countries

have ‘exploded’ in a way that is not (fully) justified by the underlying pattern of inflation, thereby breaking the close relationship between the two series that had been observed heretofore.

We consider in greater detail the case of Germany in Figure 77, as it is fully representative of the observed behaviour in several Eurozone countries (Italy and the Netherlands in our sample; equivalent plots are available for all the other countries in Appendix 1).

FIGURE 77: ACTUAL INFLATION & PERCEPTIONS, GERMANY, POST-1996

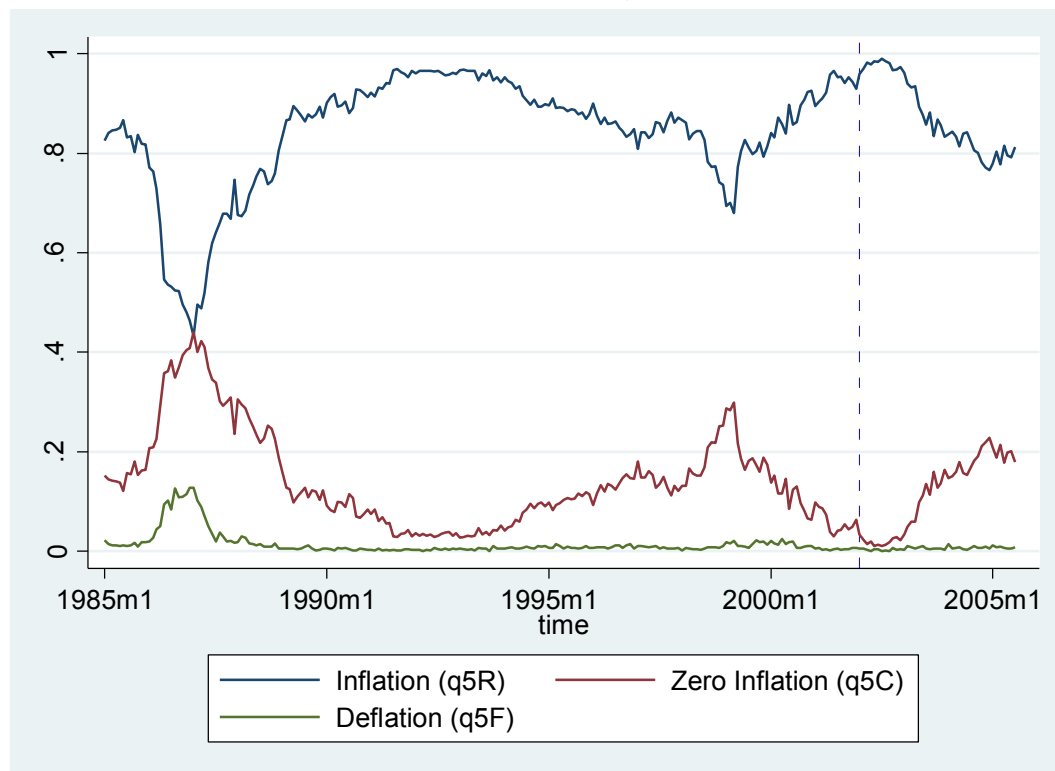


The deviation between the two series in the post-EMU period is rather large, with perceptions indicating inflation of around 3-4% (even peaking in one particular month at close to 6%) whereas actual inflation hovers just below 2%. It is remarkable that this

development is not visible in any of the EU-15 countries having opted out from the single currency so far (the UK, Sweden and Denmark⁹¹).

One might indeed have expected that the EMU period would be associated with perceptions of higher inflation (*pace* credibility theory). A reason for this was the number of news stories speculating that the conversion from the national monies to the euro would be associated with greater inflation as manufacturers and retailers took advantage of the opportunity to adjust their prices upwards. A close examination of the underlying data on response shares is supportive of this hypothesis.

FIGURE 78: GERMANY, RESPONSE SHARES (PERCEIVED INFLATION QUESTION)



⁹¹ Please refer to Curto Millet (2004) for plots of inflation perceptions and expectations for Denmark

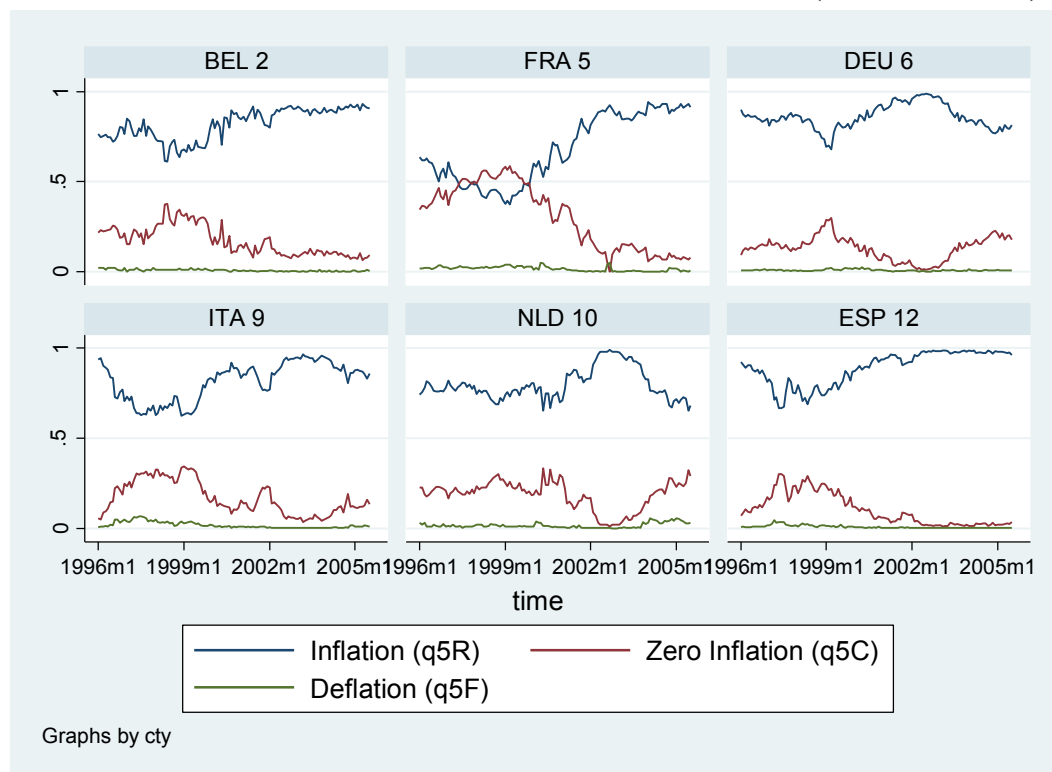
As can be seen, the post-EMU period has been characterised by a collapse of the share of respondents perceiving constant prices, to the benefit of the share perceiving positive inflation. In the case of Germany, this process started in March 1999. The shift towards a perception of inflation is so great that to accommodate the small percentages choosing the other two options, the distribution of inflation perceptions has to flatten considerably, leading to the observed increase in mean inflation perceptions. The process slows in early 2002 (after the introduction of euro notes and coins in January) and reaches its peak in May 2002 in Germany, after which it gradually starts to reverse, as reflected in the series on inflation perceptions. We note that this reversal takes place soon after the time when the euro ceased to be an effective legal tender in Germany, around late February 2002⁹².

Thus, it appears that German inflation perceptions can be characterised as having been pushed upwards by EMU, especially around the time of the introduction of the euro notes and coins; realising that their full inflationary fears had not been met, German consumers revised their perceptions downwards relatively rapidly following the euro transition process.

Having examined the German case in detail, Figure 79 presents the response shares for the question on perceptions for all Eurozone countries in our sample.

⁹² The end of legal tender of the national banknotes and coins was set to 28 February 2002 in almost all Eurozone countries. There were exceptions. Germany scheduled it for 31 December 2001; however, as was announced in the joint statement of professional associations of 22 October 1998, businesses decided to accept Deutsche Mark banknotes and coins at least until 28 February 2002. The other exceptions were: France (17/02/2002); Ireland (09/02/2002) and the Netherlands (28/01/2002).

FIGURE 79: EUROZONE COUNTRIES, RESPONSE SHARES (PERCEPTIONS)



There are clearly a number of elements in common in the behaviour of the response shares across the Eurozone. In every single country, the process at hand is characterised by a significant fall of the share of respondents perceiving constant prices, to the benefit of the share perceiving positive inflation in the post-1999 years. The differences only appear around the middle of 2002.

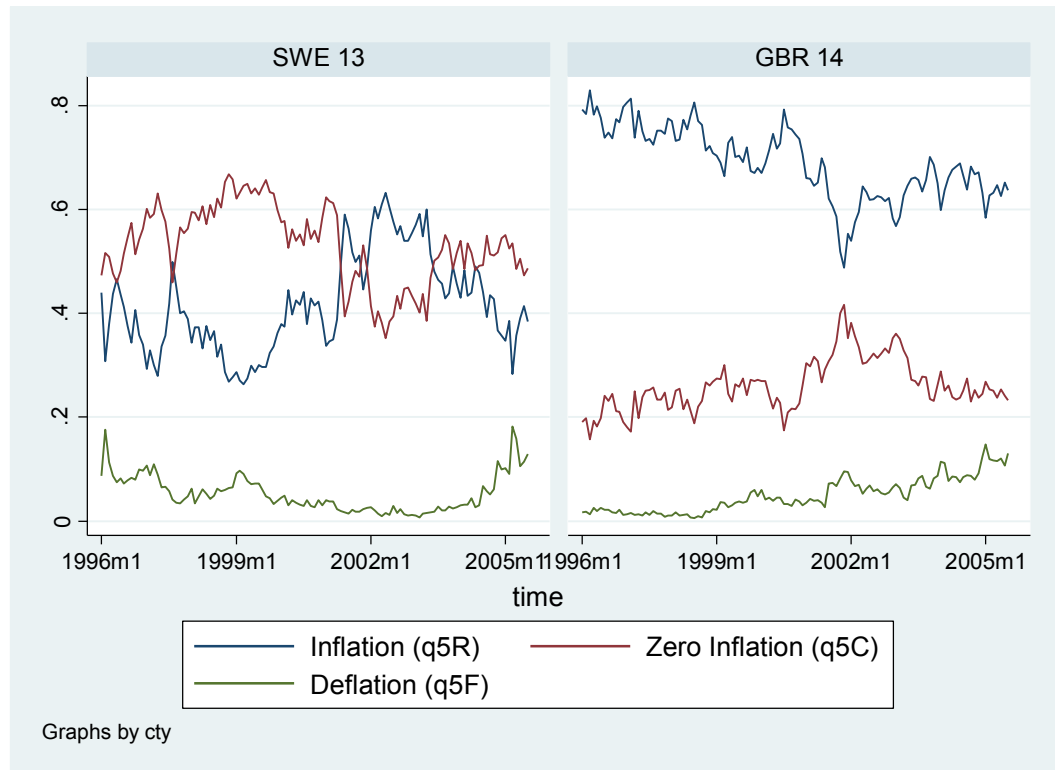
The Netherlands and Italy broadly follow the “German model”. Indeed, it is straightforward to see that the shapes traced by their response shares have the same characteristics as those of their German counterparts. The only exception is the presence of a considerably more marked “reassuring perceptions” effect in the two countries in late 2001, especially in Italy – with perceptions of inflation falling in favour of perceptions of price constancy, thereby partly reversing the previous build-up of

inflationary perceptions. This coincided with a moderation of inflation in the months between July and November/December of 2001. The effect however disappeared with the introduction of euro notes and coins, and both countries then reverted to the German pattern.

Belgium, France and Spain present similar behaviour until 2002. They each have a small “reassuring perceptions” effect in the months prior to 2002 and then revert to the “German model”. However, in these cases the diverging trend between the shares perceiving inflation and constant prices fails to reverse, although it is greatly moderated and, in the case of Spain effectively stopped in 2002. In these countries, inflationary perceptions linked with the introduction of euro notes and coins were reinforced by increasing inflation in the immediate period after the euro changeover, as opposed to the situation in the other Eurozone countries in our sample. Indeed, Figure 76 shows that the ‘high’ inflation perceptions in these countries were (eventually) vindicated by actual inflation.

For completeness, we plot the response shares for Sweden and the UK in Figure 80; these contrast with the earlier series in that they naturally fail to show any particular connection with the timeline of EMU.

FIGURE 80: UK AND SWEDEN, RESPONSE SHARES (INFLATION PERCEPTIONS)



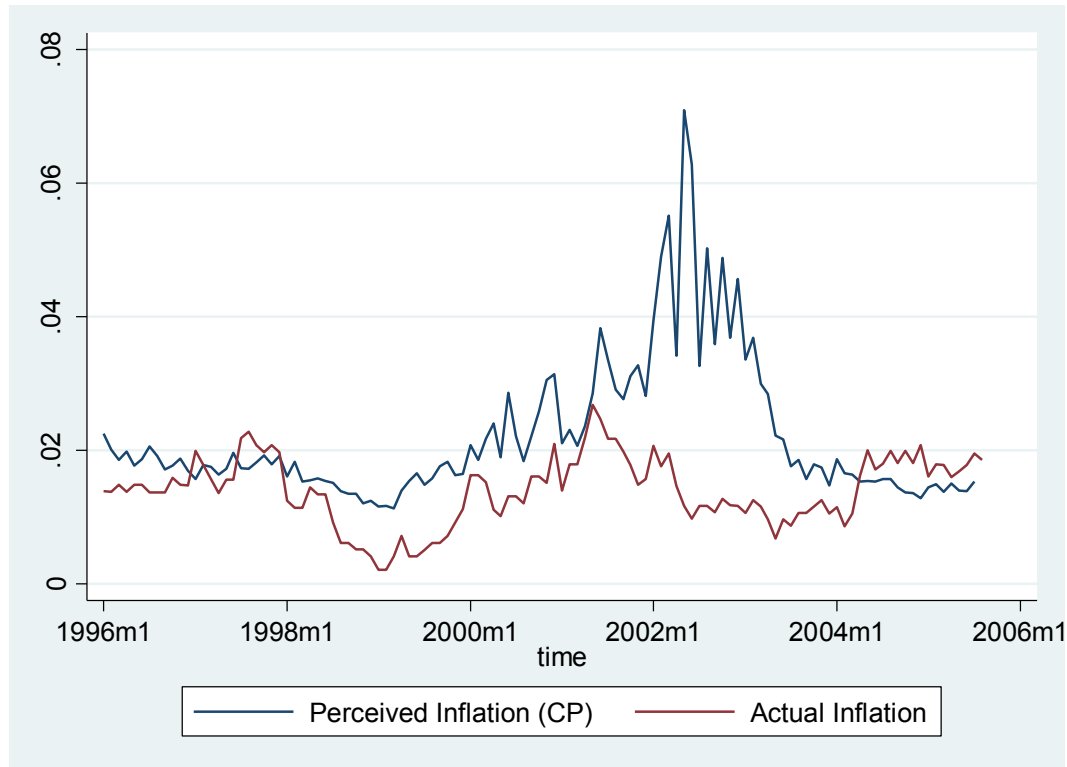
2.1.2 An additional justification for our quantification method

As shown in the previous sub-section, the introduction of the euro was generally associated with a collapse in the proportion of respondents perceiving zero inflation in favour of that perceiving positive inflation. This has translated into large deviations of inflation perceptions from the underlying inflation data.

The quantification method we have employed for our qualitative data allows for long-term trends in the indifference thresholds of the distribution (see Chapter 1, Section 3.1). How would inflation perceptions have appeared had we imposed constant indifference thresholds? This would have been the case under the standard Carlson-Parkin method (as in Berk, 1999), without the modification proposed in Curto Millet (2004) and

recommended in Chapter 1. Inflation perceptions for Germany would then appear as follows:

FIGURE 81: GERMAN INFLATION PERCEPTIONS POST-1996, CP METHOD



As can be seen by contrasting Figure 81 and Figure 77, the deviations of perceptions from actual inflation behaviour are considerably more marked under this quantification methodology. Curto Millet (2004), examining the phenomenon for a number of countries, argued that the deviations displayed by these constant-threshold quantification techniques were qualitatively correct but possibly quantitatively exaggerated.

The observed generalised breakdown of the close relationship between perceptions and actual inflation in the euro era leads to the consideration of two potential hypotheses.

One interpretation of the phenomenon would be to question the assumption of normality for our distributions. Indeed, it might be argued that the actual distribution of expectations might be more appropriately described in this period by a distribution skewed to the right. This is not without complications. First is the question of which distribution to select (e.g. chi-square, noncentral t-distribution etc.). Whereas we had the guidance of the Central Limit Theorem to point us to the normal distribution, the choice of alternatives seems unfortunately more ad hoc. Second, given that the normal distribution appears to be adequate for the earlier parts of the sample, would this require a complicated transition between two distributions? Therefore, this approach to the problem appears impractical and is in any event questioned by the reasonable success achieved by the normal distribution in past empirical research (e.g. Nielsen, 2003a; Berk, 1999).

Another interpretation of the observed outcomes, however, is that the response thresholds have shifted inwards over the EMU period, so that it takes a smaller absolute value of perceived inflation or deflation in the respondents' heads to call it as such, thereby requiring a lesser flattening of the distribution in response to the documented changes in inflation perceptions. Taking this path has the advantage of avoiding complicated distributional issues but requires some judgement. Clearly, one could replicate the actual inflation series arbitrarily well by varying the thresholds continuously in an appropriate manner. The aim, rather, is to allow for long-run changes in these parameters without risking the imposition of excessive structure. Hence the choice of applying a criterion of unbiasedness for each half of the sample separately to quantify the perceptions underlying our EXPINFLPBIS measure. The resulting series preserves all the main

qualitative features of the data whilst arguably appearing more reasonable from a quantitative perspective. Ultimately, an endorsement of the approach can only be provided by an empirical test; this was passed successfully in Chapter 1.

2.1.3 Inflation expectations

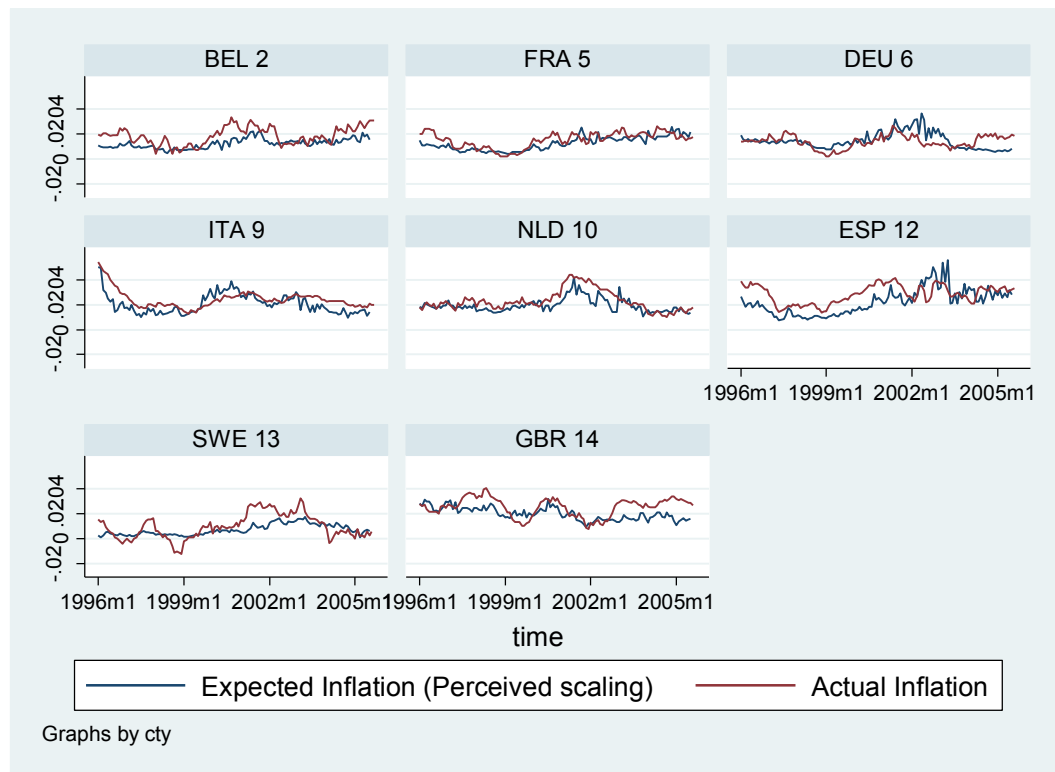
After studying the behaviour of inflation perceptions, we now turn our attention to inflation expectations.

Figure 82 plots inflation expectations *at the time they were surveyed*⁹³ against actual inflation in the post-1996 period. Individual plots for each country are available in Appendix 1.

A first observation is that inflation expectations seem to have increased, at least temporarily, in most Eurozone countries with the introduction of EMU, a fact of relevance to the interpretation of our wage equations, as was already noted in Chapter 2 (see Section 5.1).

⁹³ As opposed to the time the expectations *refer to* (i.e. over the *coming* 12 months).

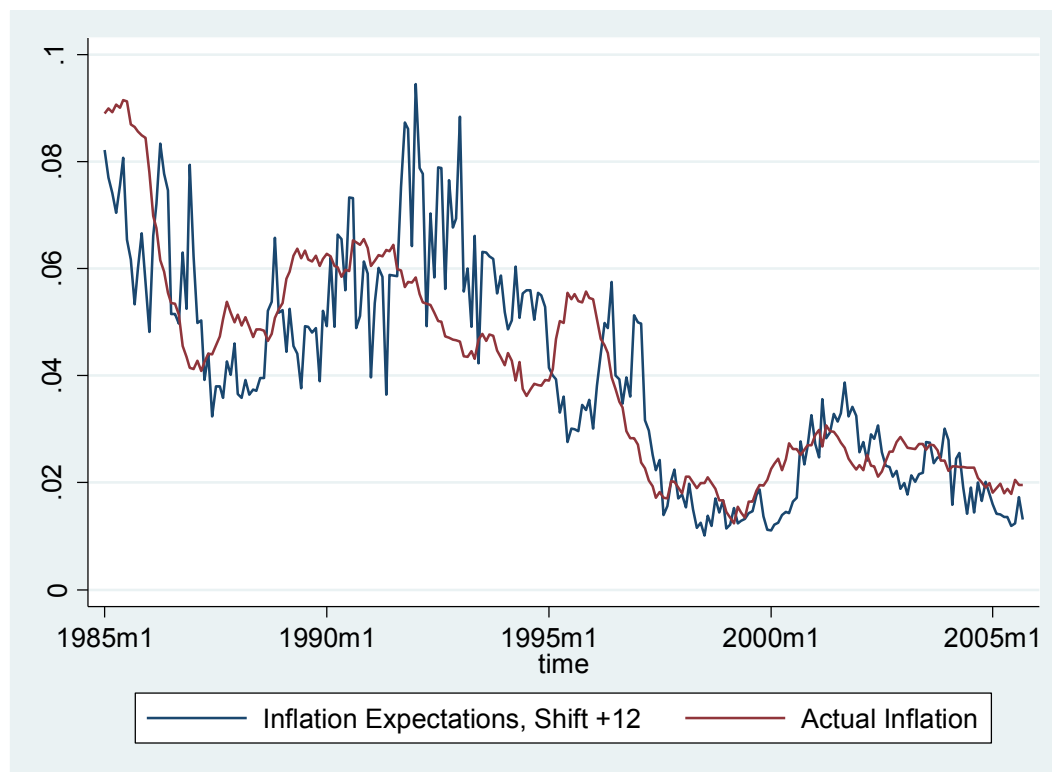
FIGURE 82: ACTUAL INFLATION & EXPECTATIONS BY COUNTRY, POST-1996



Looking closer, most expectations series appear to follow rather closely actual inflation developments, suggesting that they are mostly adaptive in nature. However, it is equally clear that they are not *only* adaptive, deviating in a way that is complex and time-varying. Large systematic gaps between the series are often evident. In some cases, these translate a correct anticipation of inflationary outturns, as is visible when the series is shifted by 12 months so as to directly compare expectations and realisations⁹⁴. This is done in Figure 83 for Italy, where it can be seen that the series anticipates reasonably well the decline in inflation in the late 1990s and its subsequent increase.

⁹⁴ Curto Millet (2004) noted that Denmark appeared to be a special case in this respect, as an undershooting of actual outturns was observed for the entire sample period. It was suggested that this may be a case in which the OECD CPI measure is rather different from the one people have in mind when answering the survey questions – for instance, people may disregard a component of the index that is quite important. Readers are referred to the discussion there and that in Section 2.3 below.

FIGURE 83: ITALY, INFLATION EXPECTATIONS VS ACTUAL OUTTURNS



The suggestive evidence presented here thus joins the results of previous research in finding that overall, although highly adaptive, expectations are not simple extrapolations (e.g. Roberts, 1997; Forsells and Kenny, 2006). The high apparent weight attached to the adaptive “component” was moreover only to be expected given the population surveyed (consumers at large) with no “inside” economic information. We note that the degree of “forward-lookingness” seems to vary across countries, but that the observations made here broadly hold for all countries.

Figure 84 plots the response shares underlying the data to highlight a crucial pattern. Response shares linked to an expectation of positive inflation are aggregated in the figure for improved visibility.

It is notable that in all Eurozone countries⁹⁵ the introduction of the euro has been initially accompanied by often sharp increases in the share of respondents *expecting* inflation, at the expense of those *expecting* zero inflation (see Section 2.1.1 for our observations on the behaviour of inflation *perceptions*). Even more remarkable is the fact that this development has largely and quickly been reversed in most countries around January 2002. The change is particularly sharp for France, Belgium and the Netherlands, and occurs *exactly* in January 2002. For Germany, the reversal starts very slowly in November 2001 and accelerates in January 2002. For Italy, the change starts in February 2002. Spain is a different case, and although a dampened version of the phenomenon visible for other countries does take place in February 2002, we have to wait until March 2003 for a large-scale reversal in expectations to take place.

It would therefore seem that in the period following the ‘virtual’ introduction of the euro in 1999, a build-up of inflationary expectations took place. These expectations were supported and fed by the simultaneously increasing inflation across the area. The associated inflationary fears appear to have been rapidly punctured by the introduction of the physical currency in all countries but Spain. This may be due to the volatility of Spanish inflation over 2002, with the country having to wait until 2003 for a significant sustained decrease in the inflation rate (with an important fall between March and April of that year).

⁹⁵ See Curto Millet (2004) for the evidence concerning Eurozone countries not covered here.

FIGURE 84: EXPECTED INFLATION, RESPONSE SHARES BY EUZ COUNTRY

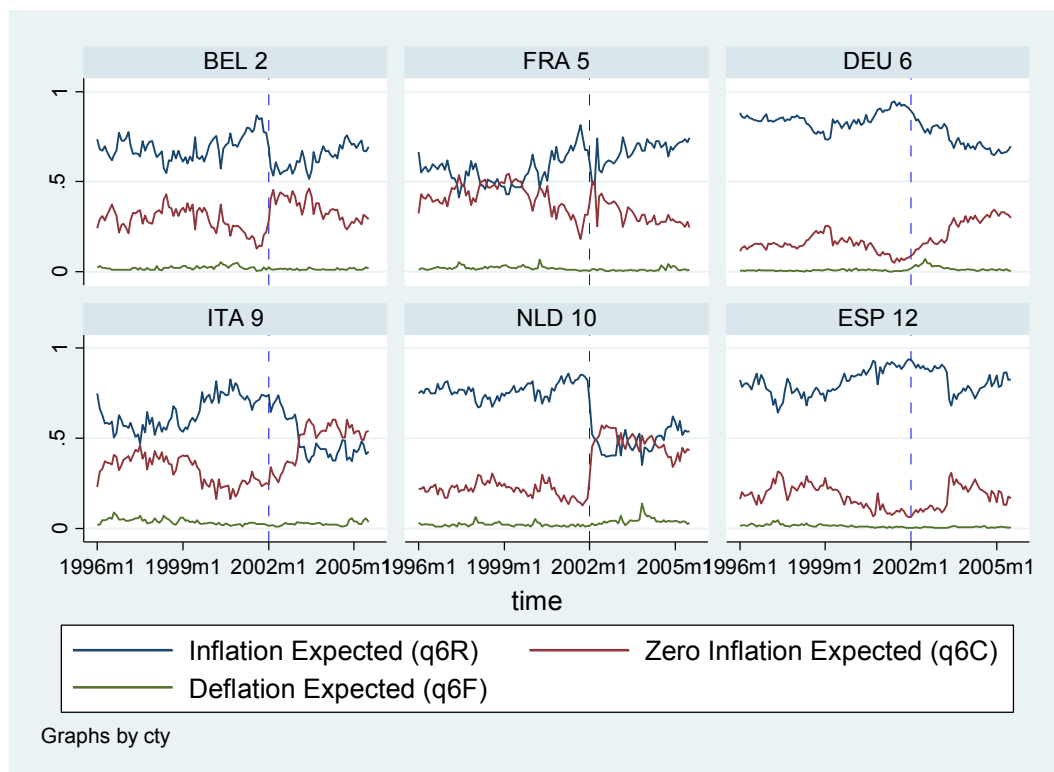
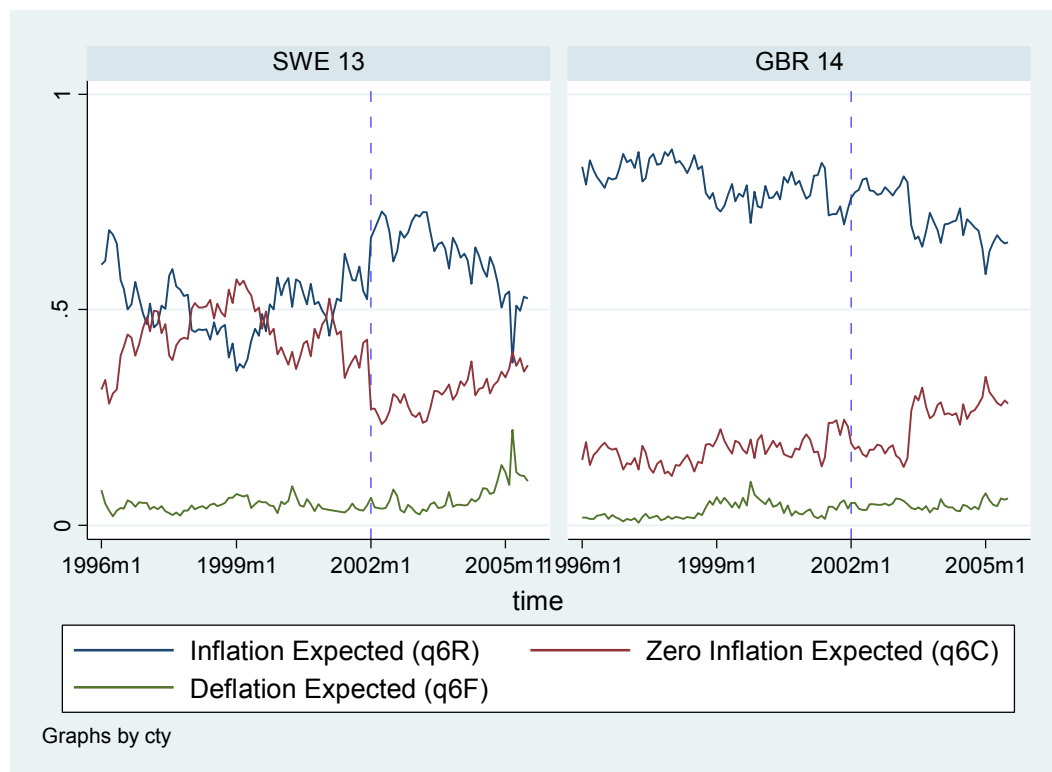


Figure 85 presents equivalent plots for the UK and Sweden for comparison. The pattern once again appears very different from that of their Eurozone counterparts. Sweden experienced the *opposite* of the other countries in January 2002, that is, an increase in the share of respondents expecting inflation and a corresponding decrease in that expecting constant prices. The UK reveals no particular pattern.

FIGURE 85: EXPECTED INFLATION, RESPONSE SHARES, SWEDEN AND UK



2.2 Understanding the link between inflation expectations and perceptions

We now turn to a further extraordinary feature of the EMU period. Not only is it the case that the historically close relationship between actual inflation and inflation perceptions broke down temporarily; it is also the case that the close relationship between inflation perceptions and expectations broke down in this period.

Historically, when aggregating response shares in terms of perceptions or expectations of (1) inflation, (2) constant prices and (3) deflation, the shares for perceptions and expectations at any point in time corresponding to each of those three categories have been very closely aligned. This relationship broke down around January 2002 with the introduction of euro notes and coins. We illustrate both of these points in Figure 86,

Figure 87 and Figure 88, which plot the absolute differences between shares corresponding to expectations and those corresponding to perceptions for each of the categories (1), (2) and (3).

FIGURE 86: ABS. DIFF. (EXPECTATIONS – PERCEPTIONS) OF INFLATION

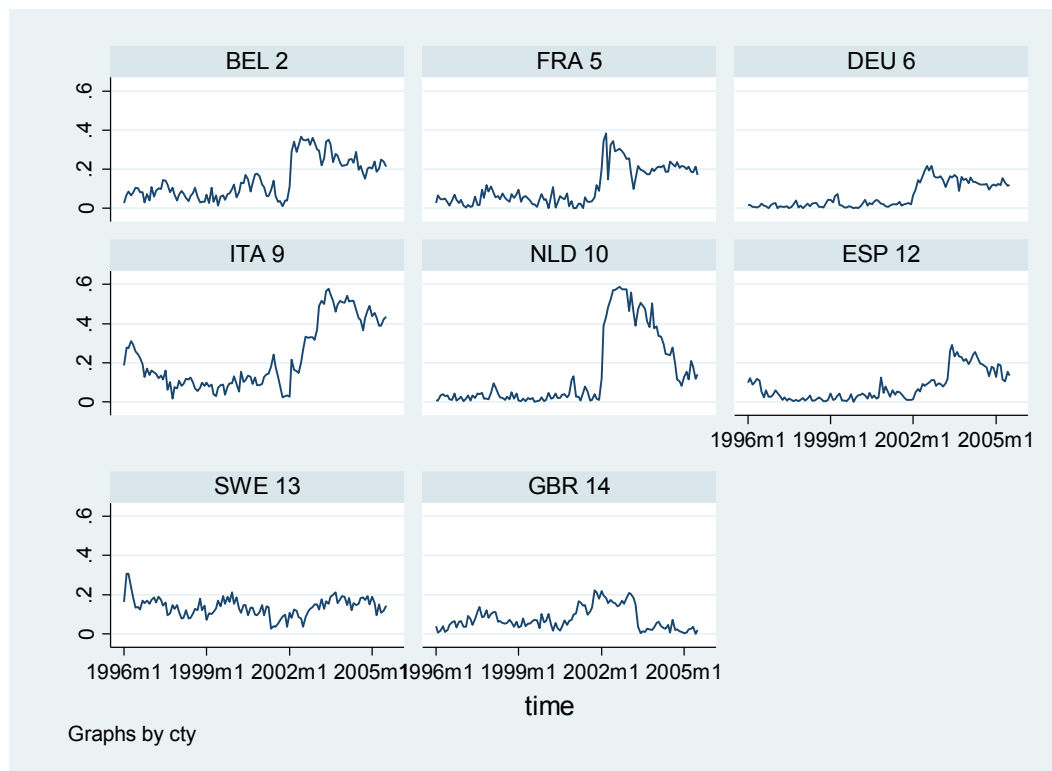


FIGURE 87: ABS. DIFF. (EXPECT. – PERCEPTIONS) OF CONSTANT PRICES

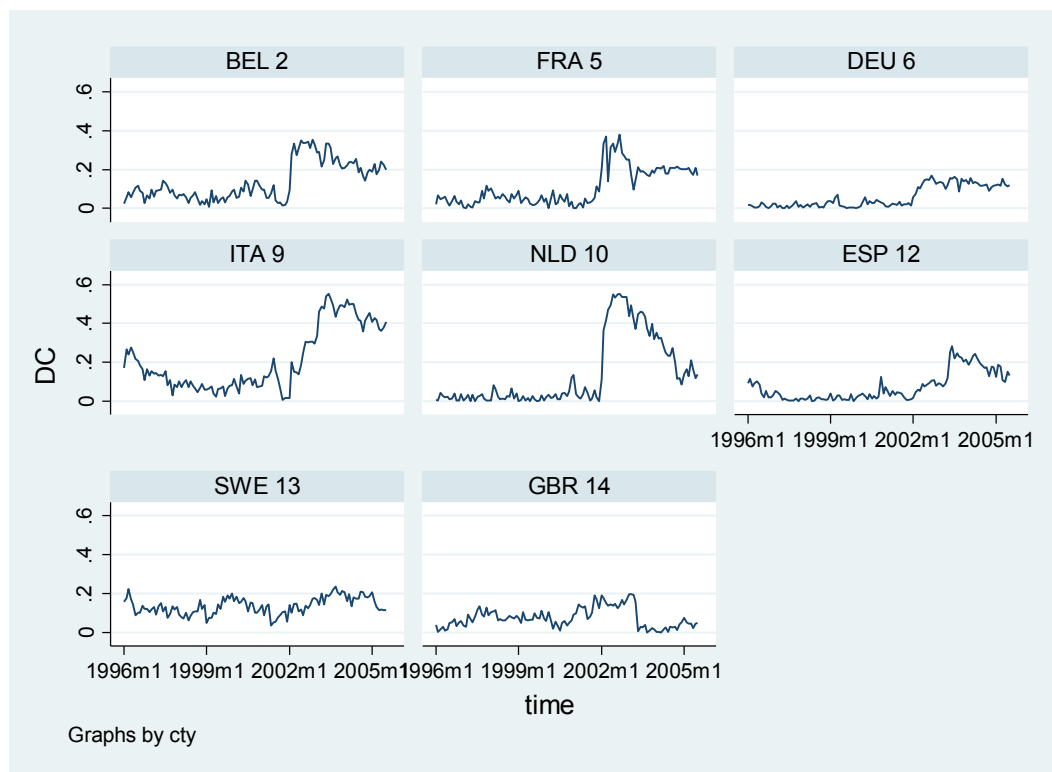
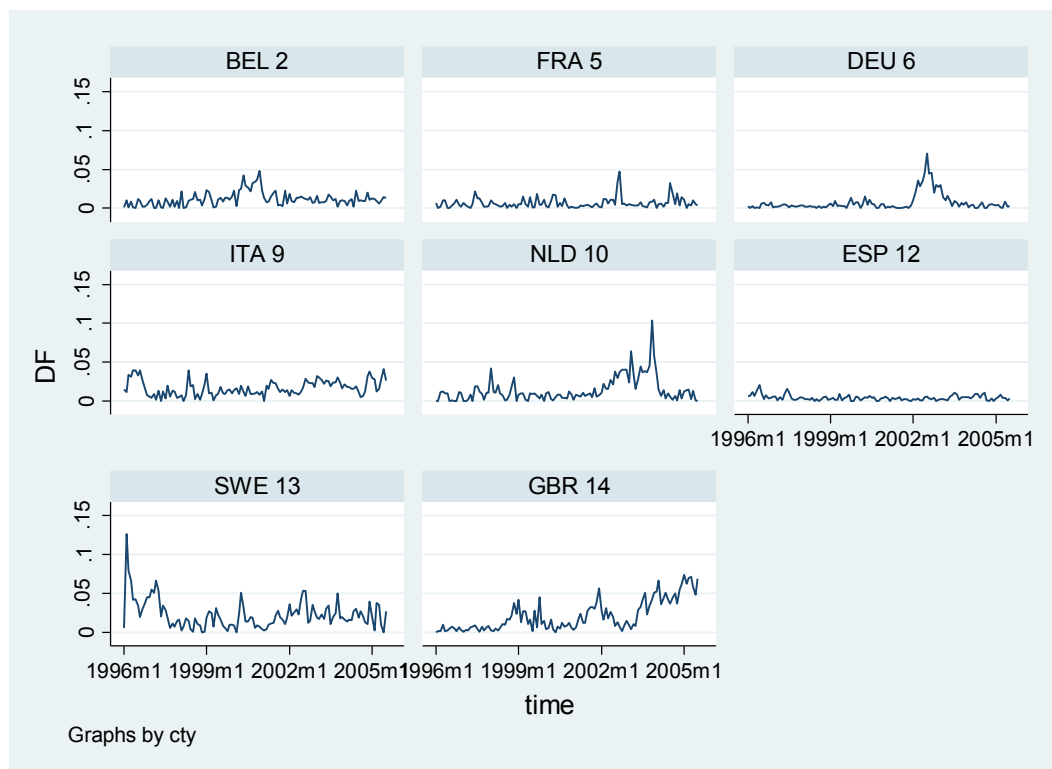


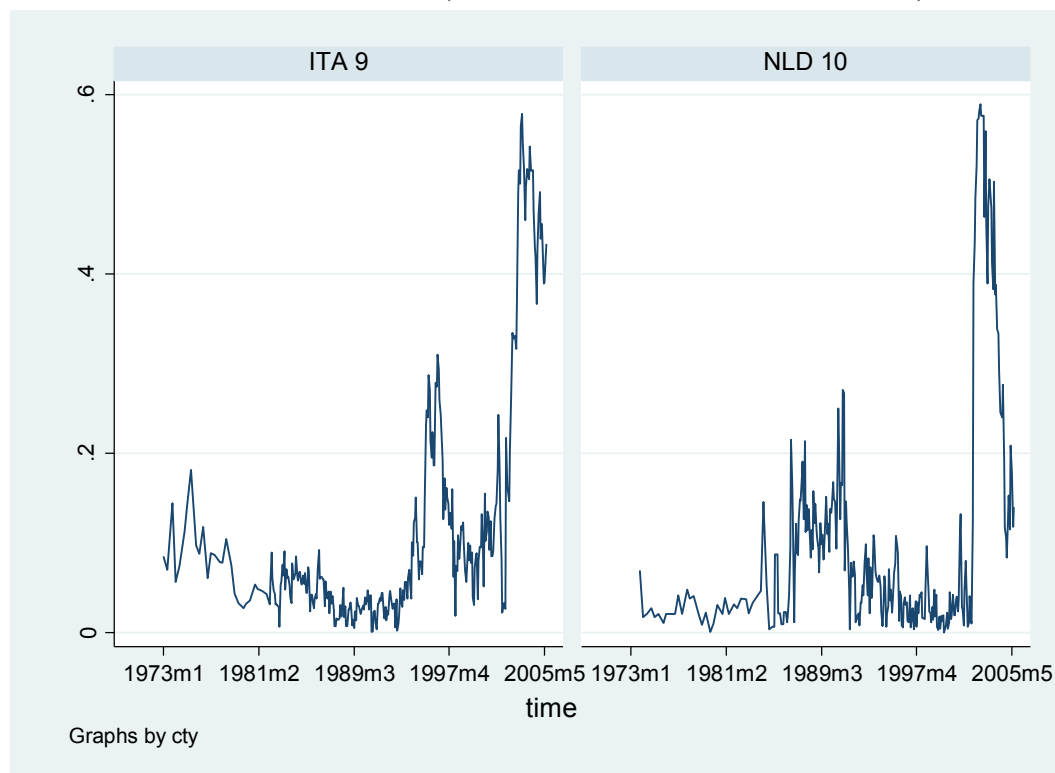
FIGURE 88: ABS. DIFF. (EXPECTATIONS – PERCEPTIONS) OF DEFLATION



The observed movements in these absolute differences reflect what we observed earlier. Namely, the introduction of the euro was associated with a build-up of *both* inflation perceptions and expectations. The build-up of inflation perceptions continued past the date of physical introduction of the euro and reversed only gradually. To the contrary, inflation expectations fell sharply upon introduction of the notes and coins, leading to the observed historical divergence.

It is no exaggeration to refer to this development as ‘historical’. To motivate this statement, Figure 89 provides plots of the absolute differences between the percentages of respondents expecting and perceiving inflation for Italy and the Netherlands, the countries for which the largest samples are available.

FIGURE 89: UK, ITA, ABS. DIFF. (EXPECTATIONS – PERCEPTIONS) OF INFLATION



Again, we note that the phenomenon just described does not concern either the UK or Sweden, confirming its euro-centricity. Furthermore, all of the change is focused on perceptions and expectations of inflation or constant prices as highlighted earlier; the shares of respondents expecting or perceiving deflation change relatively little, and neither does their difference.

It may be tempting to use this surprising development in support of some notion of ‘rational’ expectations. It is however important to be clear about what these data show and what these data do not show. It is certainly true that this phenomenon clearly demonstrates that expectations are not merely extrapolations of what people are currently perceiving, in that they have a forward-looking component and are not simply adaptive. However, this observation cannot be extended to make general claims about rationality; indeed:

1. For most of the historical record these shares have been relatively closely aligned, revealing the importance of adaptive elements.
2. The build-up of inflation perceptions behind this phenomenon does not seem particularly ‘rational’ in most countries, as it is well out of line with underlying inflationary behaviour.

Note further that the issues which generate doubts about rational expectations also question those economic theories whose validity rests upon that of rational expectations.

An important example is credibility theory (Kydland and Prescott, 1977; Barro and Gordon, 1983) which is notably criticised in Forder (2000) on these grounds.

Thus, the overall conclusion here reinforces that of the previous sub-section, in that we can again observe that consumer inflation expectations are more than a basic adaptive process but certainly fall rather short of a fully rational process.

2.3 Understanding the formation of inflation perceptions

2.3.1 Modelling inflation perceptions

An implicit assumption deployed in Chapter 1 – and made without discussion by all the literature to our knowledge – is that when consumers think about prices, they conceptualise them in the same way as the statisticians compiling the consumer price index. That is, to reach their personal notion of ‘consumer prices’, they weigh expenditure components in their heads in the same way as they are weighed in the CPI. In other words, the CPI is their reference index for inflation.

This assumption is debatable, at the very least. For instance, survey participants might respond to the question by thinking of retail price inflation rather than headline inflation. Their responses might be biased towards goods inflation, as many services are consumed less frequently; this matters given the differences between sectoral inflation rates due to technological progress differentials, for instance. One-off, administrative payments that influence the official index may also be excluded from their assessment.

Now, actual CPI inflation is in some form or another used as a scaling mechanism for expectations by all the methods presented in Chapter 1. Should the reference index used by consumers differ from CPI, this would potentially lead to a number of problems whose seriousness depends on the intended application for the quantified data. If the data are to be used essentially as control variables – as in Chapter 2 – this observation will lead to a certain amount of measurement error being introduced, the gravity of which will depend on the relationship between the reference index and the CPI. This may or may not be a serious issue. However, an application where this observation is almost certainly a grave concern is in analyses aiming to establish the ‘rationality’ of inflation expectations. Indeed, differences between the actual reference index and CPI may lead to systematic deviations between inflation expectations and *ex post* observed inflation without this depending in any way on the rationality of consumers – except if the term is also meant to capture their understanding of CPI weights.

This ‘reference index’ problem is a difficult one to address due to severe data constraints. It is straightforward enough to apply the CP method to alternative reference series for the ‘CPI’, thereby deriving alternative measures of inflation perceptions. In particular, it is possible to find the individual component series underlying the official CPI index and the official weightings applied, so that the latter can then be modified. However, it is unclear what the aim of such manipulations should be. The most straightforward answer would be to minimise deviations from the actual inflation series. However, a certain amount of bias – possibly time-varying – cannot be excluded, in ways that are likely to be ignored by the researcher. Thus, such adjustments have no clear benchmark against which they can be measured and are therefore fully inadvisable.

The situation is very different, however, when quantitative perceptions data are available. This is the case for Sweden in our sample. We can then seek to ‘explain’ the perceptions data in terms of inflation in the underlying components of a consumer price index – the HICP, for instance. This leads to a regression of the type:

$$PERCEPTIONS_t = \alpha + \beta_1 INF1 + \dots + \beta_n INF_n + \varepsilon_t \quad (63)$$

Where *INF* denotes inflation for one of the *n* components underlying the consumer price index chosen. If consumers had fully unbiased price perceptions, we would expect the constant α to be zero and the β terms to correspond to the component weights in the consumer price index. We shall refer to this as the ‘rational perceptions hypothesis’.

An alternative modelling approach consists in incorporating the official consumer price index weights (denoted by *CP*) directly into the explanatory variables, as follows:

$$PERCEPTIONS_t = \alpha + \beta_1 * (CP01_t * INF1_t) + \dots + \beta_n * (CPn_t * INF_n_t) + \varepsilon_t \quad (64)$$

As can be seen, the independent variables are now the inflation rates of the individual HICP components *multiplied* by their time-varying weights in the HICP index (whose values are listed below in Table 38). Such a transformation might be hoped to lead to a better specification, in that some of the potential instability that could be experienced might now be subsumed *within* the independent variables (Hendry, 2004). The other advantage will be in the readability of the regression output. Indeed, our null hypothesis of ‘rational perceptions’ would imply a value of zero for α and values of one for all the β

terms. Values either above or below one for the β terms would respectively imply an over or underweighting of the component in question by consumers relative to what would be ‘rational’.

We will use data from Eurostat on the components of the Swedish HICP for this exercise. The components underlying HICP are listed in Table 36, whilst the time-varying weights associated with them are displayed in Table 38 (Appendix II).

TABLE 36: HICP COMPONENTS

INF1	"Food and non-alcoholic beverages"
INF2	"Alcoholic beverages and tobacco"
INF3	"Clothing and footwear"
INF4	"Housing, water, electricity, gas and other fuels"
INF5	"Furnishings and household goods"
INF6	"Health"
INF7	"Transport"
INF8	"Communication"
INF9	"Recreation and culture"
INF10	"Education"
INF11	"Hotels, Cafés and Restaurants"
INF12	"Miscellaneous goods and services"

SOURCE: Eurostat

Table 37 presents the results of regressions based on (63) and (64). We sought a parsimonious specifications by removing insignificant and non-positive coefficients⁹⁶ from the general models:

⁹⁶ We note that component 8 displayed a significant negative coefficient under both models at the 1% level; in addition, the first model featured a significant negative coefficient for component 10 at the 5% level.

TABLE 37: MODELLING INFLATION PERCEPTIONS (SWE), 1993m1-2005m7

Equation (63): PERSCBEV		Equation (64): PERSCBEV	
Constant	0.315 *** (0.121)	Constant	0.652 *** (0.104)
INF1	0.097 *** (0.017)	WINF1	0.512 *** (0.093)
INF4	0.102 *** (0.017)	WINF4	0.407 *** (0.099)
INF5	---	WINF5	1.445 *** (0.455)
INF6	0.039 ** (0.016)	WINF6	5.432 *** (0.933)
INF7	0.048 ** (0.020)	WINF7	---
INF11	0.362 *** (0.037)	WINF11	2.530 *** (0.636)
I1995:7	1.543 *** (0.509)	I1995:7	1.695 *** (0.499)
Sigma: 0.498 R ² : 0.65 AR 1-7 test: F(7,137) = 17.477 [0.0000]** ARCH 1-7 test: F(7,130) = 3.7419 [0.0010]** Normality test: Chi ² (2) = 0.82825 [0.6609] hetero test: F(11,132) = 1.7950 [0.0608] hetero-X test: F(21,122) = 2.7127 [0.0003]** RESET test: F(1,143) = 0.032230 [0.8578]		Sigma: 0.490 R ² : 0.66 AR 1-7 test: F(7,137) = 21.862 [0.0000]** ARCH 1-7 test: F(7,130) = 5.8402 [0.0000]** Normality test: Chi ² (2) = 0.35057 [0.8392] hetero test: F(11,132) = 4.6571 [0.0000]** hetero-X test: F(21,122) = 3.9895 [0.0000]** RESET test: F(1,143) = 26.326 [0.0000]**	

NOTES: *, **, *** respectively denote significance at the 10%, 5% and 1% levels. Standard errors in parentheses; p-values in brackets.

It is transparent from the results in Table 37 that data reject the idea of ‘rational inflation perceptions’ suggested above. The constant is highly significant in both specifications and only inflation components broadly corresponding to food, non-alcoholic beverages, housing and energy, health, transport and catering and accommodation services are retained. Notice that some components are either under- (e.g. INF1) or over-weighted (e.g. INF11) relative to what we would have expected from the HICP weights. Furthermore, the models perform very poorly in terms of specification diagnostics, notably in terms of residual autocorrelation and heteroscedasticity. In addition, Chow tests reveal the regressions to be highly unstable.

In light of these failings, a more sophisticated modelling exercise was carried out through the development of alternative equilibrium correction specifications. This approach seems sensible; indeed, there may well be timing errors in the data. Furthermore, it is easy to conceive that, on a sample of monthly data, lagged perceptions of inflation would be relevant. To address such concerns, a set of nested EqCM specifications was developed, presented below from the most to the least restrictive:

$$\begin{aligned}
 H1: \Delta PERCEPTIONS_t &= \alpha_1 * \Delta HICPINFLA_t + \alpha_2 * (HICPINFLA - PERCEPTIONS)_{t-1} + \varepsilon_t \\
 H2: \Delta PERCEPTIONS_t &= \sum_{i=1}^{12} \alpha_{1i} * \Delta INF_i_t + \alpha_2 * (HICPINFLA - PERCEPTIONS)_{t-1} + u_t \quad (65) \\
 H3: \Delta PERCEPTIONS_t &= \sum_{i=1}^{12} \alpha_{1i} * \Delta INF_i_t + \sum_{i=1}^{12} \alpha_{2i} * [w_i * (INF_i - PERCEPTIONS)]_{t-1} + v_t
 \end{aligned}$$

“INF_i” denotes the inflation of the *i*th component of the HICP index, and *w_i* its associated index weight; “HICPINFLA” denotes overall HICP inflation; “PERCEPTIONS” denotes the quantitative inflation perceptions of Swedish consumers evidenced by the SCB/GfK survey.

As can be seen, model H2 imposes a restriction on the EqCM part of the equation relative to H3 (namely, that all coefficients α_{2i} are equal), whilst H1 imposes a further restriction on the dynamics of the equation (which reduce to the simple change in HICP inflation). Thus, in H3 both the dynamics and the long-run solution are allowed to be influenced distinctly by the various individual HICP components.

Having estimated all three models in their most general form, we are able to carry out nested tests on the basis of the likelihood ratio statistic, which has the asymptotic distribution:

$$LR = 2\{\log L_{Unrestricted} - \log L_{Restricted}\} \xrightarrow[D]{H_0} \chi^2(df) \quad (66)$$

The results of this exercise are presented in Table 39. As can be seen, the general H2 and H1 specifications are at the boundary of being rejected against the general H3 model. On the other hand, the restrictions implied by the parsimonious H3 specification are comfortably accepted.

TABLE 39: SWEDISH INFLATION PERCEPTIONS, NESTED TEST RESULTS

Case	No. Restrictions	LogLik. Statistic	Crit. value (5%)	Result
General H3 vs General H2	11	19.50	19.68	Marginal non-rejection
General H3 vs General H1	22	32.69	33.92	Marginal non-rejection
General H3 vs Specific H3	14	16.00	23.69	Do not reject

The non-rejection results for specifications H2 and H1 are so marginal that they hardly build much confidence in the implied restrictions. Thus, it would seem to be important to express *both* the dynamics and the long-run solution of the model in terms of the individual components of HICP inflation.

We now present the selected parsimonious H3 specification in Table 40⁹⁷. The model appears to be congruent (Figure 90, Appendix II) and quite stable overall (Figure 91, Appendix II), excepting some faint traces of instability around 2000.

⁹⁷ We note that none of the coefficients of the model were found to be significantly negative at the 5% level in the selection process.

TABLE 40: PARSIMONIOUS EqCM MODEL, INFLATION PERCEPTIONS

Modelling DPERSCBEV; Estimation Sample: 1993m2-2005m7					
	Coefficient	Std.Error	t-value	t-prob	Part.R^2
DINF6	0.0548654	0.01917	2.86	0.005	0.0549
DINF7	0.0454588	0.02045	2.22	0.028	0.0339
EqCMHICP1_1	0.187491	0.06690	2.80	0.006	0.0528
EqCMHICP4_1	0.106273	0.05452	1.95	0.053	0.0262
EqCMHICP5_1	0.579224	0.2760	2.10	0.038	0.0303
EqCMHICP7_1	0.113488	0.05861	1.94	0.055	0.0259
EqCMHICP11_1	1.07929	0.4452	2.42	0.017	0.0400
EqCMHICP12_1	0.744322	0.3026	2.46	0.015	0.0411
I1996:2	-1.27997	0.3298	-3.88	0.000	0.0965
sigma	0.314375	RSS		13.9352846	
log-likelihood	-34.6249	DW		2.07	
no. of observations	150	no. of parameters		9	
mean (DPERSCBEV)	-0.0034	var (DPERSCBEV)		0.123302	
Specification Tests					
AR 1-7 test:	F(7,134) =	0.75318	[0.6274]		
ARCH 1-7 test:	F(7,127) =	1.5047	[0.1714]		
Normality test:	Chi^2(2) =	1.7453	[0.4178]		
hetero test:	F(17,123)=	1.1647	[0.3033]		
hetero-X test:	F(45,95) =	1.2410	[0.1891]		
RESET test:	F(1,140) =	0.026845	[0.8701]		

The model in Table 40 makes clear that not all HICP components seem to matter significantly in terms of shaping consumer inflation perceptions; furthermore, heterogeneity is also evident in the influence of the components that are retained. These results therefore question the validity of the rational perceptions hypothesis – in a way that is more convincing than in Table 37, given the quality of the model.

We now turn to exploring the rationale for certain HICP components being retained rather than others. Table 41 lists these components, whether kept as part of the dynamics, the long-run solution, or indeed both.

TABLE 41: HICP COMPONENTS

Code	Concept	Retained?
INF1	"Food and non-alcoholic beverages"	YES
INF2	"Alcoholic beverages and tobacco"	NO
INF3	"Clothing and footwear"	NO
INF4	"Housing, water, electricity, gas and other fuels"	YES
INF5	"Furnishings and household goods"	YES
INF6	"Health"	YES
INF7	"Transport"	YES
INF8	"Communication"	NO
INF9	"Recreation and culture"	NO
INF10	"Education"	NO
INF11	"Hotels, Cafés and Restaurants"	YES
INF12	"Miscellaneous goods and services"	YES

We hypothesise that the importance of a particular set of products in determining the inflation perceptions of consumers is increasing in the following four factors:

1. The frequency at which the prices in question are adjusted
2. The weight of these products in the overall expenditure of consumers
3. The level of inflation applicable to those products over the sample
4. The frequency at which the consumer purchases those products

Table 42 provides data on some of these dimensions and computes a rough index of the ‘expected importance’ of particular components in consumer perceptions on the basis of this information. Unfortunately, it was not possible to find data on the fourth factor. Data on dimension 1 are based on the mean duration of prices by HICP component for Finland (Vilmunen and Laakkonen, 2004), collected and analysed in the context of the ECB’s Inflation Persistence Network. The use of Finnish data may be justified by the very similar HICP weights for both countries and great similarities in their industrial structure, business cycles and economic policies (as documented in Jonung and Sjöholm, 1999).

TABLE 42: FACTORS HYPOTHESED TO AFFECT INFLATION PERCEPTIONS

HICP	Median Duration	Frequency Rank	Avg HICP weight	Weight Rank	Avg. Inflation (Abs. Value)	Inflation Rank	Importance Index (% Max)	Overall Rank
1	3.23	4	17.32%	2	0.41%	10	95.43%	3
2	4.30	6	5.78%	7	2.93%	2	95.19%	4
3	3.41	5	7.33%	5	0.54%	9	87.03%	7
4	3.10	3	19.15%	1	1.54%	6	99.02%	2
5	5.50	8	6.42%	6	0.99%	7	80.60%	10
6	6.30	10	1.68%	11	6.56%	1	93.69%	5
7	1.51	1	16.74%	3	2.64%	3	99.54%	1
8	1.80	2	3.16%	10	0.62%	8	90.79%	6
9	4.46	7	11.63%	4	0.21%	11	82.22%	8
10	23.33	12	0.18%	12	0.01%	12	0.00%	12
11	6.76	11	5.63%	8	2.09%	5	74.58%	11
12	5.96	9	4.97%	9	2.59%	4	81.30%	9

Notes:

- Median duration is an average for the years 1997-2003 and is expressed in months. Vilmunen and Laakkonen (2004) separated the sample for the years 1997-99 and 2000-03 for technical reasons which are not a concern here. The data concern Finland, as no similar data were available for Sweden.
- HICP weights represent averages for the years 1990-2005.
- The absolute value of average inflation is computed over the period 1993m1-2005m7.
- The index of expected importance is computed as a percentage of the maximum attainable score. It is computed according to the following expression:

$$IMPINDEX = \frac{[1728 - (FREQUENCY * WEIGHT * INFLATION ranks)]}{1727} * 100 \quad (67)$$

Table 42 provides some support for our hypothesised factors influencing inflation perceptions. As can be seen, our parsimonious EqCM model retained precisely the top three factors highlighted by our ‘expected importance’ index, namely: transport, food and non-alcoholic beverages, and housing/water/energy costs. The health component (importance rank 5) was also retained; this suggests that a component requires strength on one signal alone to be spotted by consumers. Indeed, this component ranks rather low in terms of either frequency of price adjustment or HICP weight; however, it has experienced almost twice as much inflation as any other sector over the period, making it a clear outlier in this respect.

Our model has also retained certain components, despite relatively low importance scores above. The importance of the “hotels and catering services” component probably arises from the frequency at which the relevant products are purchased by consumers (almost certainly daily in the case of café and restaurant services), information that is missing from our index. The retention of the “furnishings and household goods” (component 5) and “miscellaneous goods and services” (component 12) is perhaps less clearly motivated. Component 5 essentially concerns a number of household durables, tools, maintenance and repairs as well as domestic and home care services. Component 12 notably includes personal care products and services (e.g. hairdressing), as well as banking and insurance services.

The measures underlying Table 42 are not ideal. Moreover, we make no claim that each of these concepts (frequency, weight and inflation) ought to be aggregated with equal weights into an index of importance for consumer perceptions, as is implemented above.

Notwithstanding this, the way in which this table corroborates our earlier evidence does provide support to our hypothesis on the determinants of consumer perceptions.

2.3.2 Impact on assessments of the rationality of expectations and quantification

An immediate consequence of the analysis in the previous subsection is that particular thought should be given to what is meant by the rational expectations hypothesis (REH), especially in the context of empirical tests thereof.

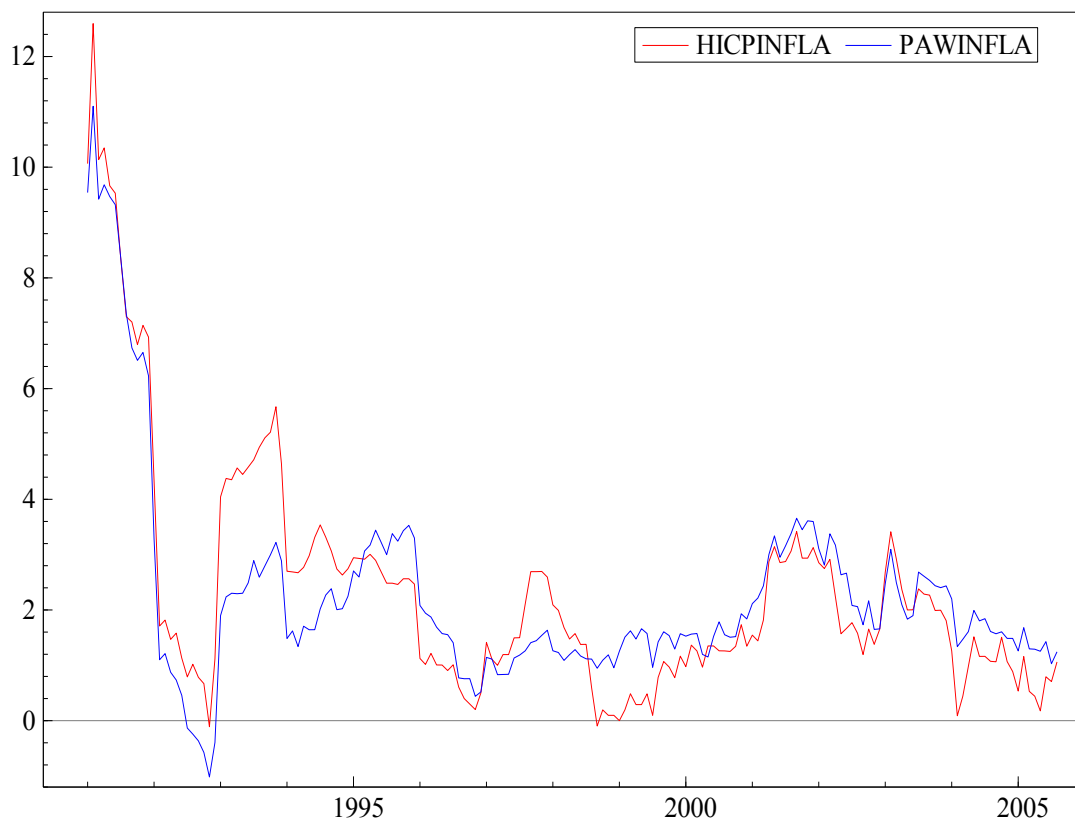
Every study of the REH published to date to our knowledge essentially compares some measure of expectations with the actual outturns of inflation, as measured by official statistics. The previous subsection, however, has shown that consumers have an incomplete and possibly distorted view of what constitutes inflation, arguably dependent on a number of factors such as the frequency of price adjustment and purchases, expenditure weights and relative inflation of the underlying expenditure components. There is therefore no guarantee that the forecasts offered by consumers relate to the entirety of the CPI. In fact, there is rather a suspicion of the contrary, as the perceptions data underlying the previous section were elicited with a question relating to prices in general.

This casts doubt on the results and interpretation of past empirical studies of the rational expectations hypothesis. Indeed, given the above, the failure of consumer forecasts of CPI-type measures cannot but be *expected* – and results to the contrary seem a priori puzzling and may raise issues of sample-dependence. The relevant interpretation, however, need not be that consumers are not rational, as this issue is confounded with the

additional null hypothesis that consumers have a correct understanding of inflation. It may be the case – although this is unlikely – that consumers make rational forecasts *on the inflation index that they perceive*. We now aim check this conjecture empirically.

We proceed by reweighting the components of HICP in a way that reflects their actual perceptions by consumers, thereby yielding a “perceptions-adjusted” inflation series (PAWINFLA). This is generated on the basis of the long-run solution to our specification in Table 40. We plot this perceptions-adjusted series against actual HICP inflation in Figure 92; the series appear to be closely related, although on occasions they do move in opposite directions.

FIGURE 92: PERCEPTIONS-ADJUSTED INFLATION vs HICP INFLATION



We then use the perceptions-adjusted series as the basis from which to generate the inflation expectations measure recommended in Chapter 1, namely EXPINFLPBIS. This allows us to consider two applications in what follows. First, we investigate the success of this new expectations measure in predicting both actual HICP inflation and future “perceptions-adjusted inflation”. Second, we test the ability of this new series to reproduce the quantitative data on expectations available for Sweden.

Application 1:

Testing rationality relative to HICP and “perceptions-adjusted inflation”

We now consider the possible impact of the findings in the previous sub-section for assessments of the rationality of consumer expectations. We will be making use of our preferred quantification methodology in this exercise: EXPINFLPBIS (see Chapter 1). However, the quantification method now relies on perceptions-adjusted inflation for the quantification of question 5 of the EC Consumer Survey.

First, we contrast the forecasting performance of the EXPINFLPBIS measure with that of a naïve benchmark which merely extrapolates the perceptions-adjusted inflation of the past 12 months ($\pi_{t+12,t}^{\text{exp}} = \pi_{t,t-12}$). The forecasts are assessed against future perceptions-adjusted inflation in Table 43.

TABLE 43: PREDICTIVE ABILITY, PERCEPTIONS-ADJUSTED INFLATION

Measure	Period	ME	MAE	RMSE	U1 %	U2 %	U3 %
EXPINFLPBIS	1997m1-2005m7	0.0044	0.0072	0.0092	22.90	19.94	57.16
Naïve	1997m1-2005m7	0.0004	0.0056	0.0067	0.38	23.18	76.43

As can be seen from the results, consumers appear to be little more rational in their forecasts than under traditional assessments of rationality despite these adjustments. The RMSE achieved here (0.0092) is only slightly improved relative to that achieved in the corresponding exercise in Table 8 (0.0107; see p.54). Moreover, the finding that naïve forecasts outperform consumer expectations appears to be robust to these modifications. The MSE decomposition achieved here is also inferior to its counterpart in Table 10.

We now contrast the predictive performance of EXPINFLPBIS for HICP inflation with that of a naïve benchmark projecting past HICP inflation as expectations.

TABLE 44: PREDICTIVE ABILITY, HICP INFLATION

Measure	Period	ME	MAE	RMSE	U1 %	U2 %	U3 %
EXPINFLPBIS	1997m1-2005m12	0.0017	0.0085	0.0103	2.74	31.08	66.18
Naïve	1997m1-2005m12	0.0001	0.0100	0.0123	0.00	53.28	46.72

Table 44 provides evidence that the modified EXPINFLPBIS series exhibits essentially the same performance as its traditional version in Chapter 1 when it is contrasted with outcomes for a complete index of consumer prices. The RMSE here is 0.0103, which is almost identical to that in Table 8 (0.0107). The naïve benchmark is outperformed in this particular exercise, but the improved RMSE remains in the same order of magnitude and as such does not seem to constitute a significant qualitative finding.

Application 2:

Testing the match of the new expectations measure and quantitative data

We now consider whether the enhanced understanding of consumer inflation perceptions gained in the previous section is able to improve the ability of quantified qualitative survey data to match actual quantitative data on inflation expectations. In other words, is the matching ability of our EXPINFLPBIS measure improved when it is quantified with the help of “perceptions-adjusted inflation”? The answer appears to be positive.

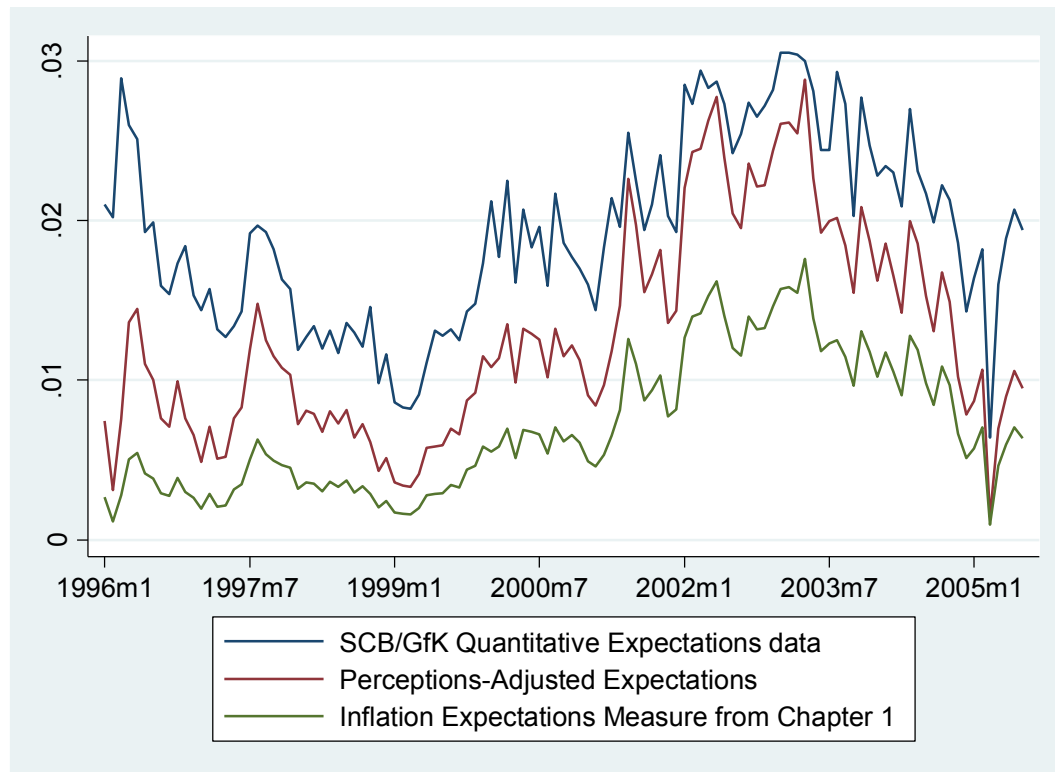
TABLE 45: ABILITY TO MATCH QUANTITATIVE DATA

Measure	Period	ME	MAE	RMSE	U1 %	U2 %	U3 %
EXPINFLPBIS	1996m1-2005m7	0.0066	0.0066	0.0071	85.14	2.77	12.09

A contrast of Table 45 and Table 17 reveals that the new version of EXPINFLPBIS has a considerably lower bias than its earlier counterpart (mean error of 0.0066 versus 0.0123 earlier). The RMSE is also around half of the one found previously (0.0126), and the MSE decomposition appears to have improved. Thus, taking into account consumers’ biased perceptions of inflation seems to be an important step in improving measures of inflation expectations derived from qualitative survey data.

Figure 93 confirms visually the improved ability of the perceptions-adjusted EXPINFLPBIS to reproduce the quantitative expectations data available in Sweden (EXPSCBEV) relative to the traditional EXPINFLPBIS as quantified in Chapter 1.

FIGURE 93: OLD / NEW EXPINFLPBIS AGAINST QUANTITATIVE EXPECTATIONS



2.3.3 Conclusion

The analysis of Swedish quantitative perceptions data carried out in this section has enabled us to reject the ‘rational perceptions hypothesis’. Indeed, only a limited number of HICP components seem to be significantly reflected in the inflation perceptions of consumers, and those retained tend to be either under- or overweighted relative to the benchmark HICP weights. An EqCM model of inflation perceptions was developed, which succeeded in capturing most of the features of the data at hand. It was notably found that allowing for a differentiated impact of individual HICP components was preferable, both in the dynamics and the long-run solution of the equation.

An important contribution of this section is the light it throws on the determinants of consumer perceptions – an exercise carried out for the first time to our knowledge. The

list of selected components is overall very sensible: it includes those that are characterised by extremely frequent purchases (e.g. catering services), frequent price changes (e.g. transport), bear a large weight on consumer expenditures (e.g. housing services, water and energy) or have experienced relatively large inflation over the sample (e.g. health).

These findings suggest that particular care should be taken when deploying expectations measures in certain applications, such as tests of the rational expectations hypothesis (REH). Indeed, the failure of the rational perceptions hypothesis implies the failure of the REH, but a priori leaves open the question of whether consumers make rational forecasts on the inflation index that best corresponds to their perceptions. The preliminary evidence presented here suggests that this is not the case, nor does the modified quantification methodology substantially improve the success of forecasting HICP inflation.

However, it was possible to show that taking into account biases in consumers' perceptions of inflation led to considerable improvements in the ability of quantified qualitative survey data to reproduce actual quantitative survey data – assumed to provide a proxy for the true expectations. This suggests that an important avenue for future improvement of expectations measures based on qualitative data will be the development of stable econometric relationships summarising our understanding of consumer perceptions in additional countries.

The issues touched upon here are obviously extremely important from a measurement perspective. The absence of attention they have elicited in the literature is justified in part by data limitations. This is fortunately about to change. Indeed, it appears that quantitative questions have recently been introduced in the EC Consumer Surveys of a number of countries, including Belgium (since May 2003)⁹⁸ and France (since January 2004; see INSEE, 2004). These data are currently subject to a test period and are unavailable at present⁹⁹. The results developed here suggest that future research would do well to explore its implications when it is released.

⁹⁸ I am grateful for a private communication by Isabelle de Greef and Jean-Paul Vonck from the *Banque Nationale de Belgique* with this information.

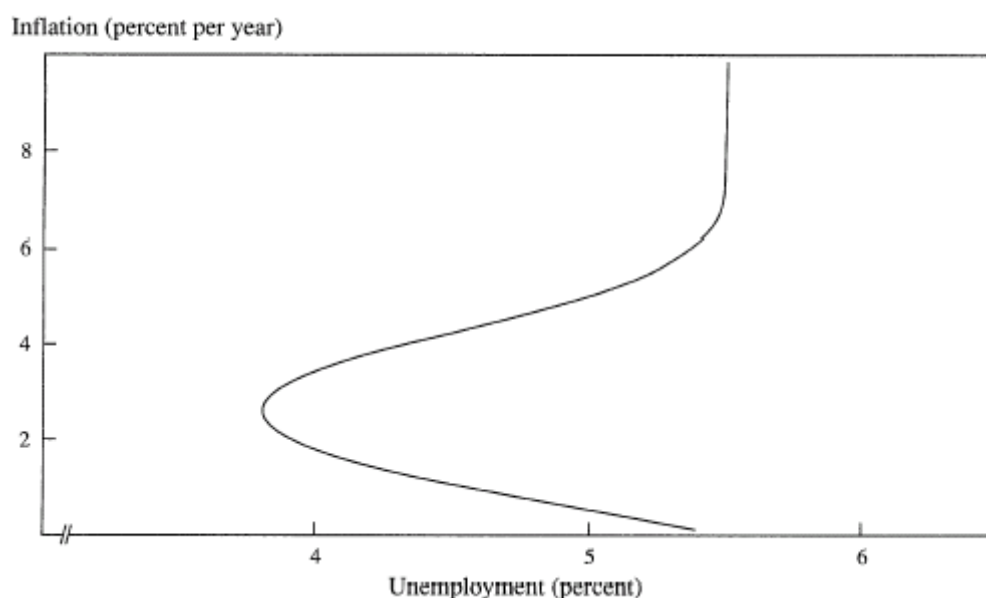
⁹⁹ See Lindén (2006) for additional details and some preliminary results.

3 TESTING THE ‘NEAR-RATIONALITY’ HYPOTHESIS

3.1 Introduction

Akerlof, Dickens and Perry (2000) (henceforth ADP) put forward the hypothesis that the population of agents in the economy is divided into two groups: whilst some form fully rational expectations, others hold “nearly rational” expectations, in that they underweight or even ignore inflation. The proportion of such agents in the economy is a function of the prevailing inflation rate, as higher rates increase the incentive to gather information about inflation. This hypothesis implies a particular shape for the Phillips curve, displayed below:

FIGURE 94: ADP HYPOTHETICAL LONG-RUN PHILLIPS CURVE



SOURCE: Akerlof, Dickens and Perry (2000)

The particular numbers used in this figure come from a calibration exercise carried out by ADP, but the argument is more general. The rationale for this shape is that at zero inflation, both rational as well as nearly-rational agents correctly expect no inflation. However, as inflation increases it is underestimated by the nearly-rational agents, leading

them to overestimate their real wage increases and substitute work for leisure. Inflation increases however also make it worthwhile for some agents to start forming expectations rationally. In the limit, all agents form expectations rationally, and the economy is again operating at the ‘natural rate’ of unemployment.

ADP presented estimates of expectations-augmented Phillips curves on the basis of US data consistent with their theory. A similar exercise was carried out for Sweden by Lundborg and Sacklén (2001). However, Bryan and Palmqvist (2005) have recently criticised those exercises as being unable to distinguish the ADP hypothesis from other competing possibilities. The authors carry out alternative empirical verifications and find that the data fail to reveal the correspondence between expectations errors and actual inflation implied by the near-rationality hypothesis, as advanced by ADP.

In what follows, we aim to reproduce the results obtained by Bryan and Palmqvist (2005) for Sweden, as well as reconsider the problem in the context of the other countries in our sample by making use of the quantitative Gallup data for the UK and the qualitative EC Consumer Survey data quantified according to the method selected in Chapter 1.

As is common in macroeconomic theory, dynamics are absent from the ADP model so that it is not clear what types of inflation “shocks” are to be considered relevant empirically. Should we expect the relationship between expectations errors and actual inflation proposed by ADP to be visible for all inflation shocks, however temporary? Or would this pattern only be appropriate in describing responses to more persistent changes in inflation?

We follow Bryan and Palmqvist (2005) by considering the problem in two steps. First, we identify a number of “inflation regimes” for every country and check whether the pattern predicted by ADP is apparent for these persistent changes in inflation. Second, we consider the ADP pattern in light of all fluctuations in inflation.

3.2 Testing Near-Rationality in the context of regime changes

We carried out Bai-Perron (1998, 2003) (henceforth BP) break-point tests on measures of inflation in order to identify potential changes in inflation regimes over our sample. The BP procedure allows for the testing of multiple structural changes, with the dates of the breaks treated as unknown variables to be estimated. Intuitively, the break-point estimators are global minimisers of an objective function based on the sum of squared residuals. Since what is of interest here is the detection of major structural changes in the mean of the series, we applied the BP procedure with only a constant as a regressor, as in Bai and Perron’s (2003a) investigation of the ex post US real interest rate. Thus, we considered the following simple linear regression model with m breaks ($m+1$ regimes):

$$y_t = \delta_j + u_t \quad \text{for } t = 1, \dots, T \text{ and } j = 1, \dots, m+1; \quad u_t \text{ is a disturbance term}$$

The authors developed a number of tests in this context and provided critical values in Bai and Perron (2003b). We will focus on two of these:

- The *double maximum test*, which considers the null hypothesis of no structural break against an unknown number of breaks bounded by some value M . The test relies on a number of “sub-tests” that check the null of no break versus a *specific* number of breaks m (inferior or equal to M). Two alternatives of the double

maximum procedure are proposed, differing in terms of the weights assigned to the underlying sub-tests:

- UDmax: equal weights are applied to the individual tests.
 - WDMax: weights are applied to the individual tests such that the marginal p-values are equal across values of m .
-
- The *test of α versus $\alpha+1$ breaks*: $\sup F_T(\alpha \mid \alpha+1)$. This test is applied sequentially to find the number of breaks. A commonly used alternative to select the dimension of a model is the application of information criteria. However, Perron (1997) presented a simulation study showing poor performance for the AIC, poor performance for the BIC in cases with serially correlated errors and a tendency for the modified Schwartz criterion of Liu et al. (1997) to underestimate the number of breaks when some are present. Furthermore, such procedures are unable to take into account potential heterogeneity across segments, unlike the sequential approach (BP, 2003a). Thus, the latter would seem to be a more reliable option and will be implemented here.

For Sweden, the sample considered aims to match that examined by Bryan and Palmqvist (2005) to check that we are able to reproduce their results (1979m1-2001m12). For all the other countries, we consider the maximum periods over which monthly data of inflation expectations are available. For EC Consumer Survey data, the starting points are listed in Table 1 (p.17). The end points of all such samples are set at 2005m8. The inflation measure considered is based on the OECD's "CPI, All Items" measure except in

the case of Sweden, for which the Core Inflation measure UNDEX is considered – again, to match Bryan and Palmqvist’s analysis.

The computations were carried out by a GAUSS program available on Pierre Perron’s website¹⁰⁰. We used a trimming value of $\varepsilon = 0.15$ (in percentage), as recommended by the authors in cases exhibiting serial correlation, and considered up to 5 breaks. The aim of this is to ensure that the size of the test is adequate under the hypothesis of no breaks – otherwise, estimations of some quantities might be carried out over insufficient numbers of observations, leading to imprecise estimates and size distortions in the tests. In the interests of generality and robustness, we allowed for serial correlation in the errors and different variances of the residuals across segments.

The primary method for determining the number of breaks in a particular series is the sequential application of the $\sup F_T(\alpha \mid \alpha + 1)$ test. However, there are instances where this sequential testing procedure can be improved. As noted in BP (2003a), in the presence of multiple structural breaks one sometimes comes across certain configurations of changes such that it is difficult to reject the null of 0 versus 1 break, whereas the rejection of the null of 0 versus a higher number of breaks is not problematic. Such cases lead to a breakdown of the sequential procedure. We follow the strategy advocated by BP (2003a) in considering whether the information of the UDMax and WDMax tests conflicts with that of the sequential procedure. When this is the case, we ignore the results of the $\sup F(1|0)$ test and choose the number of breaks m such that the tests $\sup F(\alpha \mid \alpha + 1)$ yield insignificant results for $\alpha \geq m$.

¹⁰⁰ See <http://econ.bu.edu/perron/code.html>

Table 46 presents the results of the BP tests we implemented, most notably the conclusions drawn as to the number and dates of break-points in the series:

TABLE 46: RESULTS OF BAI-PERRON TESTS FOR INFLATION REGIMES

cty	Sample	UDMax	WDMax	Sequential: Breaks found	supF(i+1 i) [where needed]	Result: Breaks
GBR	1981m11-2005m8	4.74	5.31	None	-	None
ITA	1982m1-2005m8	21.6 ***	34.3 ***	None	supF(2 1) = 43.1 ***	1985m7; 1996m5
FRA	1986m4-2005m8	24.6 ***	54.4 ***	1 ***	-	1992m6
NLD	1986m4-2005m8	6.71	11.5 **	None	supF(2 1) = 6.65	None
BEL	1985m10-2005m8	9.54 **	18.95 ***	None	supF(2 1) = 8.66 **	1988m12; 1994m10
ESP	1986m6-2005m8	54.0 ***	107.6 ***	1 ***	-	1995m12
DEU	1980m10-2005m8	2.49	4.03	None	-	None
SWE	1979m1-2001m12	51.2 ***	115.5 ***	2 ***	-	1984m6; 1992m1

Notes: *, ** and *** denote significance at the 10%, 5% and 1% levels respectively. Maximum number of breaks considered M=5. The value of WDMax changes with the significance level; the most significant value is provided here. In the cases where the sequential test procedure breaks down, the supF test with the last significant supF test is provided – or alternatively an insignificant supF(2|1) in case no such tests were significant.

As can be seen, breaks were found for the samples of all but three countries – the United Kingdom, the Netherlands and Germany. In two cases – Italy and Belgium – the sequential procedure clearly broke down and we reverted to a ‘manual’ consideration of the supF statistics. There was a slight doubt as to whether the Netherlands were also in this situation given the value of the WDMax test, but no breaks could be found in the end. The breaks we find for Sweden (1984m6 and 1992m1) are roughly consistent with those found by Bryan and Palmqvist (2005) (namely, 1982m8 and 1991m2), with the slight differences possibly due to revisions in Sweden’s UNDI_X Core Inflation measure.

We first present our replication of the Bryan and Palmqvist test in the context of Sweden in Table 47. As can be seen, our results are fully consistent with the ones they present in their Table 1. Expectations errors are calculated as expected inflation minus actual inflation.

TABLE 47: SWEDEN, EXPECTATION ERRORS & INFLATION BY REGIME

cty	Period	Avg. EXPERR (%)	Avg. UND1X INFLA (%)
SWE	1979:I-2001:12	-0.04	5.15
	1979:I-1984:II	-0.56	9.48
	1984:III-1991:IV	0.80	5.97
	1992:I-2001:12	-0.16	2.19

The pattern predicted by ADP is as follows: an expectation error close to zero for very low levels of inflation; as inflation rises, there is a pattern of increasing underprediction (i.e. negative EXPERR); this peaks at some level of inflation and subsequently under prediction diminishes back to zero as inflation rises further.

As can be seen, the pattern experienced in Sweden is inconsistent with this story. Underprediction takes place at low levels of inflation, becomes overprediction at middle levels and turns again into underprediction at high rates of inflation. This yields an “inverted U-shape” rather than the relationship predicted by ADP.

Table 48 reproduces this exercise for the other countries in our sample on the basis of the quantified EC Consumer Survey data. Due to the inability to identify breaks with the BP procedure on this sample in the case of the United Kingdom, the Netherlands and Germany, our comments will be restricted to the remaining countries.

TABLE 48: EC DATA, EXPECTATION ERRORS & INFLATION REGIMES

cty	Period	Avg. EXPERR	Avg. INFLA
UK	1981m11-2005m8	-0.20%	4.00%
ITA	1982m1-2005m8	0.48%	5.23%
	1982m1-1985m6	1.61%	12.50%
	1985m7-1996m4	0.73%	5.38%
	1996m5-2005m8	-0.07%	2.32%
FRA	1986m4-2005m8	-0.59%	2.06%
	1986m4-1992m5	-0.91%	3.00%
	1996m6-2005m8	-0.46%	1.63%
NLD	1986m4-2005m8	-0.49%	2.02%
BEL	1985m10-2005m8	-0.42%	2.07%
	1985m10-1988m11	0.96%	1.52%
	1988m12-1994m9	-0.77%	2.87%
	1994m10-2005m8	-0.50%	1.81%
ESP	1986m6-2005m8	-1.13%	4.19%
	1986m6-1995m11	-1.69%	5.55%
	1995m12-2005m8	-0.65%	2.87%
DEU	1980m10-2005m8	-0.08%	2.33%

The Italian case reveals a consistent change in the direction of overprediction as inflation climbs to higher levels in older regimes. There is thus no sign of the existence of a “U-curve” – even inverted. In France, interpretation is complicated by the availability of only one break in the sample; the increase in underprediction here is possibly though not necessarily consistent with ADP’s theory. The same observation applies to Spain. This is also the case for Belgium when considering its last two regimes; the finding of overprediction for the initial low-inflation regime is inconsistent with the theory, though.

It therefore appears that the ADP hypothesis finds no more support on this broader sample of countries than in the original tests by Bryan and Palmqvist, which concerned Sweden and the US.

3.3 Testing Near-Rationality in the context of transitory changes in inflation

We now examine whether the ADP hypothesis can find support when considering all changes in inflation as being relevant for the theory, however transitory. The following nonlinear equation is estimated for this purpose:

$$\pi_{t,t-12}^e - \pi_t = \alpha + \beta_1 \pi_t + \beta_2 \pi_t^2 + \varepsilon_t \quad (68)$$

The null hypothesis consistent with ADP's hypothesis is:

$$H_0 : \begin{cases} \alpha = 0 \\ \beta_1 < 0 \text{ and } \beta_2 > 0 \\ |\beta_1| > \beta_2 \end{cases}$$

As could be expected from a model of this sort, significant autocorrelation and heteroscedasticity are found to be present in the residual term. This would invalidate the standard errors and thus inferences drawn on their basis. We therefore proceed to a correction of the standard errors according to the Newey-West procedure; the maximum number of lags allowed was set to 12, consistent with the monthly nature of the data and an inspection of the resulting correlograms from the previous regressions.

Table 49 presents the results from these tests.

TABLE 49: TESTING ADP NEAR-RATIONALITY, ALL FLUCTUATIONS

cty	Period	α	β_1	β_2	F
GBR [EC]	1981m11-2005m8	2.11 *** (0.659)	-0.578 * (0.340)	0.007 (0.030)	15.37 ***
GBR [GLP]	1983m1-1997m1	3.86 *** (0.822)	-0.413 (0.351)	-0.026 (0.030)	39.3 ***
ITA	1982m1-2005m9	-0.424 (0.676)	0.334 (0.279)	-0.019 (0.021)	1.05
FRA	1986m4-2005m8	0.464 ** (0.216)	-0.680 ** (0.314)	0.056 (0.082)	18.7 ***
NLD	1986m4-2005m8	0.746 *** (0.209)	-0.508 *** (0.171)	-0.041 (0.038)	49.2 ***
BEL	1985m10-2005m9	1.23 (0.747)	-0.705 (0.613)	-0.028 (0.116)	26.2 ***
ESP	1986m6-2005m8	-0.836 (1.11)	0.272 (0.658)	-0.086 (0.078)	14.5 ***
DEU	1980m10-2005m8	0.857 (0.303)	-0.492 * (0.272)	0.017 (0.044)	4.92 ***
SWE [SCB]	1979q1-2001m12	-0.302 (0.523)	0.239 (0.311)	-0.029 (0.028)	1.31

NOTES: *, **, *** respectively denotes significance at the 10%, 5% and 1% levels. Newey-West standard errors in parentheses, max lags 12. [GLP] refers to UK Gallup data; [SCB] refers to quantitative SCB data.

Figure 95 (Appendix II) presents scatter plots of the expectations errors for all countries together with the fitted values from equation (68) when using the EC Consumer Survey data (except in the case of Sweden, where the quantitative SCB data are used).

Table 49 makes clear that the null hypothesis performs poorly when contrasted to the data. Only in the case of three countries (the UK with EC data, France and Germany) do the signs of β_1 and β_2 match the null. This is hardly an endorsement of the theory, however. In no case is the β_2 coefficient significant; β_1 proves to be mostly insignificant or marginally significant for the most part; and several constants are found to be significant. Furthermore, many of the coefficients appear with the “wrong” sign.

The figures reinforce the message of the regression output visually. Even for those countries exhibiting the predicted concavity for the fitted curve, it is apparent that such fitted values would imply a requirement of unrealistically large rates of inflation for the under prediction of inflation to be reduced close to zero according to the mechanism posited by ADP. Moreover, an inspection of the figures suggests that the findings are robust to the use of a simple quadratic functional form – no alternative nonlinear specification is likely to yield a conclusion more favourable to ADP.

3.4 Conclusion

We have been able to confirm the empirical results obtained by Bryan and Palmqvist (2005)¹⁰¹ for Sweden and extended the analysis to the consideration of a further seven European countries. The results obtained here underscore the robustness of their findings.

Indeed, whether interpreting the ADP near-rationality hypothesis as applying to persistent changes in inflation only or to all such fluctuations, the data reject the behavioural view embodied in the theory.

As noted earlier, the ADP hypothesis is typical of much macroeconomic theory in that it fails to specify a complete data generation process, particularly in what regards the dynamics. Equation (68), for instance, can hardly be regarded as a complete DGP – hence the findings of autocorrelation and heteroscedasticity, *inter alia*. A successful model of the expectations error will likely look very different from the one proposed by these authors.

¹⁰¹ These authors also proceed to an illuminating analysis of the link between recent inflation experience in Sweden and the US and the underlying micro data on expectations. Interested readers are referred to their paper for further details. These issues are also dealt with in this study, in Section 6.6.4 below.

Indeed, the hypothesis of a U-shaped relation exclusively involving underprediction of inflation is at odds with the expectations data examined in this study. The data show that overpredictions are a common occurrence; furthermore, the formation of beliefs regarding inflation seems to rely more heavily on some sectors of the economy than others¹⁰², and the composition of this information set can change over time and depending on circumstances. Such considerations make clear and continuous relationships of the type suggested by ADP suspicious *ex ante*. Moreover, the expectations data are characterised by heavily adaptive behaviour, making the unemployment “bliss point” suggested by ADP a temporary heaven at best. More fundamentally, ADP oversimplify the problem of modelling expectations as one involving two clear-cut groups of agents (“rational” and “less rational”); the following section will show that such a simplification is unwarranted empirically.

Thus, we were unable to find support for the view that the long-run Phillips curve might not be vertical due to ADP-type expectations effects.

¹⁰² One aspect of this issue is dealt with in Section 5 below, namely: the relationship between the forecasts of experts and consumer inflation expectations.

4 FOUNDATIONAL MODELS OF INFLATION EXPECTATIONS

4.1 Introduction

This section aims to discuss and examine empirically a number of inflation expectations processes that are commonly used as building blocks in macroeconomic models. We term these “foundational” models of expectations. Concretely, the three models we consider are the adaptive, rational and rule-of-thumb expectations hypotheses. Important applications of each of these approaches abound. For instance, adaptive expectations were deployed in Nordhaus’ (1975) theory of the political business cycle; rational expectations are a key building block of credibility theory (Kydland and Prescott, 1977; Barro and Gordon, 1983); and rule-of-thumb expectations are now commonly used as part of empirical New Keynesian Phillips curves (Galí and Gertler, 1999).

Despite the ubiquity of these models as *assumptions* in economic studies and the frequent sensitivity of the conclusions to the particular expectations process assumed, critical examinations of the foundational models have been scarce. Indeed, as Carroll (2003) recently noted, “*There has been almost no work testing alternative models of expectations using actual empirical data on expectations*”¹⁰³. An exception to this statement is the voluminous output of the research project aiming to test rational expectations on the basis of survey data. In fact, such data have been used for little else and an area of diminishing returns seems to have long been entered. Nevertheless, some useful conclusions have been reached and are evoked below.

¹⁰³ Carroll (2003), p.270.

In this contribution, we propose to provide a brief overview of past theoretical and empirical developments associated with the rise (and demise) of the three key foundational models in the history of modern macroeconomics (Section 4.2). This is directly relevant in coming to an assessment of these models, and also helps situate some of the present work in Sections 4, 5 and 6 in its broader context. Section 4.3 provides a test of the adaptive expectations model for benchmarking purposes, while Section 4.4 analyses in some detail the robustness of the most popular expectations model in much current empirical research, namely the rule-of-thumb hypothesis. Sections 4.4.5 and 4.5 provide some concluding thoughts and highlight the broader implications of the results reached here.

4.2 Overview of key developments

4.2.1 From adaptive to rational expectations

The importance of expectations in economic analysis has long been recognised. Keynes famously emphasised their key role in determining investment, output and inflation in his *General Theory* (1936). He did not, however, provide a model of their formation. Until very recently in fact, the economic literature proved similarly reluctant to address this issue precisely. Instead, the revealed preference over the past few decades was to rely on a couple of simple conjectures on this matter – regularly advocating policies on their basis.

One such conjecture is the familiar “adaptive expectations hypothesis”. This was formally introduced by Cagan (1956), Friedman (1957) and Nerlove (1958). In terms of inflation expectations, it states that each period the previous expectations are updated by a fraction α of their deviation from observed outcomes to form current expectations:

$$\pi_{t,t+1}^e = \pi_{t-1,t}^e + \alpha(\pi_t - \pi_{t-1,t}^e) \quad (69)$$

This conjecture was used extensively in the late 1950s and 1960s until the “rational expectations revolution” gained momentum from the papers of Lucas (1972) and Sargent (1973). The idea of “rational expectations” is due to Muth (1961) who hypothesised that *“expectations, since they are informed predictions of future events are essentially the same as the predictions of the relevant economic theory”*. In this view, agents’ subjective expectations of economic variables coincide with the objective mathematical conditional expectations of those variables: agents know the true underlying model of the economy and use it to inform their expectations. The popularity of this method followed both from its simplicity and the appealing idea that it concerns economically “intelligent” agents that avoid systematic mistakes.

Unfortunately, the approach was rapidly found to be problematic from an empirical perspective. Tests of the rational expectations hypothesis can broadly be placed in two categories. First, the REH was tested *directly*, using survey data on inflation expectations of the type considered in this study. In fact, a whole testing ‘industry’ developed in the literature, making this the quasi-exclusive use for survey data on expectations over the past thirty years. A variety of countries, surveys, types of data (qualitative and quantitative), quantification methodologies (where applicable) and econometric techniques were considered. Econometrically, not all the approaches commonly used in these papers are commendable (see Curto Millet, 2006, for a critique). Notwithstanding this, and although this is no place to survey such a vast literature, it can be said that the overwhelming majority of studies rejected one aspect or the other associated with

rational expectations; see, inter alia: Roberts, 1997; Bakhshi and Yates, 1998; Łyziak, 2003; Nielsen (2003b); Forsells and Kenny (2006); Baghestani (1992); Hanssens and Vanden Abeele (1987); Figlewski and Wachtel (1981, 1983); Engsted (1991); Gerberding (2001, 2006); Thomas (1999); Papadia (1983); Kokoszcyński, Łyziak and Stanisławska (2006). Such findings are hardly unexpected given the behaviour of the expectations series documented in this study (e.g. Section 2.1.3) and the failure of the ‘rational perceptions hypothesis’ evidenced in Section 2.3 (which represents a new contribution to the literature).

Second, the REH was tested *indirectly*, by implication. For instance, the finding by Ericsson and Irons (1995) of the surprisingly limited relevance of the Lucas critique in key empirical applications is consistent with the formation of expectations that are not fully optimising, but may be highly appropriate when it comes to dealing with breaks in series (Hendry, 2002). However, the most influential blow to the theory of rational expectations has perhaps come from its widely noted empirical failure when implemented in the context of the so-called New Keynesian Phillips Curve (NKPC) model. Of course, this failure is that of a joint hypothesis – of expectations being rational, and of the NKPC model being appropriate. Given the import of findings from the NKPC literature on that concerning expectations, it seems worthwhile to address these in some detail.

4.2.2 The New Keynesian Phillips Curve literature and further developments

A standard derivation of the NKPC model is based on the Calvo (1983) pricing assumption that in every period each firm faces the same probability of having a price adjustment opportunity, in which case it sets its price to its average desired level until the next adjustment. Denoting the output gap by y_t , the New Keynesian Phillips curve can be written:

$$\Pi_t = [\alpha\lambda^2 / (1 - \lambda)]y_t + E_t\Pi_{t+1} \quad (70)$$

Several well-known problems plague this formulation. The model predicts that announced, credible disinflations produce booms (Ball, 1994), it is incapable of generating inflation persistence (Fuhrer and Moore, 1995) and more generally the restrictions it imposes across impulse response functions fail to match the evidence (Mankiw, 2000). More inertia needs to be added to the above equation if the model is to shed these problems, and at least two ways related to expectations have been explored to achieve this.

Rule-of-Thumb Expectations

The first approach is based on the diagnosis that the problems the model faces arise from the assumption of rational expectations. It is *because* firms are rational that the ones having an adjustment opportunity cut their prices immediately in the face of an announced credible disinflation ahead, thereby increasing the real money stock and causing a boom. This does not happen if firms' expectations are adaptive, which also introduces inflation persistence and realistic impulse response cross-restrictions. The

model above then reduces to the traditional backward-looking model (Friedman, 1968), which is overall more successful in its data confrontation.

Given this requirement for empirical success, researchers unwilling to fully abandon the rational expectations hypothesis have attempted to strike an ad hoc compromise. A number of Phillips curves have been fitted under the assumption that there is a fraction of ‘forward-looking’ agents acting as above and a fraction of ‘backward-looking’ ones, thus introducing the required inertia. Prominent among these studies has been Galí and Gertler (1999), whose authors estimate the following hybrid Phillips curve, with s_t being a measure of real marginal costs:

$$\Pi_t = \lambda s_t + \gamma_f E_t \{\Pi_{t+1}\} + \gamma_b \Pi_{t-1} \quad (71)$$

One of their strongest results is that although backward-looking behaviour is statistically significant, their estimates suggest that forward-looking behaviour is followed by roughly sixty to eighty percent of firms. This is a tremendously forward-looking finding. However, it is plagued by a number of econometric difficulties. Ruud and Whelan (2001) show that conventional models can easily generate such results as a product of omitted variable bias. Mavroeidis (2002) argues that weak identification due to an AR(1) process for s_t can lead to upward biases in small samples for the forward-looking coefficient. Furthermore, when tested against a backward-looking EqCM specification on Norwegian data, this model is strongly rejected (Bardsen et al., 2002).

In any event, a major limitation of results in this area is that they depend on a joint hypothesis – that of ‘rule-of-thumb’ expectations and a given Phillips curve specification. To address this issue, we will consider the direct testing of the rule-of-thumb model on survey data in Section 4.4.

Sticky Information

More recently, Mankiw and Reis (2002; henceforth MR) proposed a model that is able to retain rational expectations whilst correcting the counterfactual predictions of the previous New Keynesian Phillips curves. They posited that information diffuses slowly in the economy. This idea is modelled by letting a proportion λ of firms in each period obtain new information and compute new optimal paths for prices on this basis, whilst the remaining firms continue to base their decisions on outdated information. The similarities with Calvo pricing are obvious, but the model results in a Phillips curve of the form:

$$\Pi_t = \left[\frac{\alpha\lambda}{1-\lambda} \right] y_t + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-1-j} (\Pi_t + \alpha\Delta y_t) \quad (72)$$

The striking feature of the equation is the type of inertia introduced, feeding through *past* expectations about the future. MR conduct a series of policy experiments and find sensible impulse response functions, broadly similar to those obtainable under a simple backward-looking model.

We consider direct tests of the expectations formation structure hypothesised by MR (i.e. gradual updating to rational expectations using the latest information) in Section 6.2. We will also examine a variation of the model authored by Carroll (2003) in Section 5.4.

Our next task, however, is to consider the empirical results associated with the simple adaptive expectations model, to establish a benchmark.

4.3 The Adaptive Expectations model

Rewriting equation (69) in the context of monthly data and rearranging yields the following useful formulation of the adaptive expectations hypothesis:

$$\pi_{t,t+12}^e = \alpha \pi_{t-12,t} + (1 - \alpha) \pi_{t-12,t}^e \quad (73)$$

We generalise this model further before taking it to the data by adding a constant and removing the implied constraint on its coefficients:

$$\pi_{t,t+12}^e = \alpha_0 + \alpha_1 \pi_{t-12,t} + \alpha_2 \pi_{t-12,t}^e + \varepsilon_t \quad (74)$$

The results are presented in Table 50. The coefficient constraint implied by the adaptive expectations model (namely, $\alpha_1 + \alpha_2 = 1$) is examined with an F-test under column “Test”.

TABLE 50: THE ADAPTIVE EXPECTATIONS MODEL

Equ.	Coefficients			Specification tests						
	α_0	α_1	α_2	R^2	NOBS	Autocorr.	Het. Test	ARCH	Normal	Test
FRA	0.006 * (0.003)	0.589 *** (0.141)	0.127 (0.153)	0.48	209	164.0 *** [0.00]	2.5 [0.78]	88.8 *** [0.00]	3.37 [0.19]	3.12 * [0.08]
ESP	0.006 (0.005)	0.669 *** (0.091)	0.236 ** (0.103)	0.63	207	166.5 *** [0.00]	10.6 * [0.06]	124.0 *** [0.00]	0.78 [0.68]	0.57 [0.45]
SWE	0.011 ** (0.005)	0.368 ** (0.146)	-0.661 (0.500)	0.16	92	74.4 *** [0.00]	16.6 *** [0.01]	60.2 *** [0.00]	7.57 ** [0.02]	7.56 *** [0.01]
GBR	0.017 *** (0.006)	0.653 *** (0.217)	-0.124 (0.167)	0.35	262	244.0 *** [0.00]	67.3 *** [0.00]	212.4 *** [0.00]	8.27 ** [0.02]	7.26 *** [0.01]
ITA	0.009 *** (0.002)	0.538 *** (0.114)	0.146 ** (0.079)	0.80	257	234.5 *** [0.00]	46.7 *** [0.00]	206.2 *** [0.00]	0.86 [0.65]	40.9 *** [0.00]
DEU	0.010 *** (0.003)	0.679 *** (0.091)	-0.204 ** (0.102)	0.44	275	229.7 *** [0.00]	21.0 *** [0.00]	226.3 *** [0.00]	25.8 *** [0.00]	75.7 *** [0.00]
NLD	0.015 *** (0.003)	0.519 *** (0.086)	-0.206 (0.189)	0.32	209	168.8 *** [0.00]	10.0 * [0.07]	149.4 *** [0.00]	19.0 *** [0.00]	17.1 *** [0.00]
BEL	0.016 *** (0.005)	0.373 ** (0.146)	-0.158 (0.160)	0.14	216	175.1 *** [0.00]	12.4 ** [0.03]	102.1 *** [0.00]	4.32 [0.12]	13.1 *** [0.00]

NOTE: *, **, *** respectively denote significance at the 10%, 5% and 1% levels. Newey-West correction applied to the standard errors, with twelve lags. P-values provided in brackets [] and standard errors provided in parentheses (). The autocorrelation test corresponds to the Breusch-Godfrey test for higher-order serial correlation with twelve lags. “Het. Test” corresponds to White’s (1980) test for heteroscedasticity; “ARCH” corresponds to an LM test for autoregressive conditional heteroscedasticity; “Normal” corresponds to a joint Chi2(2) test for skewness and kurtosis.

As can be seen, the model performs very poorly in terms of specification diagnostics, with almost every single ‘ailment’ possible proving highly significant. The Newey-West correction is used to adjust the standard errors for heteroscedasticity and autocorrelation. This is an ‘academic’ exercise for most countries, as the widespread failure of the diagnostics suggests that inference on the basis of this model should not be considered reliable. Indeed, counter-theoretically, the majority of α_2 coefficients are found to be *negative*. However, the models for France, Spain and Italy seem to perform better than the others – normality of the residuals cannot be rejected, and the coefficients have the correct signs. In the case of Spain it is – exceptionally – even possible to accept a homogeneity restriction for the coefficients of the explanatory variables, and the constant appears to be insignificant. Nevertheless, the findings of autocorrelation and

heteroscedasticity suggest that even in those more optimistic cases, the model is at most an incomplete story – a full one possibly involving more explanatory variables, and a different functional form.

We now consider whether the “rule-of-thumb” approach is met with more success in its empirical confrontation.

4.4 Investigating the Rule-of-Thumb approach

4.4.1 Presenting the Gerberding (2001) results

The ‘rule-of-thumb’ hypothesis is based on the idea that there is a fraction of agents in the population acting rationally, with the remainder doing ‘something else’ (typically, some form of adaptive expectations). The theoretical and empirical popularity of this approach as a convenient way to solve the problems of pure rational expectations mechanisms calls for further attention.

We will in particular examine in detail the results obtained by Gerberding (2001), who estimates a model of partly rational and partly adaptive expectations for Germany, Italy and France with the following form:

$$E_t \pi_{t+4} = c_1 E_t^m \pi_{t+4} + (1 - c_1)(E_{t-4} \pi_t + c_2 (\pi_t - E_{t-4} \pi_t))$$

$E_t^m \pi_{t+4}$ is the “mathematical” expectation of future inflation, whilst c_2 captures the speed at which agents revise their expectations. Gerberding argues that it is realistic to assume that only *lagged* inflation is known at the time of expectations formation. Furthermore, the rational expectation can be rewritten as the actual future inflation rate minus an expectation error u_{t+4} .

Highlighting the use of survey data with the superscript S, this yields:

$$E_t^S \pi_{t+4} = c_1 \pi_{t+4} + (1 - c_1)(E_{t-5}^S \pi_{t-1} + c_2(\pi_{t-1} - E_{t-5}^S \pi_{t-1})) + \varepsilon_t^S - c_1 u_{t+4} \quad (75)$$

It is important to note that the author quantifies the data from the EC Consumer Survey according to the method labelled “EXPINFLA” in Chapter 1. That is, she uses the Batchelor and Orr (1988) modification of the Carlson-Parkin tradition in combination with the assumption that consumers correctly perceive the rate of inflation experienced over the previous year¹⁰⁴. This is a strong assumption, and in any event the measure was found to be sub-optimal in Chapter 1. We will however use it initially here in order to replicate her results.

To deal with the bias that would arise from the correlation between π_{t+4} and the error term, the author uses lagged values (up to 12 lags) of the inflation rate as instruments but refrains from using lagged values of the endogenous variable. This is due to the possibility of measurement error in the survey data introducing a correlation between the lagged endogenous variable and the disturbance. Gerberding found the following coefficients:

TABLE 51: GERBERDING (2001) RULE-OF-THUMB MODEL; 1991q1-1999q4

	DEU	FRA	ITA
C₁	0.22*	0.30**	{0}
1-C₁	{0.78}	{0.70}	{1.00}
C₂	1.18**	1.01**	0.93**
Constant C₀	-	-0.45**	-

Note: * & ** denote significance at the 5% and 1% level, respectively; Brackets {} denote restricted variables.

¹⁰⁴ Please refer to Chapter 1, Section 3.1 for a detailed exposition of the approach.

The estimation method used was two-stage least squares with Newey-West correction of standard errors. In no case can the tests of overidentifying restrictions reject the null.

It is especially striking that Italy can be restricted to an entirely backward-looking specification, which seems at odds both with its inflationary history and the evidence in Chapter 2, as well as past empirical evidence (e.g. Barrell and Dury, 2001; Curto Millet, 2004). We note, however, that in the regression presented for Italy, the first order autocorrelation coefficient appears to be highly significant which may indicate misspecification.

We further note that the sample period chosen for the German regression implicitly makes use of some data applicable to West Germany only through the lagged dependent variable on the right-hand side of the equation, which may be problematic.

Similar work has been carried out by Heinemann and Ullrich (2004) in the context of data on German inflation expectations, this time from the ZEW's *Finanzmarkttest* dataset, a monthly survey of financial market experts over the period 1993:8-2003:1. The data are quantified in the CP tradition, and the authors estimate the same model for inflation expectations as deployed by Gerberding (2001), with the difference that their survey deals with a 6-month horizon. Their point estimates for c_1 and c_2 are, respectively, 0.41 and 1.06. The authors argue on the basis of recursive results that there is no evidence for a structural break in the inflation expectations process associated with EMU; we consider both the methodology used and the conclusion doubtful, and take up

this issue in Section 4.4.5. Before doing so, it is useful to reconsider the Gerberding (2001) results.

4.4.2 Replicating the Gerberding (2001) results

First, we aim to replicate the Gerberding results as a benchmark for our data over an identical time sample, taking the other countries that are the subject of this study into consideration as well.

We also make use of the same data source for expectations (the EC Consumer Survey) and the same quantification technique (leading to the series labelled ‘EXPINFLA’), up to a small technical difference. Indeed, the quantification here is based on the assumption that the underlying distribution is normal, whereas Geberding uses a logistic distribution. The differences between the two methods have been found to be minor in previous research (e.g. Nielsen, 2003a; Berk, 1999) and we do not expect them to lead to qualitative changes in the results. The CPI data used may also be slightly different – we employ OECD data concerning all goods and services; the data source is not specified in this respect in Gerberding (2001). Table 52 presents the output of our replication exercise:

TABLE 52: REPLICATION OF GERBERDING’S EQUATIONS; 1991q1-1999q4

	DEU	FRA	ITA
C₁	0.285** (0.054)	0.275** (0.044)	{0}
1-C₁	{0.715}	{0.725}	{1.00}
C₂	1.11** (0.039)	1.04** (0.09)	1.05** (0.166)
Constant C₀	-	-0.005** (0.0003)	-

Note: * & ** denote significance at the 5% and 1% level, respectively. HAC-consistent standard errors. Brackets {} denote restricted variables.

As can be seen, our results are very close and insignificantly different from the point estimates of Gerberding (Table 51), except in the case of the French constant which is considerably smaller in our estimation. Table 53 provides the diagnostics arising from these regressions, while Figure 96 plots the fitted and actual values, as well as the standardised residuals.

TABLE 53: DIAGNOSTICS FOR GERBERDING RULE-OF-THUMB MODEL

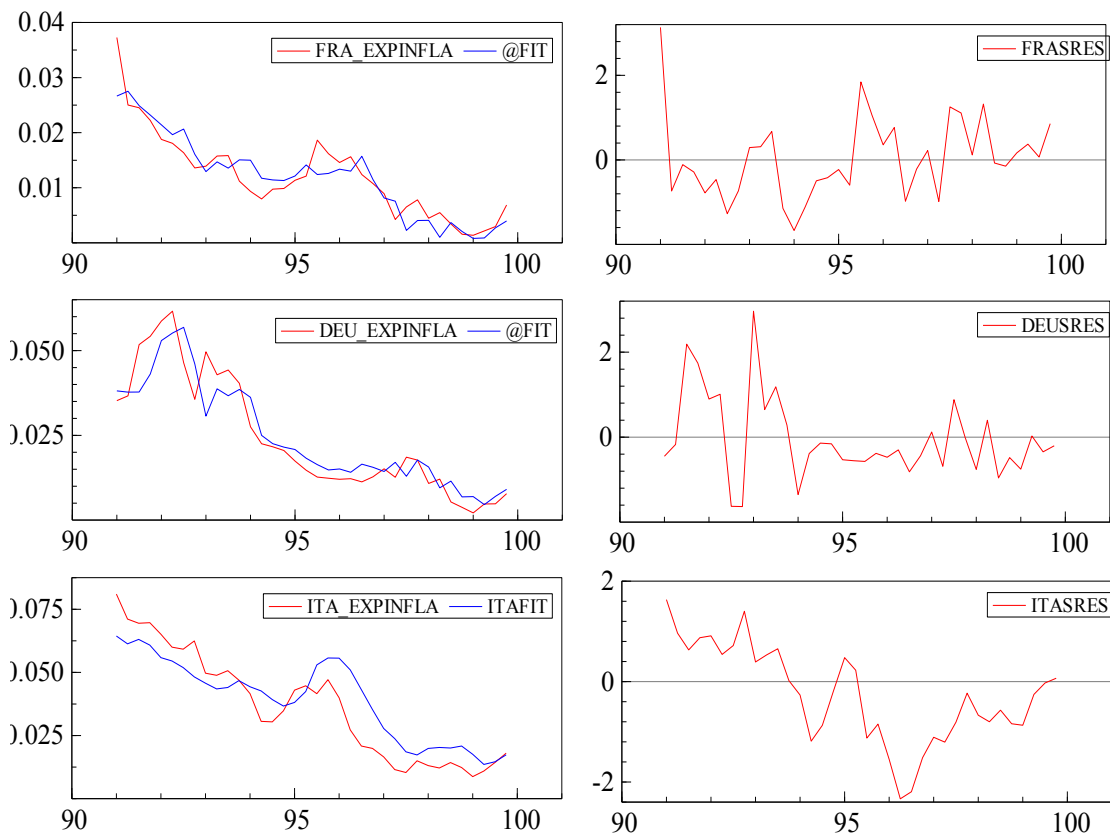
	DEU	FRA	ITA
R²	0.878	0.821	0.498
AC(1)⁺	0.292	0.279	0.763
AC(5)⁺	0.261	0.013	0.165
AC(9)⁺	-0.128	0.291	0.304
DW	1.60129 [.000,.907]	1.36725 [.000,.692]	.275188 [.000,.000]
P-value, test for overidentifying restrictions	0.939	0.902	0.962

+ ACs are residual autocorrelation coefficients. Bartlett standard error is $1/\sqrt{T}$, that is 0.17 here as T=36. For DW, the P-values relate to Durbin's h and Durbin's alternative statistic¹⁰⁵.

We find that only the first order Italian residual autocorrelation appears to be significant, confirming Gerberding's own results on this matter. This is also visible from Durbin's alternative h-statistic.

¹⁰⁵ Both of these statistics apply to the case when the right-hand side includes lagged dependent variables (LDVs). The h-statistic concerns the case of a single LDV, whilst the alternative statistic is applicable when there is one or more LDVs. The statistics are asymptotically equivalent, although the alternative may have low power in small samples.

FIGURE 96: GERBERDING MODEL, FIT AND STANDARDISED RESIDUALS



Of the models presented in Gerberding, only the French one is estimated with a constant, which is interpreted by the author as “*indicating either that a measurement error has occurred with respect to the level of the expectations variable [...] or that the formulated model does not entirely capture the expectation formation process of the polled households*”¹⁰⁶. In Table 54, we check whether we can confirm that no constants are needed for Italy and Germany, and estimate the model for the other countries in our sample except Sweden, due to insufficient observations over this particular period. Constants are only retained if found to be significant.

¹⁰⁶ See Gerberding (2001), p.31.

Coefficient C_1 was notably restricted to zero for the Netherlands. This is not due to its being insignificant (as is the case for Italy) but rather because it appears to be significantly negative and quantitatively large (-1.31) if left unrestricted, in contradiction with the model. As can be seen, we find that Germany is the only country for which a constant is not required in the specification, suggesting either widespread measurement error, misspecification or both. The diagnostics in Table 55 reveal significant first-order autocorrelations in the cases of Italy, Spain and the UK, reinforcing the suspicion of misspecification. Overidentifying restrictions cannot be rejected for any country.

TABLE 54: RULE-OF-THUMB MODEL, ALL COUNTRIES, 1991q1-1999q4

	DEU	FRA	ITA	GBR
C_1	0.285** (0.054)	0.275** (0.044)	{0}	0.368** (0.037)
$1-C_1$	{0.715}	{0.725}	{1.00}	{0.632}
C_2	1.11** (0.039)	1.04** (0.09)	0.943** (0.131)	1.17** (0.048)
Constant C_0	-	-0.005** (0.0003)	-0.00289** (0.000635)	-0.00350**

	ESP	BEL	NLD	SWE
C_1	0.468** (0.146)	0.256** (0.040)	{0}	-
$1-C_1$	{0.532}	{0.744}	{1.00}	-
C_2	0.584** (0.152)	0.971** (0.064)	1.04** (0.043)	-
Constant C_0	-0.010** (0.00104)	-0.00394** (0.00035)	-0.00177** (0.00021)	-

Note: * & ** denote significance at the 5% and 1% level, respectively. HAC-consistent standard errors. {} denote restricted variables.

TABLE 55: RULE-OF-THUMB MODEL, DIAGNOSTICS, 1991q1-1999q4

	DEU	FRA	ITA	GBR
R²	0.878	0.821	0.498	0.814
AC(1)⁺	0.292	0.279	0.752	0.393
AC(5)⁺	0.261	0.013	0.060	-0.055
AC(9)⁺	-0.128	0.291	0.273	-0.236
DW	1.601 [.000,.907]	1.367 [.000,.692]	0.316 [.000,.000]	1.298 [.000,.601]
P-value, test for overidentifying restrictions	0.939	0.902	0.925	0.893

	ESP	BEL	NLD	SWE
R²	0.832	0.839	0.752	-
AC(1)⁺	0.684	0.011	0.053	-
AC(5)⁺	-0.142	0.091	0.147	-
AC(9)⁺	-0.294	0.069	0.153	-
DW	0.818 [.000,.055]	2.01 [.003,.997]	0.955 [.000,.150]	-
P-value, test for overidentifying restrictions	0.909	0.907	0.942	-

+ ACs are residual autocorrelation coefficients. Bartlett standard error is $1/\sqrt{T}$, that is 0.17 here as T=36. For DW, the P-values relate to Durbin's h and Durbin's alternative statistic.

In qualitative terms, these regressions suggest that expectations formation is not very forward-looking for this sample of countries – the highest value for C1 is found in Spain, 0.468. On the other hand, the adaptive part of the expectations mechanism appears to be strongly responsive to past forecast errors, with coefficients of around 1 – except again for Spain, which displays a value of 0.584.

Figure 97 (Appendix II) plots fitted and actual values of the new regressions presented in Table 54, whilst Figure 98 (Appendix II) displays the corresponding standardised residuals. Only Germany and France present punctually large standardised residuals (around 3 standard errors in size; see Figure 96).

4.4.3 Sample sensitivity

Our aim now is to check the robustness of the model in terms of its stability by examining the sensitivity of the results to the sample period chosen. We therefore re-estimate the models obtained above for the largest sample period available in each country. In Germany, we limit the sample to the post-1991 data, as earlier values refer to West Germany only, and the transition could potentially be an independent source of instability.

TABLE 56: RULE-OF-THUMB MODEL, FULL SAMPLE ESTIMATION

	DEU	FRA	ITA	GBR
C₁	0.381** (0.067)	-0.070 (0.126)	{0}	0.226** (0.054)
1-C₁	{0.619}	{1.07}	{1.00}	{0.774}
C₂	1.00** (0.095)	0.832** (0.067)	1.29** (0.121)	1.13** (0.035)
Constant C₀	-	-0.00597** (0.00064)	0.013* (0.0061)	-0.0035** (0.0012)
Period	1992q2-2004q2	1986q2-2004q2	1974q2-2005q2	1981q3-2004q2

	ESP	BEL	NLD	SWE
C₁	0.04 (0.087)	0.228** (0.061)	{0}	0.011 (0.054)
1-C₁	{0.96}	{0.772}	{1.00}	{0.989}
C₂	0.933** (0.045)	1.00** (0.038)	1.20** (0.10)	0.753** (0.026)
Constant C₀	-0.013** (0.00062)	-0.0043** (0.00039)	-0.0033* (0.0014)	-0.0019** (0.00041)
Period	1987q4-2004q2	1986q2-2004q3	1975q1-2005q2	1997q2-2004q2

Note: * & ** denote significance at the 5% and 1% level, respectively. HAC-consistent standard errors. {} denote restricted variables.

TABLE 57: RULE-OF-THUMB MODEL, DIAGNOSTICS, FULL SAMPLE

	DEU	FRA	ITA	GBR
R²	0.849	0.698	0.640	0.864
AC(1)⁺	0.130	0.462	0.864	0.601
AC(5)⁺	0.050	-0.024	0.383	0.068
AC(9)⁺	-0.142	0.243	0.415	-0.203
DW	1.54 [.000,.633]	1.00 [.000,.001]	0.365 [.000,.000]	0.751 [.000,.000]
P-value, test for overidentifying restrictions	0.852	0.603	0.734	0.662
Observations	49	73	125	92

	ESP	BEL	NLD	SWE
R²	0.852	0.736	0.778	0.706
AC(1)⁺	0.269	0.099	0.694	0.128
AC(5)⁺	0.228	0.059	0.158	-0.321
AC(9)⁺	0.213	0.130	-0.067	0.258
DW	1.39 [.000,.163]	1.62 [.001,.431]	0.60 [.000,.000]	1.65 [.000,.987]
P-value, test for overidentifying restrictions	0.684	0.594	0.494	0.930
Observations	67	74	122	29

+ ACs are residual autocorrelation coefficients. Bartlett standard error is $1/\sqrt{T}$. For DW, the P-values relate to Durbin's h and Durbin's alternative statistic.

Contrasting Table 56 with Table 54 reveals a number of important differences in coefficients for all countries, except Belgium, the Netherlands and Germany (although in the latter's case, C1 has substantially increased). For both Spain and France, the coefficient C1 – formerly significant at the 1% level – has collapsed and is now insignificant; C2 is also significantly different (smaller for France, larger for Spain). For Italy, both C2 and C0 appear to be significantly higher in the full sample. The UK's coefficient C1 is significantly smaller here.

There is now also widespread evidence of misspecification, with Spain, the UK, the Netherlands and France all exhibiting significant first-order autocorrelation¹⁰⁷. For Italy, all the autocorrelation coefficients are now significant.

Ignoring the specification problems, the qualitative message of these regressions seems to be that none of the countries considered displays very forward-looking expectations – the highest value of C1 being achieved by Germany, 0.381. On the other hand, the adaptive correction for past forecast errors appears to be very strong, the lowest C2 coefficient being 0.753 for Sweden, with most of the others being around 1.00. Thus, despite the important changes due to sample periods, the qualitative message is broadly similar to that discussed earlier.

4.4.4 Sensitivity to expectations measure

We now aim to check the impact of using a different quantification methodology on the results of the rule-of-thumb model. To verify the level of robustness along this dimension, we keep the original sample of 1991q1-1999q4 considered by Gerberding (2001) and use the inflation expectations measure EXPINFLPBIS (the preferred measure in this study) instead of EXPINFLA. Table 58 presents the results of this exercise.

¹⁰⁷ France also exhibits a significant 9th order autocorrelation, although it is not clear that this is meaningful.

TABLE 58: RULE-OF-THUMB MODEL, EXPINFLPBIS, 1991q1-1999q4

	DEU	FRA	ITA	GBR
C_1	0.597** (0.051)	0.539** (0.079)	{0}	0.725** (0.094)
$1-C_1$	{0.403}	{0.461}	{1.00}	{0.275}
C_2	0.077 (0.112)	0.520** (0.168)	0.98** (0.187)	0.507 (0.272)
Constant C_0	-	-0.00565** (0.0010)	-0.00176 (0.00109)	0.00456** (0.00113)

	ESP	BEL	NLD	SWE
C_1	0.533* (0.225)	0.212** (0.053)	{0}	-
$1-C_1$	{0.467}	{0.788}	{1.00}	-
C_2	0.204 (0.363)	0.168** (0.051)	0.481** (0.068)	-
Constant C_0	-0.00961** (0.00133)	-0.00267** (0.00023)	-0.00359** (0.00067)	-

Note: * & ** denote significance at the 5% and 1% level, respectively. HAC-consistent standard errors. {} denote restricted variables.

TABLE 59: DIAGNOSTICS FOR EXPINFLPBIS RULE-OF-THUMB MODEL

	DEU	FRA	ITA	GBR
R^2	0.764	0.598	0.538	0.482
$AC(1)^+$	0.644	0.617	0.691	0.722
$AC(5)^+$	-0.645	0.144	0.085	0.090
$AC(9)^+$	-0.060	0.497	0.120	0.422
DW	1.03 [.000,.233]	0.687 [.000,.014]	.635 [.000,.007]	.478 [.000,.001]
P-value, test for overidentifying restrictions	0.947	0.945	0.923	0.889

	ESP	BEL	NLD	SWE
R^2	0.605	0.813	0.238	-
$AC(1)^+$	0.788	0.307	0.818	-
$AC(5)^+$	-0.164	-0.108	0.229	-
$AC(9)^+$	-0.200	-0.109	-0.237	-
DW	.677 [.000,.013]	.862 [.000,.079]	.341 [.000,.000]	-
P-value, test for overidentifying restrictions	0.948	0.922	0.978	-

+ ACs are residual autocorrelation coefficients. Bartlett standard error is $1/\sqrt{T}$, that is 0.17 here as $T=36$. For DW, the P-values relate to Durbin's h and Durbin's alternative statistic.

Comparing Table 58 and Table 54 reveals significant differences in coefficients for all the countries considered. For Germany and the UK, C1 becomes significantly higher whilst C2 drops into insignificance (whereas it was significant at the 1% level earlier). The same findings are made for France, except that C2 remains significant. In Italy C0 becomes insignificant despite its earlier 1% significance. C0 increases significantly in the UK. In Spain, C1 increases (but is now only significant at the 5% level instead of 1%) whilst C2 drops to insignificance from its earlier highly significant status (at 1%). For Belgium and the Netherlands, C2 becomes significantly lower, and significant changes also occur for C0 (positively and negatively, respectively).

Residual autocorrelation is detected for all countries except Belgium. All of these display first-order autocorrelation. In addition, Germany has a significant 5th order autocorrelation coefficient, whilst France and Italy have significant 9th order autocorrelation coefficients. It is unclear whether these higher-order results are meaningful, however, or whether they should be attributed to overfitting of the autocorrelation test, for instance.

In qualitative terms, the new results provide a greatly modified interpretation relative to that applicable for the former measure of inflation expectations. Important increases were detected in all the countries in which we allowed the C1 coefficient to vary, except Belgium where it remained roughly stable. The coefficient is the largest for the UK, with a value of 0.725. Furthermore, there are sharp drops of the C2 coefficient in all countries except Italy. The predominant picture is now one of more balanced forward and

backward-looking behaviour, and a considerably lower speed of adaptation to past forecast errors in the adaptive part of the mechanism.

In Table 60 and Table 61 we repeat the above exercise for the quantitative expectations data available for the UK and Sweden, and compare the results with those obtained from the quantified expectations data (with the method EXPINFLA used by Gerberding and the method EXPINFLPBIS recommended in this study). The same observations again stand out: the results are highly sensitive to the sample period chosen and the expectations measure used. We also note that although not a perfect replica, EXPINFLPBIS appears overall to give closer results to the quantitative data than EXPINFLA.

TABLE 60: RULE-OF-THUMB MODEL, QUANTITATIVE DATA

SWEDEN	EXPSCBEV	EXPSCBEV	EXPINFLA	EXPINFLPBIS
Period	1980:4-2004:2	1997:2-2004:2	1997:2-2004:2	1997:2-2004:2
Observations	95	29	29	29
C₁	0.893** (0.136)	0.088** (0.024)	0.011 (0.054)	0.077** (0.029)
1-C₁	{0.107}	{0.912}	{0.989}	{0.923}
C₂	1.74 (1.95)	0.418** (0.036)	0.753** (0.026)	0.230** (0.013)
Constant C₀	0.00345 (0.0023)	0.00424** (0.00040)	-0.0019** (0.00041)	0.000096 (0.00017)

UK	EXPGLP	EXPINFLA	EXPINFLPBIS
Period	1984:2-1996:4	1984:2-1996:4	1984:2-1996:4
Observations	51	51	51
C₁	0.253** (0.035)	0.194** (0.074)	0.446** (0.076)
1-C₁	{0.747}	{0.806}	{0.554}
C₂	0.425** (0.034)	1.14** (0.035)	0.448** (0.080)
Constant C₀	0.00639** (0.00072)	-0.00168 (0.00092)	-0.000478 (0.00122)

Note: * & ** denote significance at the 5% and 1% level, respectively. HAC-consistent standard errors.

TABLE 61: RULE-OF-THUMB, DIAGNOSTICS, QUANTITATIVE DATA

SWEDEN	EXPSCBEV	EXPSCBEV	EXPINFLA	EXPINFLPBIS
R²	0.777	0.748	0.706	0.810
AC(1)⁺	0.797	0.295	0.128	0.285
AC(5)⁺	0.104	-0.634	-0.321	-0.732
AC(9)⁺	0.031	0.213	0.258	0.384
DW	0.523 [.000,.000]	1.40 [.000,.910]	1.65 [.000,.987]	1.44 [.000,.932]
P-value, test for overidentifying restrictions	0.875	0.923	0.930	0.931

UK	EXPGLP	EXPINFLA	EXPINFLPBIS
R²	0.891	0.844	0.660
AC(1)⁺	0.523	0.606	0.592
AC(5)⁺	-0.264	0.034	0.114
AC(9)⁺	-0.225	-0.315	-0.113
DW	1.02 [.000,.030]	0.824 [.000,.002]	0.813 [.000,.002]
P-value, test for overidentifying restrictions	0.877	0.878	0.833

+ ACs are residual autocorrelation coefficients. Bartlett standard error is $1/\sqrt{T}$. For DW, the P-values relate to Durbin's h and Durbin's alternative statistic.

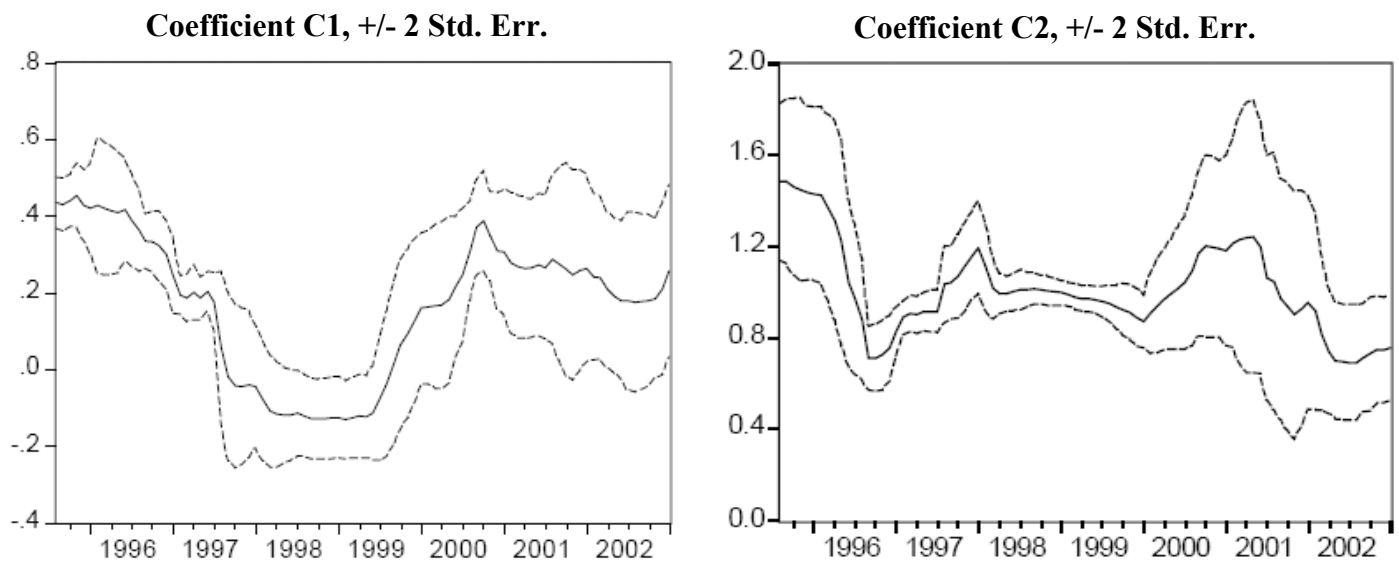
4.4.5 Conclusion and further implications of our 'rule-of-thumb' results

As could only be expected from a simple reduced-form model of the type considered here, evidence of coefficient instability and misspecification is not hard to find, making such models fragile and inappropriate descriptions of consumer inflation expectations. In particular, we have provided evidence of severe sample-dependence in the results as well as measure-dependence. We find it impossible to draw firm conclusions regarding the relative degree of forward and backward-lookingness from this model, for instance. Here as elsewhere, the General-to-Specific modelling approach advocated by Hendry (1995) would appear to be invaluable.

Our results have some very practical consequences for empirical research. In particular, they strongly suggest that inferences based on the simple rule-of-thumb model are likely to be very poor and unreliable.

For instance, Heinemann and Ullrich (2004) tried to assess the impact of the transition from the Bundesbank's to the ECB's credibility on German inflation expectations using the same model as Gerberding (2001). The authors proposed to do this by conducting a sample-splitting exercise and by studying the recursive behaviour of the coefficients of this rule-of-thumb model. The period considered spanned 1993m8-2003m1. Even within this relatively limited period, the recursive estimates clearly showed a number of important breaks. Figure 101 illustrates this by reproducing the recursive plots presented in their paper.

FIGURE 101: HEINEMANN AND ULLRICH (2004) RECURSIVE COEFFICIENTS



SOURCE: Heinemann and Ullrich (2004), Figure 6a (left panel) and Figure 6b (right panel), p.19.

Nevertheless, the authors eventually concluded that “*the regime change is not associated with a significant structural break in the formation of inflation expectations*”¹⁰⁸. Taken literally, this statement is clearly contradicted by the data. The authors probably meant (as they emphasise earlier) that the breaks experienced around the advent of EMU acted in opposite directions, so that by the end of the sample coefficient C1 has ‘reverted’ to its earlier values¹⁰⁹.

Be that as it may, no sound conclusion can actually be drawn from this model. If anything, the results of Heinemann and Ullrich (2004) confirm the chronic instability of models of this type documented above, and their failure to achieve congruence (e.g. in terms of autocorrelation¹¹⁰, and even in terms of a rejection of overidentifying restrictions in the case of their first subsample, at the 0.1% significance level).

Thus, it seems impossible to draw robust conclusions from exercises involving these types of rule-of-thumb models as one can actually *expect* instability in the equation, and therefore biased results – especially when the interpretation is based *precisely* on the stability properties of the equation.

¹⁰⁸ Heinemann and Ullrich (2004), p.20.

¹⁰⁹ Heinemann and Ullrich (2004), p.19.

¹¹⁰ We note, however, that this finding may be related to an overlapping data issue and is not by itself necessarily indicative of model failure.

4.5 Conclusion

The mainstream economic literature has cycled through a number of hypotheses concerning the formation of inflation expectations – from adaptive to rational, and from rational to rule-of-thumb or composite expectations. Rational expectations have been tested in countless papers, both directly and indirectly, and have been found to be wanting. Adaptive expectations prove similarly disappointing when the model is applied to survey data, as in Section 4.3. Even the ‘rule-of-thumb’ model that has become so popular in much theoretical and applied work (especially in the context of the New Keynesian Phillips Curve) appears to be inappropriate – notably proving to be both sample- and measure-dependent, among other signs of misspecification. In some respects, this is not surprising and brings further weight to this side of the case in a trial in progress. Indeed, Roberts (1997) had already noted on the basis of US survey data that the observed deviations from rationality could not be adequately captured by a simple weighted-average of rational and adaptive expectations. The warning, however, appears to have been largely unheeded given the continued use of this ‘trick’ in empirical work (e.g. Galí and Gertler, 1999).

The processes modelled here are very complex and it is admittedly difficult to specify a satisfactory conditional density for inflation expectations. The ‘animal spirits’ property of inflation expectations means that they will easily be subject to breaks in ways that cannot readily be predicted by econometric models (consider, for instance, the descriptive analysis of inflation expectations in Section 2.1.3 above). Prediction of such processes, in these circumstances, is likely to be particularly hazardous. However, there is certainly still scope for econometric models to shed some light on the factors important

in the formation of expectations. Indeed, although the “foundational” models of inflation expectations do not fare well when contrasted to the data, other interesting theoretical and empirical developments have taken place recently.

For instance, Branch (2004a,b) takes the concept of bounded rationality seriously and builds upon earlier literature by adding a further logical layer of complexity to the models – namely, by considering the decision of agents regarding which forecasting rule to use. Branch (2004a) proposes a Rationally Heterogeneous Expectations (RHE) model. Agents are uncertain about the appropriate model to use and each period make a discrete choice between alternatives in a multinomial logit model. The choice set is assumed to contain a VAR forecast (which approximates ‘rational’ expectations), adaptive expectations and naïve expectations. The choice itself is based on the MSE performance of the various measures, which is contrasted with an (estimated) ‘cost’ of using a particular forecasting methodology. Although this approach is particularly satisfactory from a theoretical economic perspective (namely, conceptualising agents as making rational choices based on cost/benefit calculations) it is not without its problems. First, it is relatively data-expensive – Branch uses repeated cross-sectional data, although this can be adapted to less-rich data sources by modifying the likelihood function of the problem. Second, and more damagingly, a considerable number of ad hoc assumptions have to be made. For instance, it is not clear what specific VAR should be used to suitably approximate ‘rational’ forecasts. More critically, the results depend crucially on the specification of the ‘menu’ of predictors available – the composition of which is more than debatable. Branch (2004b) himself considers a completely different menu from that in his earlier paper (Branch, 2004a), now assuming that expectations are formed as a

discrete choice between a finite set of forecasting functions which differ by their recursive updating frequencies. When contrasting both models to the data with the Kullback-Leibler distance measure, the author concludes that elements of both seem relevant. Despite the undoubted novelty and interest of the approach, it is likely overestimating the rationality of agents forming their inflation expectations. Section 6 below will consider the properties of the distribution of inflation expectations, as well as new evidence regarding the way agents are approaching the problem of forming expectations. Those findings seem difficult to square with the purposeful agents picking items out of a menu of predictors envisaged by Branch – even if that is interpreted in terms of an ‘as if’ story.

Two particular recent views about inflation will receive greater attention below¹¹¹.

Section 6.2 will consider the approach proposed by Mankiw and Reis (2002), the basic intuition of which was already mentioned earlier. According to this view, agents are involved in a process of discretely updating their expectations to what is rational, when they get the chance of doing so (in a “Calvo” sense). A related story is provided by Carroll (2003), with the important difference that the updating envisaged here is one towards the expectations of professional forecasters. Carroll’s model is considered empirically in the next section. Its central idea is interestingly different from much of the literature – the general public may essentially be “outsourcing” the process of forming expectations to the experts.

¹¹¹ Of course, in a strict sense both the Carroll and Mankiw/Reis frameworks are encompassed by the Branch approach. However, as already commented, this is both the greatest virtue and the worst shortcoming of the method: it is so general that it can encompass *any* mechanism of expectations formation. The expectations of the general public, however, seem to be formed on a simpler and less well-defined basis – see Section 6.

Models of the formation of inflation expectations are undoubtedly useful. Achieving an understanding of this process could shed light on many interesting questions; in turn, this may help policymakers avoid outcomes such as inflationary spirals featuring the feedback of expectations into decisions. Indeed, in this as in other cases, the remedy is rather dependent on the nature of the illness – most notably, on its rationality content. Nevertheless, it is important not to overstate the importance of *models* relative to *data*. With the data sources available today, it is perfectly possible to carry out valuable empirical work involving expectations without necessarily having a highly definitive, congruent understanding of the process of expectations formation. Chapter 2 provided an example of the kind of empirical application expectations data can contribute to. Thus, it is even possible for expectations to behave in the way of “animal spirits” from the modeller’s perspective, yet be of consequence to economic outcomes. Actual data on expectations seem by far preferable to many of the ad hoc expectations aids that have been strapped to numerous empirical models in the past.

5 THE EXPECTATIONS OF CONSUMERS & ECONOMIC EXPERTS

5.1 Introduction

The present section proposes to develop our understanding of consumer inflation expectations by considering the way in which they might be influenced by the forecasts of professional forecasters and the diffusion of these through the mass media.

Carroll (2003) made a pioneering contribution in this context in his study of US consumer expectations from the Michigan SRC survey and expert forecasts from the Survey of Professional Forecasters. We propose here to reconsider his findings in a European context, with a larger sample of seven countries, namely: France, Spain, Italy, the Netherlands, the UK, Sweden and Germany¹¹². We will rely on data from the EC Consumer Survey – quantified according to the methodology selected in Chapter 1 – to obtain consumer inflation expectations, and use data from the “Consensus Forecasts” surveys to obtain the expectations of professional forecasters. This work is fully novel in the context of European data. Indeed, although Gerberding (2001, 2006) has considered both types of data previously, she did not examine the *interaction* between them and focused mainly on the analysis of rational expectations.

It is useful to briefly expose the main working hypotheses used in Carroll (2003) at this stage, although a fuller description will be provided in Section 6.2 when presenting Mankiw and Reis’ (2002) sticky information model. Indeed, the model proposed by Carroll (2003) can be interpreted as a microfoundation for that of Mankiw and Reis. In essence, Carroll assumes that people obtain their macroeconomic views from information

¹¹² Belgium is unfortunately excluded from this section due to the unavailability of the relevant expert data.

supplied in the media, which is absorbed probabilistically. In particular, the updating of expectations is assumed to follow the Calvo (1983) structure. Given this structure Carroll's derivation eventually reaches the following equation:

$$CONSEXP_t[\pi_{t,t+4}] = \lambda N_t[\pi_{t,t+4}] + (1 - \lambda)CONSEXP_{t-1}[\pi_{t-1,t+3}] \quad (76)$$

Where *CONSEXP* denotes consumer expectations, and *N* denotes the forecasts reported in newspapers. The latter will be proxied by the data on expert forecasts, thus taking the view that these are reported and transmitted to the general public through the media.

Section 5.2 introduces the Consensus Forecasts survey. The forecasting power of consumers and experts is subsequently compared in Section 5.3, before formally considering the relationship between both sets of expectations in Section 5.4. Section 5.5 directly considers the role of news coverage in the diffusion of expert forecasts and examines its implications in the context of the UK. Finally, Section 5.6 concludes.

5.2 Introducing the Consensus Forecasts survey

The London-based macroeconomics survey firm Consensus Economics has conducted a monthly survey of leading forecasters¹¹³ for a number of industrialised countries since the autumn of 1989. These renowned experts are asked to provide forecasts for the development of a range of important macroeconomic variables for their country, among which is inflation. On the basis of this information, a “consensus forecast” is calculated

¹¹³ As an example, the following UK forecasters participated in the January 2006 edition of Consensus Forecasts: Barclays Capital, Lombard Street Research, Williams de Broe, Lloyds TSB Financial Markets, Goldman Sachs, Citigroup, Credit Suisse First Boston, DTZ Research, Experian Business Strategies, RBS Financial Markets, UBS, Morgan Stanley, ABN Amro, Confederation of British Industry, Global Insight, HBOS, ITEM Club, JP Morgan, Oxford – LBS, Cambridge Econometrics, Deutsche Bank, HSBC, ING Financial Markets, Lehman Brothers, Liverpool Macro Research, Schroders, Economic Perspectives.

for each macroeconomic variable. This is a simple arithmetic average of all of the individual predictions collected for a given economic indicator.

Unfortunately for our purposes, the monthly data compiled by Consensus Economics consists of “fixed event” forecasts with changing forecast horizons – i.e. forecasting the inflation for the current and following year. However, Consensus Economics does collect on a quarterly basis additional forecasts from its panellists for key macroeconomic variables – including consumer price inflation – for each of the following one to six quarters. The means of these forecasts are published in the firm’s *Consensus Forecasts* reports.

The public availability of these data, and therefore the size of our sample, is limited by the reports stored in the British Library. Furthermore, the quarterly forecasts were not collected by Consensus Economics for all countries at the outset. Table 62 summarises the situation in terms of the time periods that are available for consideration here.

TABLE 62: CONSENSUS FORECASTS, SAMPLE USED IN THIS STUDY

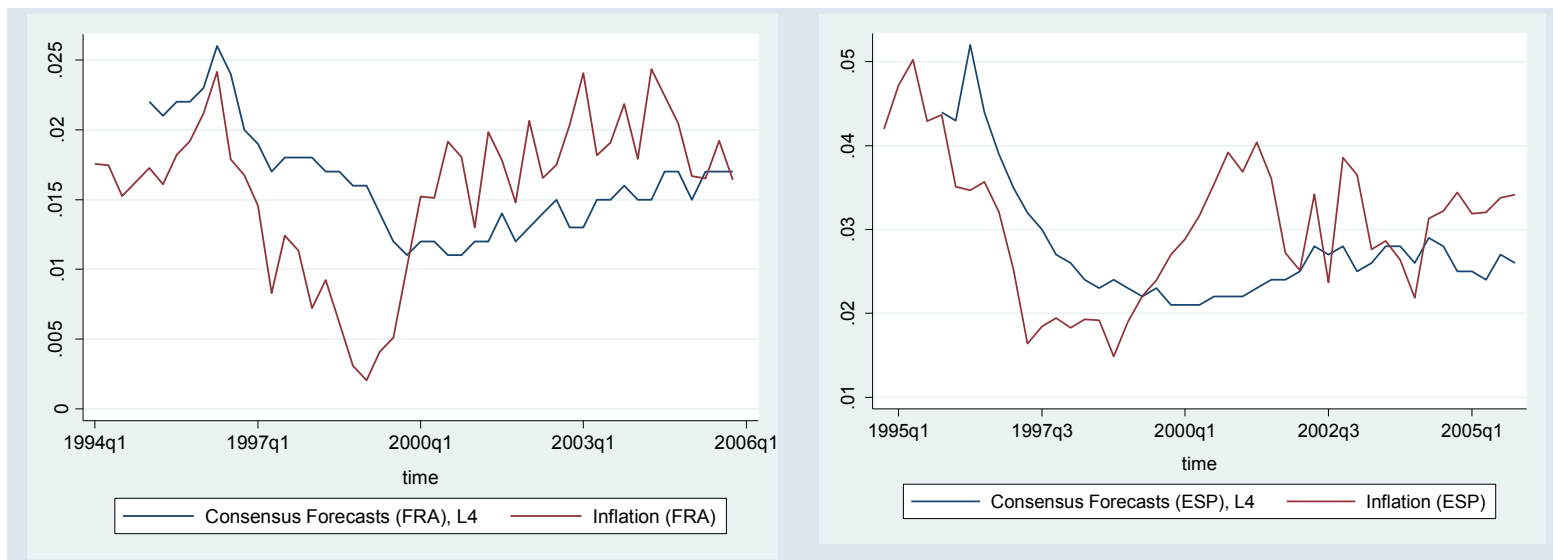
Country	Sample Start	Sample End
Germany	1994q1	2005q4
France		
United Kingdom		
Italy		
Netherlands	1994q4	2005q4
Spain		
Sweden		
Belgium	None	None

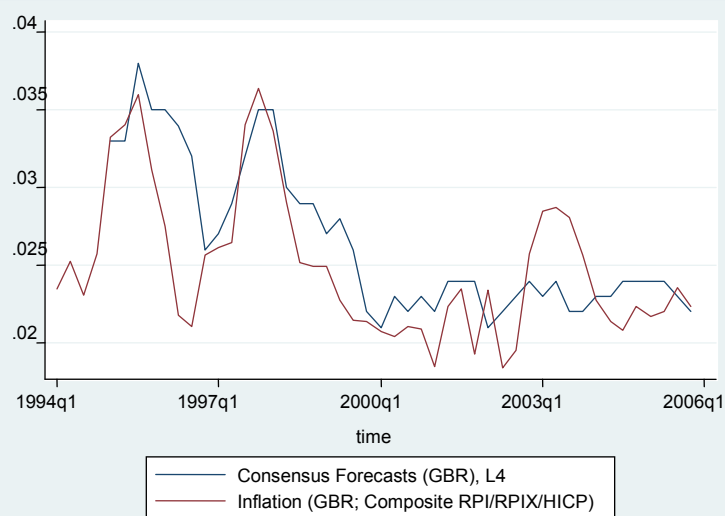
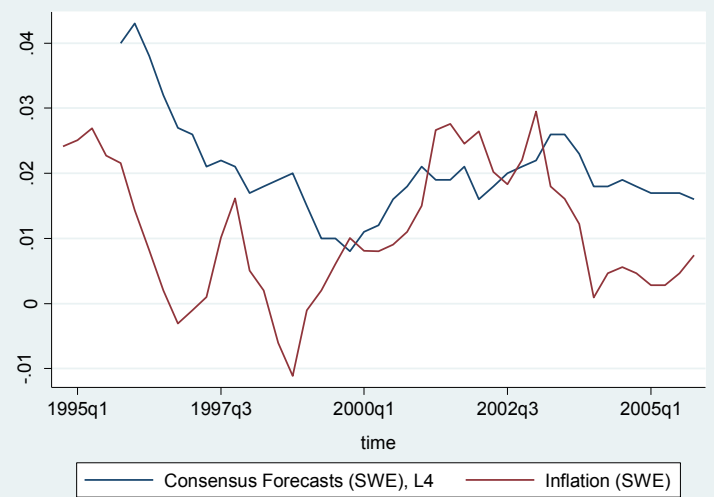
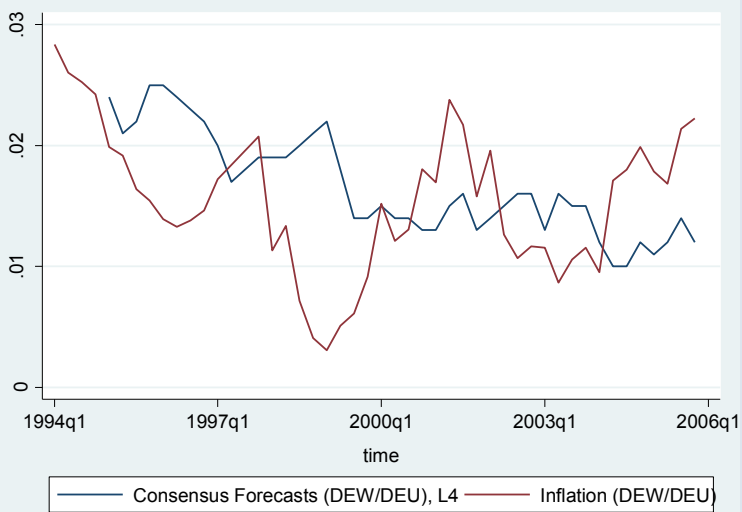
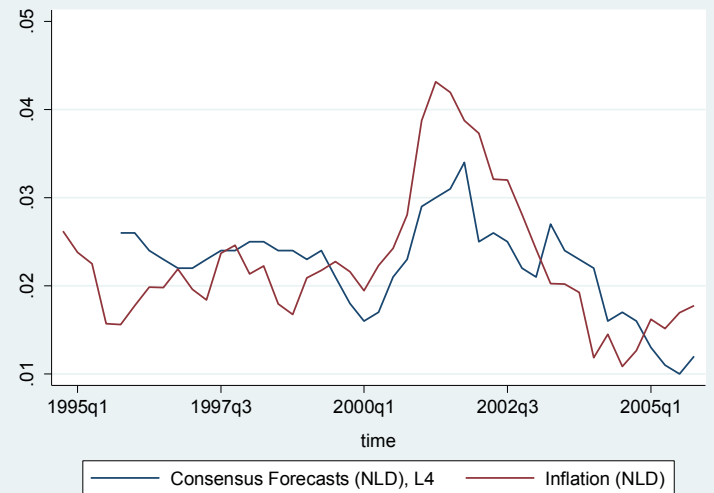
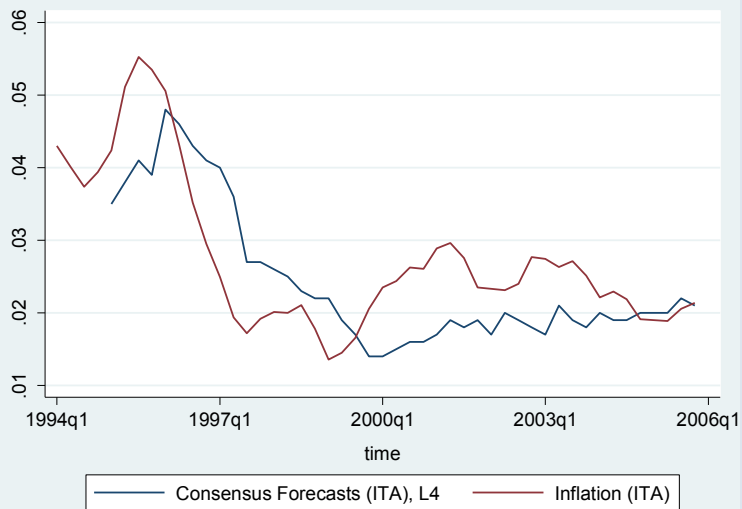
It is important to note that there are a number of “discontinuities” in the data:

- Forecasts in Germany switch from concerning West-German consumer prices only to German consumer prices in 1996q4.
- Forecasts in the UK switch from concerning the headline RPI rate to RPIX inflation in 1997q2. There is one further switch in 2005q1 towards HICP inflation.

Figure 102 plots the Consensus Forecasts *at the time they refer to* against actual inflation, so that the gaps between both series are expectations errors.

FIGURE 102: CONSENSUS FORECASTS (AT TIME REFERRED) AND INFLATION





**[BELGIUM IS MISSING DUE TO
ABSENCE OF QUARTERLY
CONSENSUS FORECASTS]**

It is interesting to note that professional forecasters on average consistently failed to predict the very marked fall in inflation that was characteristic of most countries at the end of the 1990s (with the only exception of the Netherlands in this sample). Similarly, there is a generalised underprediction of the increase in inflation in Eurozone countries in the post-EMU period (with the exception of Germany).

5.3 Forecasting power of consumers' and experts' expectations

We now compare the forecasting performance of consumers with that of professional forecasters on the basis of some descriptive statistics already deployed in Chapter 1: the Mean Error (ME), RMSE and the MSE decomposition of equation (36) (see Chapter 1, Section 4.2). Table 63 presents the results. Consumer forecasts are made on the basis of the optimal quantification measure recommended in Chapter 1 (EXPINFLPBIS) and are contrasted with a CPI benchmark. The expert forecasts based on the data from Consensus Economics are contrasted with the variable the professional forecasters were predicting (CPI in all cases except the UK, which uses several indices – RPI, RPIX and HICP – over the sample). A naïve benchmark is also provided, which simply predicts the same rate of inflation for the coming year as that observed over the previous year on the inflation measure used by Consensus Economics professional forecasters.

TABLE 63: PREDICTIVE ABILITY OF CONSUMERS AND EXPERTS

	CONSUMER FORECASTS					
Country	ME	RMSE	U1	U2	U3	NOB
France	0.0040	0.0067	35.0	8.9	56.1	44
Spain	0.0066	0.0123	28.8	37.8	33.3	41
Sweden	0.0028	0.0103	7.2	5.7	87.1	36
UK	0.0031	0.0099	9.7	37.4	52.9	44
Italy	0.0025	0.0085	8.3	13.3	78.5	44
Netherlands	0.0038	0.0085	19.2	6.9	73.9	41
Germany	-0.0014	0.0086	2.6	63.5	33.9	44

EXPERT FORECASTS						
Country	ME	RMSE	U1	U2	U3	NOB
France	-0.0005	0.0063	0.7	20.5	78.8	44
Spain	0.0014	0.0089	2.4	34.1	63.5	41
Sweden	-0.0106	0.0154	46.6	14.4	39.0	41
UK	-0.0015	0.0038	15.3	11.2	73.5	44
Italy	0.0021	0.0082	6.4	10.0	83.7	44
Netherlands	0.0006	0.0060	0.8	0.0	99.2	41
Germany	-0.0022	0.0075	8.5	46.8	44.7	44

NAIVE FORECASTS						
Country	ME	RMSE	U1	U2	U3	NOB
France	0.0001	0.0061	0.0	28.2	71.7	44
Spain	-0.0012	0.0095	1.6	45.6	52.8	41
Sweden	-0.0020	0.0126	2.4	42.0	55.6	41
UK	-0.0002	0.0068	0.1	49.8	50.2	44
Italy	-0.0018	0.0098	3.4	24.1	72.5	44
Netherlands	-0.0005	0.0085	0.4	27.6	72.0	41
Germany	-0.0006	0.0073	0.6	51.8	47.6	44

As can be seen from the mean errors, expert forecasts are less biased than consumer forecasts in all but two countries (Sweden and Germany). Similarly, they are more accurate in RMSE terms for all countries but Sweden. We note that the qualitative findings in the Swedish case are not driven by the slightly different sample used for experts (starting in 1994q4) and consumers (starting in 1996q1). The naïve forecasts beat both consumer and expert forecasts in terms of bias for every country; however, this does not translate straightforwardly into RMSE dominance. Table 64 summarises the situation by providing the average performance for each type of forecast.

TABLE 64: PREDICTIVE ABILITY CONSUMERS/EXPERTS, SUMMARY

Data	Avg. ME	Avg. RMSE	Avg. U1	Avg. U2	Avg. U3
Consumer forecasts	0.0031	0.0093	15.8	24.8	59.4
Expert forecasts	-0.0015	0.0080	11.5	19.6	68.9
Naïve forecasts	-0.0009	0.0086	1.2	38.4	60.3

As could be expected, expert forecasts dominate in RMSE terms and also present the most favourable MSE decomposition. This result is in line with intuition. Nevertheless, their performance is surprisingly close to that of consumers and naïve measures, remaining well within the same order of magnitude.

A question worthy of consideration is the extent to which these forecasts have further predictive power beyond that which is contained in past inflation data (Carroll, 2003). As observed above, a naïve forecast performs relatively well in RMSE terms, and we already highlighted earlier (e.g. Section 2.1.3) that consumer inflation expectations presented highly adaptive features (although they went beyond that at points, and were certainly not completely “silly”). The question is therefore: to what extent is the predictive performance of such measures based on persistence in the inflation process, and to what extent do they bring information beyond that?

We examined this by regressing the inflation rate over the coming year on the inflation rate over the past year and either the consumer or expert inflation expectations, before running a non-nested test between both sets of forecasts. In terms of specification diagnostics, we considered a number of tests for these regressions, including: the Portmanteau (Q) test for white noise; Bartlett's periodogram-based test for white noise; the Breusch-Godfrey test for higher-order serial correlation; the Breusch-Pagan/Cook-Weisberg test for “general” heteroscedasticity and an LM test for autoregressive conditional heteroscedasticity; and finally, skewness and kurtosis tests for the normality of residuals.

The tests revealed widespread autocorrelation and ARCH findings, and a few null rejections on the general heteroscedasticity test; the behaviour of the residuals seemed to be well characterised by a normal distribution, however. Accordingly, Newey-West corrections were implemented for the standard errors below, with an inspection of the correlograms revealing that a lag of four was a fully appropriate specification. Table 65 presents the results. “EC” denotes expectations from the EC Consumer Survey, whilst “CF” denotes the Consensus Forecasts.

TABLE 65: CONSUMERS & EXPERTS, INTRINSIC PREDICTIVE ABILITY

Equ.	France: $\pi t, t+4$					Spain: $\pi t, t+4$				
	Constant	$\pi t-5, t-1$	CFt	ECt	R ²	Constant	$\pi t-5, t-1$	CFt	ECt	R ²
1	0.0104 * [0.09]	0.3585 [0.17]	-0.0197 [0.96]	-	0.12	0.0216 *** [0.00]	0.1150 [0.61]	0.1437 [0.56]	-	0.24
2	0.0087 ** [0.05]	0.0321 [0.89]	-	0.5551 * [0.07]	0.23	0.0219 *** [0.00]	0.1330 [0.44]	-	0.1381 [0.50]	0.09
3	0.0065 [0.29]	-0.0517 [0.83]	0.1790 [0.58]	0.6097 ** [0.03]	0.24	0.0206 *** [0.01]	0.0440 [0.84]	0.1449 [0.57]	0.1385 [0.50]	0.29

Equ.	Germany: $\pi t, t+4$					Italy: $\pi t, t+4$				
	Constant	$\pi t-5, t-1$	CFt	ECt	R ²	Constant	$\pi t-5, t-1$	CFt	ECt	R ²
1	0.0185 *** [0.00]	0.1885 [0.33]	-0.4154 [0.21]	-	0.06	0.0107 * [0.06]	-0.4059 [0.27]	1.1283 *** [0.01]	-	0.51
2	0.0158 *** [0.00]	0.1070 [0.55]	-	-0.1898 [0.40]	0.04	0.0098 * [0.06]	0.0093 [0.97]	-	0.6827 *** [0.00]	0.46
3	0.0205 *** [0.00]	0.3257 ** [0.05]	-0.4545 * [0.09]	-0.2201 [0.27]	0.11	0.0083 [0.10]	-0.6023 * [0.01]	0.9377 *** [0.01]	0.5299 *** [0.01]	0.64

Equ.	United Kingdom: $\pi t, t+4$					The Netherlands: $\pi t, t+4$				
	Constant	$\pi t-5, t-1$	CFt	ECt	R ²	Constant	$\pi t-5, t-1$	CFt	ECt	R ²
1	0.0118 ** [0.02]	-0.3406 *** [0.00]	0.8178 *** [0.00]	-	0.64	0.0007 [0.91]	-0.0167 [0.94]	1.0109 *** [0.01]	-	0.43
2	0.0244 *** [0.00]	-0.3048 * [0.09]	-	0.3588 *** [0.00]	0.25	0.0120 *** [0.00]	-0.0321 [0.86]	-	0.6067 * [0.06]	0.45
3	0.0093 ** [0.03]	-0.2994 *** [0.00]	1.0718 *** [0.00]	-0.2292 * [0.06]	0.68	-0.0027 [0.66]	-0.2589 [0.27]	0.9810 *** [0.01]	0.5170 * [0.09]	0.49

Sweden: $\pi_t, t+4$					
Equ.	Constant	$\pi_{t-5, t-1}$	CFt	ECt	R ²
1	0.0025 [0.70]	0.1248 [0.63]	0.3416 [0.44]	-	0.06
2	0.0072 [0.14]	0.1784 [0.52]	-	0.1244 [0.86]	0.05
3	0.0019 [0.72]	0.0857 [0.78]	0.3433 [0.43]	0.1295 [0.86]	0.06

NOTE: *, **, *** respectively denote significance at the 10%, 5% and 1% levels. Newey-West correction applied to the standard errors, with four lags. P-values provided in brackets.

Carroll (2003) ran the same regression presented in Table 65 in the context of the United States, using data from the Survey of Professional Forecasters (SPF) and the Michigan survey of consumers. He found that both sets of expectations were highly significant when considered on their own, whereas only the expert forecasts proved to be so in a non-nested test. Carroll concluded that there was therefore no information contained in the consumer expectations that was not also included in the forecasts by the experts, and that the latter forecasts had significant predictive power not contained in the consumer expectations. The author further concluded that the professional forecasts were more “rational” than those of consumers.

The evidence presented here over seven countries offers a more contrasted story. The United Kingdom and the Netherlands broadly follow the “pattern” found by Carroll for the US, and Italy does to a lesser extent. The opposite is however true of France, although it is difficult to assess the meaning of this given the very significant problem of measurement error in the consumer data of this country evidenced in Chapter 1. Only insignificant results are available in the cases of Spain and Sweden, whilst the German equations display the “wrong sign”.

Obviously, the samples used here are relatively small (at best including 44 observations). Nevertheless, it is possible to find some (mild) support on European data for the findings of Carroll. It is difficult to assess whether the absence of stronger support is due to sample specificities (e.g. regime shifts in European data) or to underlying agent behaviour in certain countries. In any event, the finding of RMSE-dominance noted earlier (with the exception of Sweden) stands. In this light, we turn to examining the relationship between expert and consumer expectations, and in particular whether the former can be seen to influence the latter in a way that conforms to the model proposed by Carroll.

5.4 The relationship between the expectations of consumers and experts

The theory proposed by Carroll suggests that expectations should spread from experts to consumers. This observation implies that professional forecasts should Granger-cause consumer inflation expectations, but that the reverse finding should not hold. Carroll found that this was indeed the case in the US.

Table 66 reconsiders the issue in the context of our sample of seven countries. We regressed consumer and expert forecasts on four lags of both variables and a constant. When considering the same specification tests highlighted in the previous sub-section, we found here only very occasional traces of autocorrelation or heteroscedasticity – contrary to what was the case earlier. There were also a few isolated findings of kurtosis or skewness. Where appropriate, we considered Newey-West standard errors. We also considered an alternative specification using eight lags of both variables. The qualitative conclusion remained unaltered to these checks. We therefore present the results of the uncorrected regressions with four lags, but the findings are general.

TABLE 66: GRANGER CAUSALITY, CONSUMER & EXPERT FORECASTS

Dep. Var.	France				Spain			
	Constant	$\Sigma(\text{CFt-1} \dots \text{CFt-4})$	$\Sigma(\text{ECt-1} \dots \text{ECt-4})$	R^2	Constant	$\Sigma(\text{CFt-1} \dots \text{CFt-4})$	$\Sigma(\text{ECt-1} \dots \text{ECt-4})$	R^2
CFt	0.0015 [0.23]	0.8688 *** [0.00]	0.0368 [0.92]	0.83	0.0072 *** [0.00]	0.6479 *** [0.00]	0.0707 ** [0.01]	0.91
ECt	0.0020 [0.25]	-0.0145 ** [0.02]	0.9033 *** [0.00]	0.85	0.0052 [0.28]	-0.1363 [0.58]	0.9363 *** [0.00]	0.87

Dep. Var.	Germany				Italy			
	Constant	$\Sigma(\text{CFt-1} \dots \text{CFt-4})$	$\Sigma(\text{ECt-1} \dots \text{ECt-4})$	R^2	Constant	$\Sigma(\text{CFt-1} \dots \text{CFt-4})$	$\Sigma(\text{ECt-1} \dots \text{ECt-4})$	R^2
CFt	0.0019 [0.14]	0.8741 *** [0.00]	-0.0117 [0.64]	0.81	0.0015 [0.18]	0.8846 *** [0.00]	0.0325 * [0.07]	0.93
ECt	0.0021 [0.37]	0.0227 [0.41]	0.8325 *** [0.00]	0.78	0.0044 * [0.06]	0.1367 ** [0.02]	0.6816 *** [0.00]	0.82

Dep. Var.	United Kingdom				The Netherlands			
	Constant	$\Sigma(\text{CFt-1} \dots \text{CFt-4})$	$\Sigma(\text{ECt-1} \dots \text{ECt-4})$	R^2	Constant	$\Sigma(\text{CFt-1} \dots \text{CFt-4})$	$\Sigma(\text{ECt-1} \dots \text{ECt-4})$	R^2
CFt	0.0037 * [0.08]	0.7546 *** [0.00]	0.0986 [0.13]	0.83	0.0037 [0.16]	0.8024 *** [0.00]	0.0278 [0.33]	0.83
ECt	-0.0018 [0.62]	0.2908 [0.21]	0.7138 *** [0.00]	0.80	0.0022 [0.55]	0.3508 [0.12]	0.4853 ** [0.03]	0.55

Dep. Var.	Sweden			
	Constant	$\Sigma(\text{CFt-1} \dots \text{CFt-4})$	$\Sigma(\text{ECt-1} \dots \text{ECt-4})$	R^2
CFt	0.0072 *** [0.01]	0.5155 *** [0.00]	0.1705 * [0.09]	0.78
ECt	0.0012 [0.52]	-0.0185 [0.27]	0.9004 *** [0.00]	0.89

NOTES: *, **, *** respectively denote significance at the 10%, 5% and 1% levels. P-values are reported in brackets. For the “sums of coefficients”, these correspond to the p-value results of an F-test of joint significance.

As can be seen from glancing at the table, no strong case can be made for *either* variable Granger-causing the other. In particular, the results provide no clear evidence of expert forecasts feeding into consumer expectations, in contrast with Carroll’s findings. The apparently significant coefficient for Italian expert forecasts in the consumer expectations equation is achieved with traces of residual heteroscedasticity and autocorrelation and becomes insignificant when the Newey-West correction is applied.

Given this, a number of hypotheses are possible. First, the results could straightforwardly indicate that the feedback mechanism posited by Carroll does not exist – or else is very weak – in the European countries examined and over the sample period studied. Second, it may simply be that the sample considered does not allow us to detect such feedback. Indeed, Carroll benefited from a larger sample (1981q3-2000q2) that also proved more interesting in terms of the behaviour of inflation. The sample covered in the exercise above broadly embraces a period of relatively low and stable inflation in the countries considered. In such conditions, the incentives for consumers to pay attention and absorb expert opinion may be considerably reduced – as may be the *availability* of such expert opinion, an observation which we address in the following section. Finally, it is also possible that the Consensus Forecasts do not represent an appropriate proxy for the “expert” opinion absorbed by consumers – although it is hard to imagine another plausible proxy (e.g. OECD forecasts or government forecasts?) that would not be highly correlated with this – in such a way that qualitatively different findings would be obtained.

We nevertheless proceeded to testing Carroll’s model for the formation of consumer inflation expectations, which was summarised earlier in equation (76). We deployed the Consensus Forecasts as our proxy for the “newspaper forecasts” of that model. This is arguably a reasonable proxy, as the experts and institutions behind the Consensus Forecast often feature prominently in news stories about inflation and inflation forecasts, as is easily verified by searches in the Lexis Nexis newspaper database for the UK. The model presented in (76) was adapted to the present context and generalised, yielding the following equation:

$$EC_t[\pi_{t,t+4}] = \alpha_0 + \alpha_1 CF_t[\pi_{t,t+4}] + \alpha_2 EC_{t-1}[\pi_{t-1,t+3}] + \alpha_3 \pi_{t-5,t-1} + \varepsilon_t \quad (77)$$

Where EC_t denotes the expectations arising from the European Commission's Consumer Survey at time t , CF_t denotes the Consensus Forecast at time t and π denotes the inflation rate. As can be seen the equation generalises the model in (76) by allowing for a constant (α_0), by not imposing ($\alpha_1 + \alpha_2 = 1$) and by introducing α_3 – which gives consumers the possibility of updating to the most recently published past inflation rate. Carroll (2003) suggested that this option might be useful, given that a substantial part of news coverage about inflation is associated with the reporting of past inflation statistics.

Three specifications of equation (77) were successively estimated for insight;

1. The “baseline model” which imposes $\alpha_0 = 0$ and $\alpha_3 = 0$.
2. The “baseline model” augmented by a constant, whose significance may notably be indicative of specification problems.
3. The full model, as expressed in equation (77).

These specifications were estimated for the seven countries present in our sample and the maximum time series coverage available in each case (see Table 62). Table 67 presents the results. The specifications were initially estimated using ordinary least squares, and a battery of specification diagnostics was checked, namely:

- The Breusch-Godfrey test for higher-order serial correlation (with four lags, the test being performed separately for each order);
- White's (1980) test for heteroscedasticity, which regresses the squared residuals on all distinct regressors, cross-products, and squares of regressors;
- An LM test for autoregressive conditional heteroscedasticity;
- A joint Chi2(2) test for skewness and kurtosis to detect deviations from normality in the residuals.

In the cases where significant autocorrelation and/or heteroscedasticity were detected, Newey-West corrected standard errors were employed, allowing for four lags (which was shown to be appropriate in all cases from inspection of the relevant correlograms). Finally, an F-test of the null hypothesis that $\alpha_1 + \alpha_2 = 1$ (specifications 1 and 2) or that $\alpha_1 + \alpha_2 + \alpha_3 = 1$ (specification 3) was also implemented and its results are reported in the rightmost column of Table 67.

TABLE 67: REGRESSION RESULTS FOR CARROLL'S (2003) MODEL

France										
Equ.	α_0	α_1	α_2	α_3	R^2	Autocorr.	Het. Test	ARCH	Normal	Test
1	-	0.103 ** (0.044)	0.883 *** (0.049)	-	0.97	12.1 ** [0.02]	6.5 [0.26]	0.02 [0.89]	5.84 * [0.05]	0.35 [0.56]
2	0.0004 (0.001)	0.082 (0.067)	0.876 *** (0.057)	-	0.76	11.8 ** [0.02]	6.3 [0.28]	0.05 [0.82]	6.12 ** [0.05]	0.29 [0.59]
3	0.0010 (0.002)	-0.033 (0.120)	0.729 *** (0.108)	0.197 * (0.105)	0.78	6.71 [0.15]	7.3 [0.61]	0.43 [0.51]	4.64 * [0.10]	0.71 [0.22]

Spain										
Equ.	α_0	α_1	α_2	α_3	R^2	Autocorr.	Het. Test	ARCH	Normal	Test
1	-	0.079 (0.068)	0.904 *** (0.080)	-	0.97	3.6 [0.47]	12.2 ** [0.03]	2.02 [0.16]	1.93 [0.38]	0.27 [0.61]
2	0.0026 (0.003)	-0.005 (0.122)	0.895 *** (0.086)	-	0.79	3.3 [0.51]	12.8 ** [0.03]	0.51 [0.48]	5.06 * [0.08]	1.24 [0.27]
3	0.0026 (0.003)	0.002 (0.139)	0.897 *** (0.110)	-0.007 (0.131)	0.79	3.3 [0.51]	15.5 * [0.08]	0.51 [0.48]	4.99 * [0.08]	1.11 [0.30]

Germany										
Equ.	α_0	α_1	α_2	α_3	R^2	Autocorr.	Het. Test	ARCH	Normal	Test
1	-	0.142 * (0.078)	0.811 *** (0.104)	-	0.96	12.9 ** [0.01]	12.3 ** [0.03]	0.54 [0.46]	8.23 ** [0.02]	1.59 [0.21]
2	0.0019 (0.002)	0.053 (0.138)	0.788 *** (0.106)	-	0.70	12.0 ** [0.02]	13.3 ** [0.02]	0.91 [0.34]	7.6 ** [0.02]	1.71 [0.20]
3	0.0019 (0.002)	0.054 (0.178)	0.788 *** (0.099)	-0.002 (0.101)	0.70	12.2 ** [0.02]	13.9 [0.13]	0.91 [0.34]	7.61 ** [0.02]	1.67 [0.20]

Italy										
Equ.	α_0	α_1	α_2	α_3	R^2	Autocorr.	Het. Test	ARCH	Normal	Test
1	-	0.220 (0.145)	0.745 *** (0.155)	-	0.96	11.6 ** [0.02]	15.5 *** [0.01]	2.05 [0.15]	4.13 [0.13]	0.88 [0.35]
2	0.0019 (0.002)	0.175 (0.183)	0.718 *** (0.143)	-	0.74	13.4 *** [0.01]	13.9 ** [0.02]	1.63 [0.20]	4.23 [0.12]	1.28 [0.26]
3	0.0026 (0.002)	0.419 *** (0.131)	0.916 *** (0.163)	-0.400 (0.161)	0.77	5.8 [0.22]	24.4 *** [0.00]	5.82 ** [0.02]	2.42 [0.30]	0.94 [0.33]

United Kingdom										
Equ.	α_0	α_1	α_2	α_3	R^2	Autocorr.	Het. Test	ARCH	Normal	Test
1	-	0.266 *** (0.085)	0.681 *** (0.095)	-	0.98	3.3 [0.51]	4.7 [0.46]	0.03 [0.87]	0.18 [0.91]	6.78 ** [0.01]
2	-0.0033 (0.002)	0.416 *** (0.14)	0.652 *** (0.096)	-	0.83	3.8 [0.43]	3.3 [0.65]	0.00 [0.99]	0.17 [0.92]	0.55 [0.46]
3	-0.0046 (0.003)	0.413 *** (0.14)	0.632 *** (0.101)	0.070 (0.105)	0.83	4.0 [0.41]	11.8 [0.22]	0.09 [0.76]	0.07 [0.97]	0.98 [0.33]

The Netherlands										
Equ.	α_0	α_1	α_2	α_3	R^2	Autocorr.	Het. Test	ARCH	Normal	Test
1	-	0.259 *** (0.087)	0.695 *** (0.099)	-	0.96	3.5 [0.47]	3.7 [0.60]	2.2 [0.14]	8.28 ** [0.02]	2.17 [0.15]
2	0.0029 (0.003)	0.189 * (0.111)	0.627 *** (0.119)	-	0.50	2.79 [0.59]	4.4 [0.50]	2.02 [0.16]	10.91 *** [0.00]	1.76 [0.19]
3	0.0041 * (0.002)	0.092 * (0.049)	0.343 (0.273)	0.272 (0.170)	0.55	2.52 [0.64]	22.9 *** [0.01]	3.74 * [0.05]	2.04 [0.36]	4.35 ** [0.04]

Sweden										
Equ.	α_0	α_1	α_2	α_3	R^2	Autocorr.	Het. Test	ARCH	Normal	Test
1	-	0.076 ** (0.033)	0.834 *** (0.073)	-	0.96	6.9 [0.14]	4.2 [0.52]	0.00 [1.00]	6.33 ** [0.04]	3.82 * [0.06]
2	-0.0015 (0.001)	0.160 ** (0.075)	0.827 *** (0.072)	-	0.86	7.8 [0.10]	4.4 [0.50]	0.19 [0.67]	4.50 [0.11]	0.03 [0.87]
3	-0.0002 (0.001)	0.078 (0.063)	0.718 *** (0.062)	0.088 ** (0.043)	0.87	12.0 ** [0.02]	9.0 [0.44]	0.03 [0.87]	1.32 [0.52]	3.78 * [0.06]

NOTE: *, **, *** respectively denote significance at the 10%, 5% and 1% levels. P-values provided in brackets [] and standard errors provided in parentheses ().

Before commenting on these results, it is interesting to note the broad findings reached by Carroll (2003) for the United States, on the basis of Michigan Consumer Survey data and the Survey of Professional forecasters (used as a proxy for “newspaper forecasts”). This author found that the constant appeared to be significant (a finding interpreted in terms of misspecification) whereas lagged inflation was insignificant. His regressions appeared not to suffer from autocorrelation according to Box-Ljung tests, although they presumably suffered from heteroscedasticity given that Newey-West standard errors were used. Other specification diagnostics (e.g. those concerning the normality of residuals) were not considered.

A number of notable findings emerge from Table 67:

First, the constant in these equations never reaches significance at the 5% level, and only achieves it at the 10% in one case, for the Netherlands. Nevertheless, there is widespread evidence of misspecification for many countries, mostly involving autocorrelated or heteroscedastic residuals and violations of the normality assumption. Although the findings of autocorrelation may be to some extent due to an overlapping data issue, they may also be a symptom of the model being “too simple” in most cases – with the behaviour of the residuals possibly resulting from the unaccounted and persistent effects of omitted variables. This behaviour may also be reflective of issues relating to the linear functional form deployed in this case. Indeed, the conditional density of consumer inflation expectations is not a straightforward concept to analyse empirically, and nonlinearities cannot be excluded a priori.

Second, the key variable in the analysis is consistently found to be the lagged EC Consumer Survey expectation, highlighting the strong persistence of this process. Indeed, its associated coefficients are large and the variable is found to be significant at the 1% level in every single specification save the third Dutch regression (whose results are most probably due to an issue of multicollinearity).

Third, although it is possible to detect some significant influence from the expert forecasts on consumer expectations (consistent with the thrust of Carroll's approach), their impact is nonetheless considerably weaker than the autoregressive component of the equation.

Fourth, lagged inflation appears to be insignificant for the most part (the cases of France and Sweden excepted). It is important to stress that this finding does not in the least indicate the unimportance of updating to past inflation for consumers. Indeed, as has already been highlighted in several places in this study (e.g. Section 2.1.3 in this chapter) the inflation expectations of consumers are very closely related to actual inflation. The insignificance of the past inflation coefficient rather translates the fact that its information has already been subsumed in the autoregressive component of the equation, which also contains additional information.

Finally, it is worthwhile to examine the case of the United Kingdom in detail. There is no doubt that the model performs impressively well in this context considering its extreme simplicity. The specification diagnostics provide no grounds on which to question the congruence of the model. Furthermore, both the constant and lagged inflation appear to

be superfluous to the model in terms of significance. However, in the context of specification 1, the coefficient restriction implied by Carroll's model can be rejected (unlike in specification 2). Nevertheless, we impose the restriction for interest and therefore constrain the coefficients as follows in both specifications:

$$\alpha_1 = \lambda \text{ and } \alpha_2 = (1 - \lambda)$$

The estimation is carried out through nonlinear least squares. In the context of specification 1, λ is found to equal 0.18 whereas upon re-estimation of specification 2 it takes the value of 0.34 (and the constant is now significant). Interpreting these results in the framework of Carroll's model, this would suggest that in each quarter between one fifth and one third of the population update their forecast for the coming year. These results are close to (and indeed, insignificantly different from) those of Mankiw and Reis (2002) – who set $\lambda=0.25$ in their simulations – and Carroll (2003), who finds $\lambda=0.27$ when imposing the above restriction.

The qualitative similarity between UK and US results is very marked, in such a way that it may be worthy of attention. Indeed, in Table 65, we noted that the “intrinsic predictive ability” (i.e. controlling for inflation) of consumers and experts in the UK followed the US pattern. In Table 66, although it was not possible to establish unidirectional Granger causality from expert forecasts to consumer expectations, the coefficient found on the expert forecasts was relatively large in the sample of countries considered, and its insignificance may have been related to the small and “quiet” period of inflation history covered by our data. The similarity was further emphasised by the results of Table 67 above. Finally, the *a contrario* results obtained for all the other countries relative to the

US benchmark are suggestive that there might be a common set of reasons behind the US and UK findings. We will provide a possible interpretation for this in Section 5.6. However, it is useful to examine some further evidence before that.

Indeed, given the particular position of the UK, it is especially interesting to consider data on news coverage of inflation in the country and its relationship with expectations, and check whether the similarities with the US hold in this respect as well. This is the focus of the next section.

5.5 Inflation news coverage and inflation expectations in the UK

The diffusion of economic information through the news is an important mechanism posited by Carroll (2003) in his model of inflation expectations. In this section, we aim to replicate the empirical exercises he carries out for the US in the context of the United Kingdom and thereby check the robustness of his conclusions.

Carroll argues that if the view that the public derives its expectations of inflation from news stories is correct, it follows that people should be “better informed” in periods of more intensive news coverage. By “better informed”, the author means that the public should have expectations that are closer to those of professional forecasters in such periods. An alternative interpretation (not examined by Carroll) is that the public should be making smaller forecast errors on average in such periods. This possibility will also be considered below.

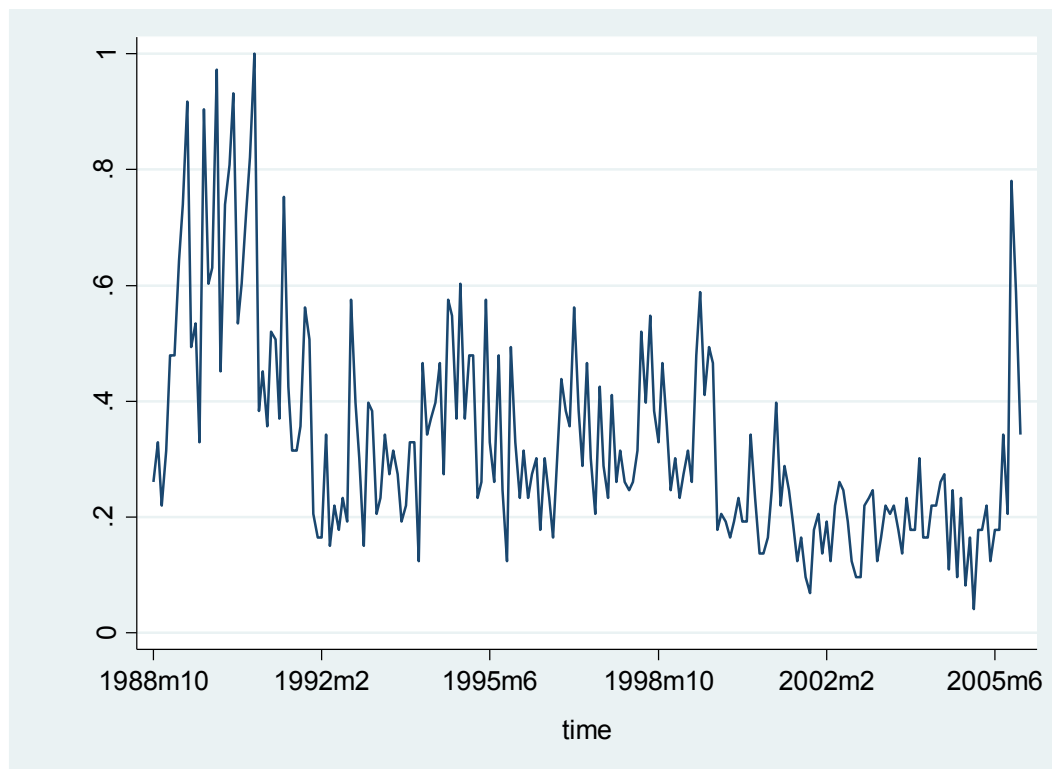
To check these hypotheses empirically, it is necessary to build an index of the intensity of news coverage. For this purpose, we collected monthly data on Lexis Nexis Professional regarding the number stories concerning UK inflation published in three major national newspapers: The Times, The Guardian and The Independent. These are major publications, with a joint average net circulation in March 2006 of about 1,240,000 copies according to the Audit Bureau of Circulation (ABC). They were selected on the basis of both their size and the backward sample coverage available on Lexis Nexis, which in the case of these publications allows us to build a monthly index for the period from October 1988 until December 2005.

A news story was counted as relevant for our purposes if it matched the following criteria:

- The story is classified in the system as concerning “inflation” and the “United Kingdom”.
- Inflation was a major component of the story, in that the word appeared a minimum of three times.

The index of intensity of news coverage was then created by expressing the number of news stories in any given month as a percentage of the maximum number achieved in any month in the sample. Hence, a value of 25% indicated that in that particular month, there were only a quarter as many stories about inflation as there were during the sample peak.

FIGURE 103: INDEX OF NEWS COVERAGE INTENSITY, 1988M10-2005M12



The mean value of the index over the sample is 33% – thus, roughly a third of the maximum coverage is attained on average. Interestingly, the peak intensity of news coverage on inflation is reached in October 1990. This is both the month with the highest inflation in our sample, and the time at which the UK joined the ERM. As the ERM disinflation proceeded, the number of news stories fell sharply. With the collapse of British membership of the Exchange Rate Mechanism on Black Wednesday (16 September 1992), there is a localised jump in inflation stories to about twice the sample mean, but this effect does not outlast the month. At the end of the 1990s (concretely, in December 1999) the coverage index collapses further (from 46% in November to 17% in December) and remains relatively stable at that low level. By then, UK inflation was itself at a low and moderately steady level, and most of the institutional innovations that

may have given journalists a reason to write about inflation in the general press (the Bank of England Act, the reduction of the period between MPC decisions and publication of minutes to two weeks etc.) had already taken place. A spectacular deviation from this pattern took place in October 2005, when the index jumps to nearly 80% of its maximum value. The month contained an unusual amount of news providing opportunities to write about inflation. Scanning the headlines reveals a variety of reasons behind the proliferation of articles, among which:

- The National Institute of Economic and Social Research raised its forecasts for inflation but cut its forecast for economic growth for the year and over the medium term¹¹⁴.
- There was intense speculation as to what the behaviour of interest rates would be. Some articles reported that the Bank of England governor, Mervyn King, “*moved to quell any lingering hopes of an interest rate cut arriving next month by emphasising that the Bank is targeting inflation, not overall growth or retail sales*”¹¹⁵. Others headlines however noted that “*King offers hope of interest-rate cut by the spring*”¹¹⁶ after his testimony to the Lords Economic Affairs Committee. The uncertainty in the press surrounding this period is well-captured by the following headline: “*MPC sharply split as it votes 6-3 against cut in rate*”¹¹⁷.

¹¹⁴ The Independent, October 28, 2005; The Times, October 28, 2005.

¹¹⁵ The Guardian, October 26, 2005.

¹¹⁶ The Times, October 26, 2005.

¹¹⁷ The Times, October 6, 2005.

- Several articles expressed worries about the behaviour of inflation in the context of soaring oil prices; one article noted, for instance: “*With oil prices at near-record levels, the MPC has been worried as much about rising inflation, which stands at the highest for eight years, as faltering growth.*”¹¹⁸
- Other miscellaneous items also generated articles about inflation. For instance, it was noted that Sir John Gieve was appointed as a deputy Governor at the Bank of England, succeeding Sir Andrew Large¹¹⁹.

Following this extraordinary month of news, the index fell to values of 59% and 34% successively.

It is important to highlight upfront a subtlety of interpretation that may be important in the econometrics below. Indeed, it is easily conceivable that the number of stories about inflation in the media is endogenously related to the level of the inflation rate, with more stories accompanying periods of higher inflation. This point suggests caution in making causal interpretations regarding bivariate relationships between news coverage and the accuracy of consumer expectations (or their similitude to those of the experts)¹²⁰. Indeed, were such a relationship to be found, it could arguably be reflecting the mechanism of Akerlof et al. (2000), according to which more consumers choose to become informed about inflation when the variable rises. Although the issue is raised by

¹¹⁸ The Times, October 22, 2005.

¹¹⁹ The Independent, October 18, 2005.

¹²⁰ It may be the case that for well-publicised events which generate media attention (e.g. a sales tax hike) consumers and experts independently agree on the inflationary impact, without there being any real “transmission”. We try to control in what follows for endogeneity arising from the level of inflation, but aim to address more carefully other types of endogeneity in future research, possibly through refinements to the news index.

Carroll (2003), the author makes no attempt to distinguish this interpretation from an independent role of the media in conveying information about inflation. In what follows, we choose to run additional regressions to increase the robustness of our inferences to this point.

We first examine the hypothesis that increases in news coverage translate into more accurate expectations on the part of consumers. This analysis is carried out using both Gallup and EC Consumer Survey monthly data. The dependent variable is defined in both cases as the absolute value of the forecast error:

$$\begin{aligned} \text{ABSERRBIS}_t &= \text{Abs}(\text{EXPINFLPBIS}_{t,t+4} - \text{INFLA}_{t,t+4}) \\ \text{ABSERRGLP}_t &= \text{Abs}(\text{EXPGLP}_{t,t+4} - \text{INFLA}_{t,t+4}) \end{aligned}$$

With this aim, we regress these variables on a constant and our news coverage index (specification 1) before adding the current inflation rate to the regression (specification 2). The equation we are considering is therefore:

$$\text{Forecast Error}_t = \alpha_0 + \alpha_1 \text{NewsIndex}_t + \alpha_2 \text{INFLA}_{t-4,t} + \varepsilon_t \quad (78)$$

According to the null hypothesis, coefficient α_1 should be significantly negative.

The results are presented in Table 68. We consider the same specification tests described in the previous sub-section. As can be seen the equations perform poorly in terms of specification diagnostics, a finding which is not unexpected given the simplicity of the regression considered. Accordingly, the standard errors are corrected for autocorrelation

and heteroscedasticity following the Newey-West approach allowing for twelve lags – which is found to be appropriate from a consideration of the correlograms.

TABLE 68: DOES NEWS COVERAGE IMPROVE FORECAST ACCURACY?

	Coefficients			Specification					
ABSERRGLP	α_0	α_1	α_2	R^2	NOBS	Autocorr.	Het. Test	ARCH	Normal
1	0.021 *** (0.005)	0.003 (0.011)	-	0.00	100	80.6 *** [0.00]	0.85 [0.66]	50.6 *** [0.00]	9.36 *** [0.01]
2	0.018 *** (0.005)	-0.014 (0.011)	0.212 * (0.117)	0.14	100	76.7 *** [0.00]	21.0 *** [0.00]	38.8 *** [0.00]	3.18 [0.20]
ABSERRBIS	α_0	α_1	α_2	R^2	NOBS	Autocorr.	Het. Test	ARCH	Normal
1	0.007 *** (0.002)	0.015 ** (0.007)	-	0.12	191	66.2 *** [0.00]	25.9 *** [0.00]	11.1 *** [0.00]	13.4 *** [0.00]
2	0.005 *** (0.002)	0.002 (0.005)	0.170 *** (0.052)	0.26	191	52.0 *** [0.00]	61.0 *** [0.00]	57.0 *** [0.00]	1.79 [0.41]

NOTE: *, **, *** respectively denote significance at the 10%, 5% and 1% levels. Newey-West correction applied to the standard errors, with twelve lags. P-values provided in brackets [] and standard errors provided in parentheses (). The autocorrelation test corresponds to the Breusch-Godfrey test for higher-order serial correlation with twelve lags. GLP sample: 1989m10-1998m1; EC sample: 1989m10-2005m8.

The findings are not particularly encouraging for the importance of the news coverage intensity index in improving the accuracy of consumer forecasts. The index appears to be insignificant in the light of the Gallup data, and only appears as significant in the EC data (with the wrong sign) artificially, as a product of the failure to control for the role of actual inflation. We also ran the same regressions with a three-month moving average of the news index to account for “memory” effects; the qualitative findings above remained unchanged, however.

We now turn to the hypothesis according to which greater news coverage leads to a closer alignment of consumer and expert forecasts due to greater and/or better updating of the former on the latter’s forecasts. This is much more directly related to Carroll’s model. The dependent variable we consider is simply the absolute difference between the

Consensus Forecasts and the expectations of consumers¹²¹, and the regression here is analogous to that in (78):

$$\begin{aligned} DIFCONSBIS_t &= Abs(ConsensusForecast_{t,t+4} - ECForecast_{t,t+4}) \\ DIFCONSBIS_t &= \alpha_0 + \alpha_1 NewsIndex_t + \alpha_2 INFLA_{t-4,t} + \varepsilon_t \end{aligned} \quad (79)$$

The results are presented in Table 69:

TABLE 69: NEWS EFFECT ON THE CONSUMER/EXPERT FORECAST GAP

DIFCONSBIS	Coefficients			Specification					
	β_0	β_1	β_2	R^2	NOBS	Autocorr.	Het. Test	ARCH	Normal
1	0.008 *** (0.001)	-0.009 *** (0.003)	-	0.14	46	13.5 [0.34]	1.25 [0.53]	3.43 * [0.064]	1.23 [0.54]
2	0.011 *** (0.002)	-0.008 ** (0.003)	-0.123 ** (0.060)	0.22	46	16.1 [0.19]	4.26 [0.51]	1.73 [0.19]	1.52 [0.47]

NOTE: *,**,*** respectively denote significance at the 10%, 5% and 1% levels. P-values provided in brackets [] and standard errors provided in parentheses (). Sample: 1994q1-2005q2.

Carroll (2003) found in the context of the US that his news index was significant at the 5% level over the sample 1981q3-2000q2. However, his Durbin-Watson statistic (the only diagnostic provided) revealed significant autocorrelation at the 1% level, and accordingly the author resorted to a Newey-West correction. The findings presented here are stronger and more encouraging. Both specifications above are surprisingly free of misspecification (at the conventional 5% level), excepting a hint of ARCH behaviour in the first regression. The news coverage index is highly significant and correctly signed in both equations and displays explanatory power independent to that of inflation itself. Hence, the media in the United Kingdom would indeed appear to be an important vehicle for the updating behaviour of consumers towards professional forecasts, to the extent that this happens.

¹²¹ The Gallup data is not used here, as the sample overlap with the Consensus Forecasts is too reduced.

Given this encouraging finding, it seems particularly interesting to consider the implication of news coverage for Carroll's variant of the sticky information model. Indeed, Carroll (2003) reports finding a higher speed of updating in the sub-sample characterised by above-average news coverage. We reconsider his findings in the case of the UK, for which a sample covering 1994q1-2005q2 is available. We examine two possibilities to characterise periods with "intense" news coverage:

1. Quarters with a number of news stories in excess of the full news sample average (1988q3-2005q4) [denoted as "mean1" below]. According to this definition, months in "high activity" quarters have over 24 news stories about inflation on average in the publications considered.
2. Quarters with a number of news stories in excess of the average in the period of the regression (1994q1-2005q2). This average is denoted as "mean2" below. According to this definition, months in these "high activity" quarters have over 20 news stories about inflation on average in the publications considered..

The rationale for considering two definitions is to check the robustness of any findings that may arise. Thus, definition one might be characterised as concerning months of "very high activity" whereas definition two would then apply to months of merely "high activity".

The model examined was already introduced earlier as equation (77). We consider the equation in its "baseline model" version, allowing for a constant but imposing the

constraint $\alpha_1 = \lambda$ and $\alpha_2 = (1 - \lambda)$, as it was found to be reasonable in the UK context in the previous section (5.4). The implied specification is reproduced here for convenience:

$$EC_t[\pi_{t,t+4}] = \alpha_0 + \lambda CF_t[\pi_{t,t+4}] + (1 - \lambda) EC_{t-1}[\pi_{t-1,t+3}] + \varepsilon_t$$

Rearranging yields the following equation:

$$(EC_t[\pi_{t,t+4}] - EC_{t-1}[\pi_{t-1,t+3}]) = \alpha_0 + \lambda (CF_t[\pi_{t,t+4}] - EC_{t-1}[\pi_{t-1,t+3}]) + \varepsilon_t \quad (80)$$

Table 70 presents the results of estimating (80) over the full sample and a number of sub-samples.

TABLE 70: NEWS COVERAGE AND SPEED OF UPDATING, MODEL 1

Equ.	Sample	α_0	λ	Specification					
				R^2	NOBS	Autocorr.	Het. Test	ARCH	Normal
1	All observations	-0.002 *** (0.001)	0.35 *** (0.092)	0.24	46	7.21 [0.51]	2.10 [0.35]	0.00 [0.97]	0.07 [0.97]
2	NEWS > mean 1	-0.001 (0.001)	0.40 ** (0.164)	0.36	13	11.40 [0.18]	6.45 ** [0.04]	0.04 [0.85]	1.61 [0.45]
3	NEWS < mean 1	-0.002 ** (0.001)	0.34 *** (0.113)	0.22	33	3.74 [0.88]	1.26 [0.53]	0.02 [0.89]	0.32 [0.85]
4	NEWS > mean 2	-0.001 * (0.001)	0.36 ** (0.132)	0.30	20	8.52 [0.38]	7.97 ** [0.02]	0.55 [0.46]	0.09 [0.96]
5	NEWS < mean 2	-0.002 ** (0.001)	0.39 ** (0.141)	0.25	26	4.36 [0.82]	1.78 [0.41]	0.47 [0.49]	0.31 [0.86]

NOTE: *, **, *** respectively denote significance at the 10%, 5% and 1% levels. P-values provided in brackets [] and standard errors provided in parentheses ().

As can be seen, the equations overall perform well in terms of specification diagnostics – excepting some traces of heteroscedasticity found when considering sub-samples, although it should be noted that these contain a small number of observations. Focusing on specifications 1-3, it can be seen that the speed of adjustment for the “above average” news periods ($\lambda = 0.40$) exceeds those found for the full sample ($\lambda = 0.35$) and for the “below average” news periods ($\lambda = 0.34$). These qualitative findings correspond to those

in Carroll (2003), although in our case the difference between high and low activity periods is not significant. Unfortunately, the relatively low number of observations available makes the drawing of definitive conclusions difficult. Specifications 4 and 5 provide a further cautionary note on the temptation to claim that Carroll's findings are reinforced. Indeed, using this alternative definition, a lower speed of updating coefficient is found for "high activity" news periods than for "low activity" periods. The difference is once again insignificant, however, highlighting that more definitive conclusions are hidden behind the standard errors.

We attempted to explore the issue further with the following nonlinear model, in which the weights on the expert and consumer forecasts are allowed to vary with the news index:

$$EC_t[\pi_{t,t+4}] = C_0 + (\alpha_0 + \alpha NewsIndex)CF_t[\pi_{t,t+4}] + (1 - \alpha_0 - \alpha NewsIndex)EC_{t-1}[\pi_{t-1,t+3}] + \varepsilon_t \quad (81)$$

TABLE 71: NEWS COVERAGE AND SPEED OF UPDATING, MODEL 2

C_0	α_0	α	R^2	NOBS
-0.0016 *** (0.0005)	0.072 (0.229)	0.821 (0.628)	0.84	46

NOTE: *, **, *** respectively denote significance at the 10%, 5% and 1% levels. Standard errors provided in parentheses ().

The coefficients found have the correct signs and magnitudes. The important size of the coefficient on the news index variable tends to suggest that expert forecasts gain their importance in consumers' assessments mostly when they are appropriately reported by then news media. However, once again, we must come to the conclusion that too much is hidden behind the standard errors to make robust judgements.

5.6 Conclusion

This section has provided some support for the intuitive view that expert forecasts outperform consumer ones. This was broadly confirmed by the RMSE evidence in Section 5.3. However, non-nested tests of both sets of forecasts generated mixed results, with the hypothesis of expert forecasts strictly dominating consumer ones receiving support in some countries but not in others. Similarly, the results regarding Granger causality and Carroll's sticky information model (Section 5.4) were unable to provide for most countries a clear-cut picture of information being transmitted from experts to consumers, in the way that was posited at least. One country, however, stood out in the analysis – namely, the UK. These findings require interpretation.

Due to limitations on the public availability of the Consensus Forecasts data, the samples available for study here have been relatively short ones; furthermore, they have been restricted to cover a period with relatively low variability in terms of the inflation environment. Such conditions make insignificant results unsurprising, and these were indeed widespread in the analysis above. The other face of the coin, of course, is that they make significant results all the more impressive. In this respect, the UK appears to be a “model student”, as was already noted, matching the US results of Carroll (2003) in terms of: the relative intrinsic predictive ability of consumers and professional forecasters; providing sensible (although not significant) coefficients when the possibility unidirectional Granger causality from expert to consumer expectations was examined; displaying characteristics of a well-specified Carroll-type sticky information model despite the simplicity of the specification involved; and confirming the importance of news coverage in the transmission of expert to consumer expectations.

The differential ease with which such results could be obtained for the UK relative to continental European countries and their proximity with those holding for the US suggests a number of hypotheses:

1. Institutional considerations

One possible explanation for the exceptional UK results relative to those for continental Europe is that British consumers are *better informed* or at least *are informed on a more continuous basis* than their continental counterparts regarding inflation, making it easier to detect the effects posited by Carroll (2003) over even relatively limited data samples.

A reason for this may simply lie in particular historical circumstances which may have brought inflation to the forefront of public consciousness in the country. After the high inflation of the 1970s, the topic became a major political issue and occupied the centre stage of the public scene. In response to this, the Thatcher government applied a set of much-publicised policies branded as the “Medium-Term Inflation Strategy”, which notably aimed to control the growth of monetary aggregates in an attempt to bring inflation under control¹²². Inflation subsided but spiralled upwards once again in the wake of the credit-driven Lawson boom of the late 1980s. ERM membership of the country and the associated tight monetary policy helped secure a rapid disinflation, although at the cost of prolonging the recession. Since then, inflation has remained an important news item, most notably due to the announcements of various inflation targets in 1992, 1995 and 2003, as well as the institutional arrangements surrounding the establishment of the Bank of England’s independence (see Table 81 on page 451 for a more detailed account). This has undoubtedly done much to shape the public

¹²² Although see Hendry (2002) for some empirical evidence as to the *actual* success of this strategy.

consciousness of inflation, in a way that is historically closer and presumably more present than the traumatic 1923 hyperinflation that shaped the German consciousness in this respect, for instance.

Another possible reason is linked to differences in the UK and US labour markets relative to those in continental Europe over the relevant sample period. These labour markets are relatively decentralised (Nickell et al., 2005; Ochel, 2001), as is confirmed by Table 72 (which presents averages over the period 1994q1-2005q2). They have low levels of union density; most remarkably however, they have by far the lowest levels of adjusted bargaining coverage and the lowest possible score in terms of the bargaining coordination index¹²³. This implies that wage bargaining will be decentralised and carried out in a less discontinuous fashion. Accordingly, this may imply that information about inflation will be relatively more widely and more continuously diffused over time.

A further contributory factor to our findings may lie in the housing market. As is well-known, the UK housing market is characterised by a high rate of owner-occupied housing, with floating-rate mortgages being the predominant type. Table 72 shows that only Spain has a higher rate of owner-occupied households as a proportion of total households in our sample (Oswald, 1996; Nickell and Nunziata, 2001). Furthermore, the table also provides information regarding the *predominant* type of household mortgage interest rate (Borio, 1995; Debelle, 2004; ECB, 2003). The UK stands out as one of the few countries displaying a predominance of floating mortgages in our sample, with Spain and Sweden. In fact, it can be said that “*the proportion of fixed - as opposed to variable - rate mortgages is markedly low in the UK in comparison with many European countries*”

¹²³ Please refer to Section 3.6 of Chapter 2 for precise definitions of these concepts.

and the USA¹²⁴ as noted by HM Treasury, an observation which prompted the commissioning of a review of the UK fixed rate mortgage market to Professor David Miles (2004). Such circumstances may lead consumers to pay even closer attention to the situation of inflation in the economy, and contribute to the findings above.

TABLE 72: LABOUR AND HOUSING MARKET INDICATORS

Country	Labour Market			Housing market	
	Union density	Coverage	Coordination	Mortgages	% Owner-Occupied
United States	15.0%	17.0%	1.0	Fixed	64%
United Kingdom	30.6%	36.9%	1.0	Variable	68%
Belgium	55.6%	90.0%	2.0	Fixed	62%
Spain	14.8%	78.0%	2.2	Variable	78%
France	9.7%	96.6%	1.5	Fixed	54%
Italy	35.6%	83.9%	2.5	Mixed	67%
Sweden	82.9%	89.0%	2.0	Variable	42%
Germany	25.6%	92.0%	2.5	Fixed	38%
The Netherlands	23.8%	86.9%	3.0	Fixed	44%

NOTES: the figures for union density, adjusted bargaining coverage and coordination are calculated as averages over the period 1994q1-2005q2, except for the US where their 1995 values are displayed. Sources include the LMIDB database of Nickell and Nunziata (2001), Ochel(2001) and updates by the author (see Chapter 2 for details). The predominant type of mortgages is derived from Debelle (2004), Borio (1995) and ECB (2003). The % of owner occupied households is expressed as a proportion of total households and the data are derived from Oswald (1996) and the LMIDB.

2. Differential expert ability

Another possible explanation behind the exceptional UK results may be related to the forecasting ability of its experts relative to that of their continental counterparts. Higher quality forecasts by experts would obviously contribute to their salience in the non-nested test carried out against the predictive ability of consumers (Table 65, page 356). It may also facilitate their diffusion among other economic agents, their accuracy making them more widely reported and also increasing the incentives of agents to use them. This

¹²⁴ HM Treasury (2003), “Improving Housing Supply And Simplifying Planning”, Press Notice 09, 9 April 2003. Emphasis added.

may go some way towards explaining the good performance of Carroll's version of the sticky information model in the UK (Table 67), as well as the importance of the news channel in diffusing expert forecasts (Table 70).

It does indeed turn out that the forecasts made by British professional forecasters are superior in RMSE terms to those available for the other countries in our sample. Indeed, as was already shown in Table 63 (page 353), the RMSE for British forecasts is close to 40% lower than that of the closest "competitor" in this respect (the Netherlands). It may simply be that professional forecasters in the UK are using better econometric models than those on the continent. However, they may also be helped somewhat by the circumstances. This possibility is considered in the following point.

3. Differential ease of forecasting due to economic environment

The inflation environment in the UK over the period covered by our sample can be aptly characterised as stable¹²⁵. The shocks affecting the economy over this period have been relatively moderate, and were countered rather effectively by the monetary policy framework in place. Furthermore, the transparency of the latter may well have increased the understanding of policy among private forecasters, a development which would in turn have been reflected positively in their output. In contrast, all of the other countries in our sample (excepting Sweden) have been subject to the uncertainty surrounding the introduction of EMU and the associated destabilising effects for forecasting, one of which has been studied earlier – namely, its impact of consumer perceptions and expectations (see Section 2 above). Indeed, even ignoring the temporary shocks

¹²⁵ Accordingly, naïve forecasts perform relatively well in the UK – see TABLE 63.

associated with the EMU transition, it can be argued that the forecasting problem has become more complex in the Eurozone, as the inflation process in any given country becomes increasingly connected with that in the others.

Thus, the results in this section have provided new evidence on the relationship between expert and consumer forecasts, raising in the process interesting questions as to the exceptional behaviour of the UK model. The availability of longer data samples in the future will be invaluable to explore these issues further, especially as the Eurozone economies reach a steady state after EMU.

Having by now studied the *mean* of consumer expectations in some detail, it seems appropriate to consider what may be learnt from their *distribution*. This is the object of the next section.

6 THE ECONOMICS OF DISAGREEMENT

6.1 Introduction

In their seminal paper, Mankiw, Reis and Wolfers (2003) (henceforth MRW) present a strong case for the consideration of *heterogeneity* in inflation expectations. The authors make two key points in this study.

Firstly, textbook macroeconomic theory assumes only too often (whether explicitly or implicitly) that agents have a common expectation for future inflation, formed from a common information set¹²⁶. This worldview is in sharp contradiction with the data, a message that will be considerably reinforced below. People *do* indeed disagree about inflation expectations.

Secondly (and crucially) ignoring the dispersion of the expectations distribution can only lead to a poor understanding of the macroeconomy. The authors notably document the co-movement of inflation expectations with other aggregate variables, leading them to the conclusion that its disagreement dimension is “*itself an interesting variable for students of monetary policy and the business cycle*” and that it may even be “*a key to macroeconomic dynamics*”. Indeed, in earlier work, Mankiw and Reis (2002) put forward a “sticky information” model of expectations formation centred around the idea of disagreement, leading to a very different model of inflation from that provided by the benchmark of the New Keynesian Phillips Curve.

¹²⁶ There are exceptions, of course. Lucas’ (1973) “islands model” is probably the most famous, although the extent of disagreement in the model is generated exogenously.

This section aims to address a number of issues for which the distribution of expectations has some bearing. Much of this research is made possible through the availability of quantitative Gallup data, which provide distributional information between January 1982 and January 1997.

First, we will present the sticky information model of Mankiw and Reis (2002) in Section 6.2 and contrast one of its key predictions with the behaviour of the data in the context of two extraordinary disinflation processes: the Volcker disinflation of 1979-1982 in the US and the ERM disinflation of 1991-1992 in Britain. We will argue that the data contradict the model of the authors partially, due to their reliance on the rational expectations hypothesis. Section 6.3 then provides a critique of claims in van Lelyveld (2000), according to which the expectations question of the EC Consumer survey data *builds-in* bimodality into the survey results. Following this, Section 6.4 will seek to ‘explain’ empirically the extent of disagreement in the data by considering a set of variables highlighted by MRW. Section 6.5 will turn to the question of whether measures of disagreement available through a quantification of qualitative data using the Carlson-Parkin method can be considered to be reliable. Finally, Section 6.6 provides an in-depth qualitative examination of the distributional evidence available to us; the hypothesis that consumers form their expectations *qualitatively* to a certain extent is explored, as is the impact of policy changes (e.g. inflation targeting) on the expectations distributions.

6.2 The “Sticky Information” model of Mankiw and Reis (2002)

6.2.1 Overview

In a seminal paper, Mankiw and Reis (2002) introduced a model of “sticky information”, managing to retain rational expectations whilst correcting for several counterfactual predictions of more traditional New Keynesian Phillips Curves.

They argued that costs of information acquisition or reoptimisation¹²⁷ implied a slow diffusion of information in the economy, a concept which the authors formalised through the Calvo (1983) geometric structure: at any period in time, only a fraction λ of the population updates its information set and computes economic plans on this basis; the rest continue to act on the basis of outdated information. Crucially, when updating does take place, it does so *according to rational expectations*. Mankiw et al. (2003) implemented this idea in the context of inflation expectations, implying the following equation for the aggregate expectation $\bar{\pi}_{t+12}^e$:

$$\bar{\pi}_{t+12}^e = \lambda \sum_{j=0}^{\infty} (1-\lambda)^j \hat{\pi}_{j,t+12} \quad (82)$$

Where j denotes how many months ago the information set was last updated for a proportion $\lambda(1-\lambda)^j$ of agents and $\hat{\pi}_{j,t+12}$ is the rational forecast made j months ago for inflation between t and $t+12$. The latter is approximated by the forecast from a VAR model.

¹²⁷ Reis (2005) has recently provided some microfoundations for this “inattentiveness”.

Carroll (2003) proposes a model which can be interpreted as a microfoundation for the research by Mankiw et al. People are assumed to obtain their macroeconomic views from information supplied in the media, which is absorbed probabilistically. In particular, the updating of expectations is once again assumed to follow the Calvo (1983) geometric structure. The author assumes that the economy has at any point an underlying “fundamental” inflation rate π_t^f , which follows a random walk, justifying this by what looks like (near) unit root behaviour of inflation in the post-1959 period:

$$\begin{aligned}\pi_t &= \pi_t^f + \varepsilon_t \\ \pi_{t+1}^f &= \pi_t^f + \eta_{t+1}\end{aligned}$$

Where ε_t and η_t are transitory and permanent shocks, respectively. The author further assumes that experts have some ability to estimate ε through period t and η through period $t+1$. The general public has access to expert forecasts through newspaper articles, which are assumed to contain a projection of the fundamental inflation rate over the coming year. The object of these particular assumptions on consumer beliefs is to rigorously obtain an equation for expectations that can be taken to the data. Indeed, the Mankiw et al. updating assumptions on their own, when considered in this context, imply the following for consumer inflation expectations over the coming year ($M_t[\pi_{t,t+4}]$):

$$M_t[\pi_{t,t+4}] = \lambda N_t[\pi_{t,t+4}] + (1 - \lambda)M_{t-1}[\pi_{t,t+4}]$$

Where N_t denotes newspaper forecasts. However, the quantity $M_{t-1}[\pi_{t,t+4}]$ is generally not available from the data. Carroll's assumptions imply that $M_{t-1}[\pi_{t,t+4}] = M_{t-1}[\pi_{t-1,t+3}]$ ¹²⁸ and the previous equation is therefore usefully rewritten as:

$$M_t[\pi_{t,t+4}] = \lambda N_t[\pi_{t,t+4}] + (1 - \lambda)M_{t-1}[\pi_{t-1,t+3}]$$

This states that the mean measured consumer inflation expectations for the coming year is a weighted average of the current newspaper expectations N_t and last period's mean measured consumer inflation expectations. We recognise the shape of equation (76), which was already tested in Section 5.4.

Although the Carroll and Mankiw/Reis stories are similar, they are certainly not equivalent – unless one equates professional forecasts with rational expectations¹²⁹. Section 5.4 provided some support for Carroll's model in the case of the UK but failed to do the same in the other countries in our sample. We now consider the performance of the Mankiw/Reis version of the model.

6.2.2 Testing the model across the Atlantic: the Volcker and ERM disinflations

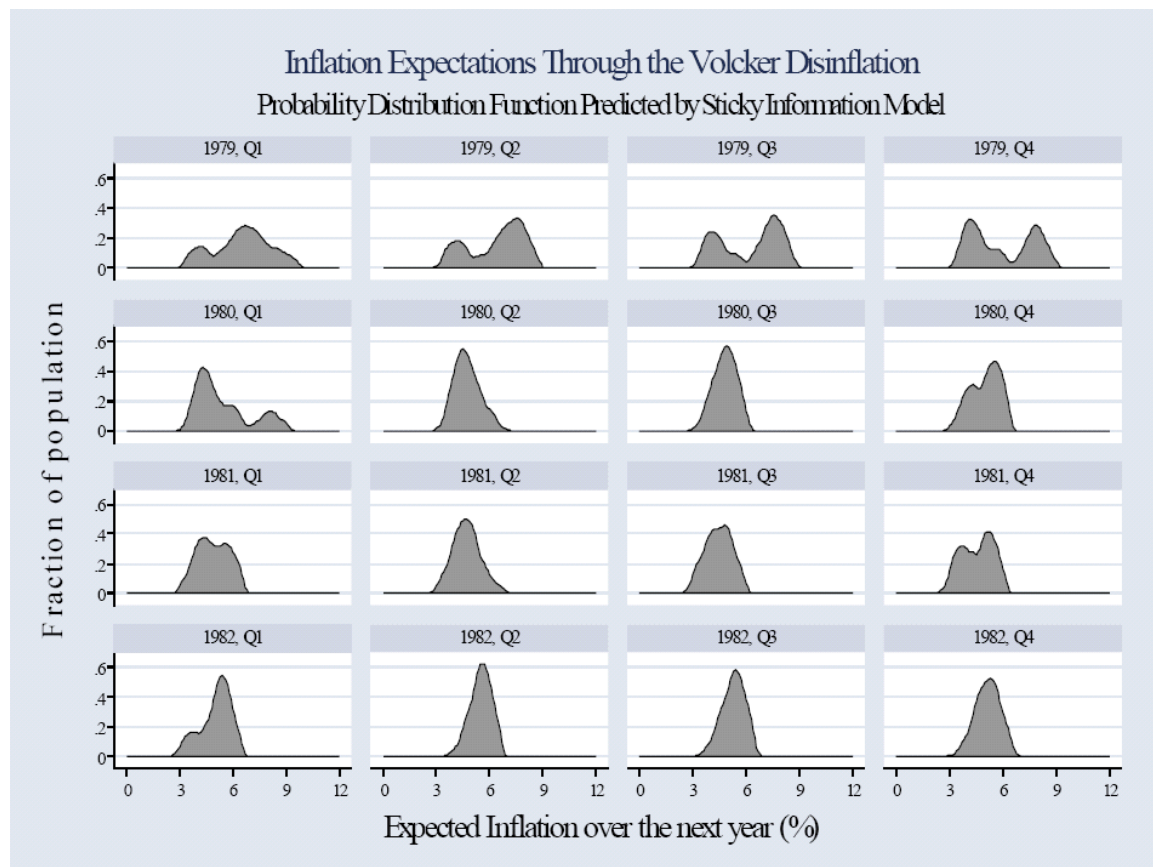
As a test of the sticky information model, MRW (2003) offered a case study of the “Volcker disinflation”, which followed the appointment of Paul Volcker as chairman of the Board of Governors of the Federal Reserve Bank in August 1979. The period following his appointment is of particular interest as it witnessed a very sharp fall in

¹²⁸ Carroll notes correctly that even if one chooses to disbelieve his particular derivation, there is likely to be a very high degree of correlation between those two terms, so that the difference may not be of consequence in practical terms.

¹²⁹ This is a leap that the history of economic forecasting patently fails to justify as a general rule; see Clements and Hendry (1999).

inflation from 11% to around 4% in 1982; the empirical consensus is that there was a serious structural break towards more conservative policy (e.g. Clarida et al., 1998). In such a context, the sticky information model makes clear predictions as to the changes that one should expect for the distribution of inflation expectations. The predictions of the model are presented in Figure 104. These were generated by approximating the ‘rational forecasts’ required by the sticky information model as the output of a VAR including monthly CPI inflation, the interest rate on 3-month Treasury bills and a measure of the output gap. The parameter λ of the model was calibrated by maximising the correlation between the interquartile range of inflation expectations in the survey data with that predicted by the model, leading the authors to set $\lambda=0.10$.

FIGURE 104: VOLCKER DISINFLATION, “STICKY INFORMATION” PREDICTIONS



SOURCE: Mankiw, Reis and Wolfers (2003), p.47.

A number of distributional predictions of the model for an episode of strong disinflation are clear from the figure:

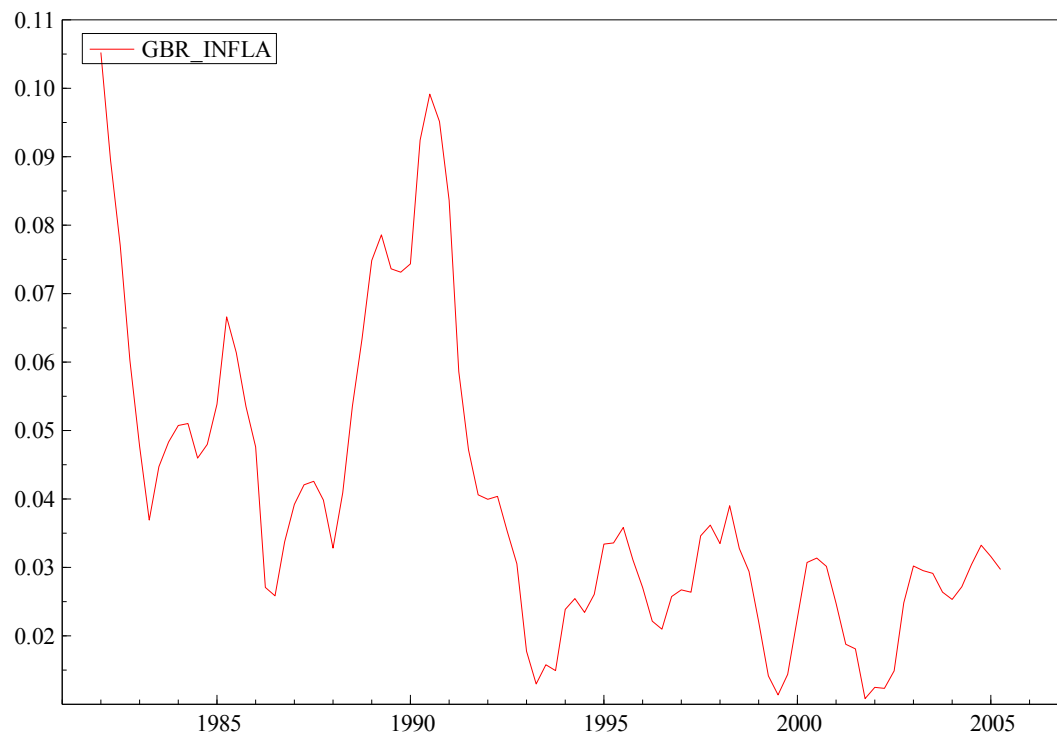
1. The shift towards the new distribution is gradual in nature
2. The distribution of expectations flattens temporarily – its dispersion increases
3. In the process, the distribution becomes clearly bimodal, with a very visible trough between the two modes.

When contrasting these predictions with the actual developments in the data from the Michigan survey of consumers, the authors recognise that the model generates predictions that are “*too sharp*”. However, this thought is taken no further and overall, MRW conclude that the model “*successfully accounts for broad features of the evolution of the distribution of inflation expectations during the Volcker disinflation*”¹³⁰.

Before examining these claims, we propose to throw new light on the issue through the consideration of an even sharper disinflation episode across the Atlantic – that associated with Britain’s membership of the Exchange Rate Mechanism (ERM), starting on 8 October 1990 and ending with ‘Black Wednesday’ on 16 September 1992. From a starting point of over 10% in August 1990, inflation fell to roughly 1.7% by January 1993. Most of the decline took place in 1991, with inflation being reduced to 4% by January 1992. Figure 105 shows these developments graphically.

¹³⁰ Mankiw, Reis and Wolfers (2003), p.48.

FIGURE 105: UK CPI INFLATION, ALL ITEMS; RECENT HISTORY



The data collected monthly by Gallup on samples of British employees – which were already presented in Section 5.1 of Chapter 1 – contain some distributional information that allows us to study the evolution of disagreement through this historical episode. In particular, the *Gallup Political Index* and *Gallup Political and Economic Index* reports provide the proportions of respondents falling into a number of quantitative ranges. Table 73 lists these by providing the data for January 1991 as an example.

TABLE 73: UK INFLATION EXPECTATIONS, DISTRIBUTION DATA, 1991m1

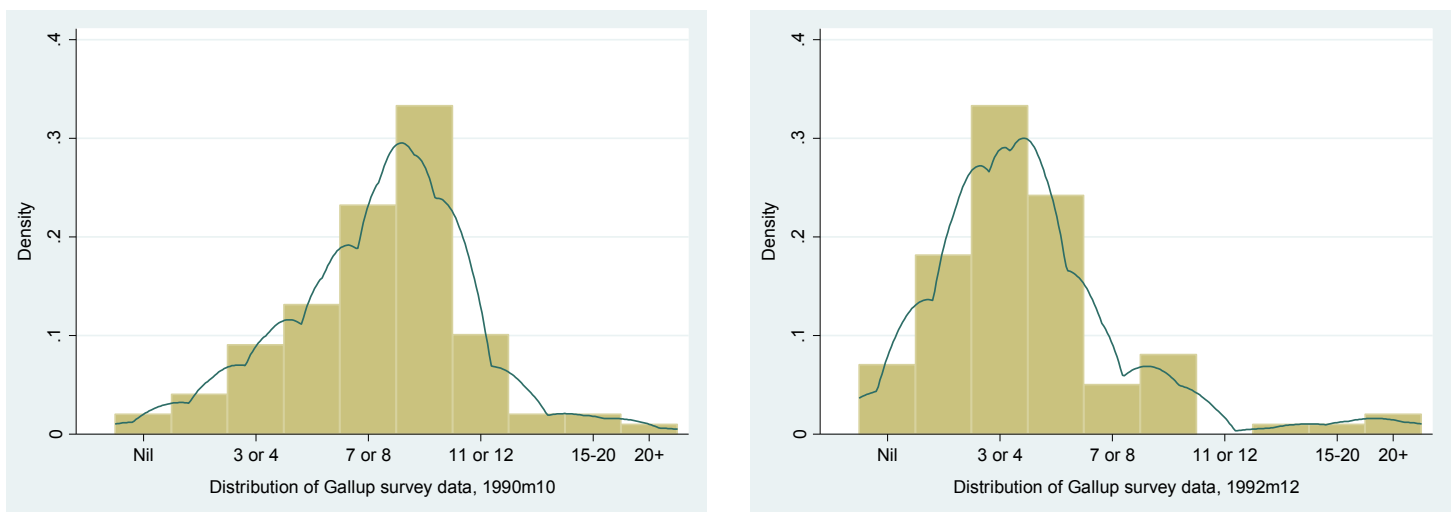
Category	Proportion Responses (%)	Category	Proportion Responses (%)
Nil	2 %	11 or 12 %	6 %
1 or 2 %	3 %	13 or 14 %	2 %
3 or 4 %	4 %	15-20 %	4 %
5 or 6 %	14 %	20+ %	1 %
7 or 8 %	16 %	Don't know	20 %
9 or 10 %	28 %		

SOURCE: Gallup Political and Economic Index Report, January 1991

For the following exercise, we dropped the “don’t know” category and only considered the answers of the respondents having addressed the question, re-weighting the proportions so that they would sum to 100. We propose to analyse the changes in the distribution of expectations from the start of the disinflation to the end in some detail. To provide some guidance as to the shape of the underlying distribution, we estimated univariate kernel densities and plotted them alongside the relevant histograms. We used the Epanechnikov kernel function to perform these particular calculations.

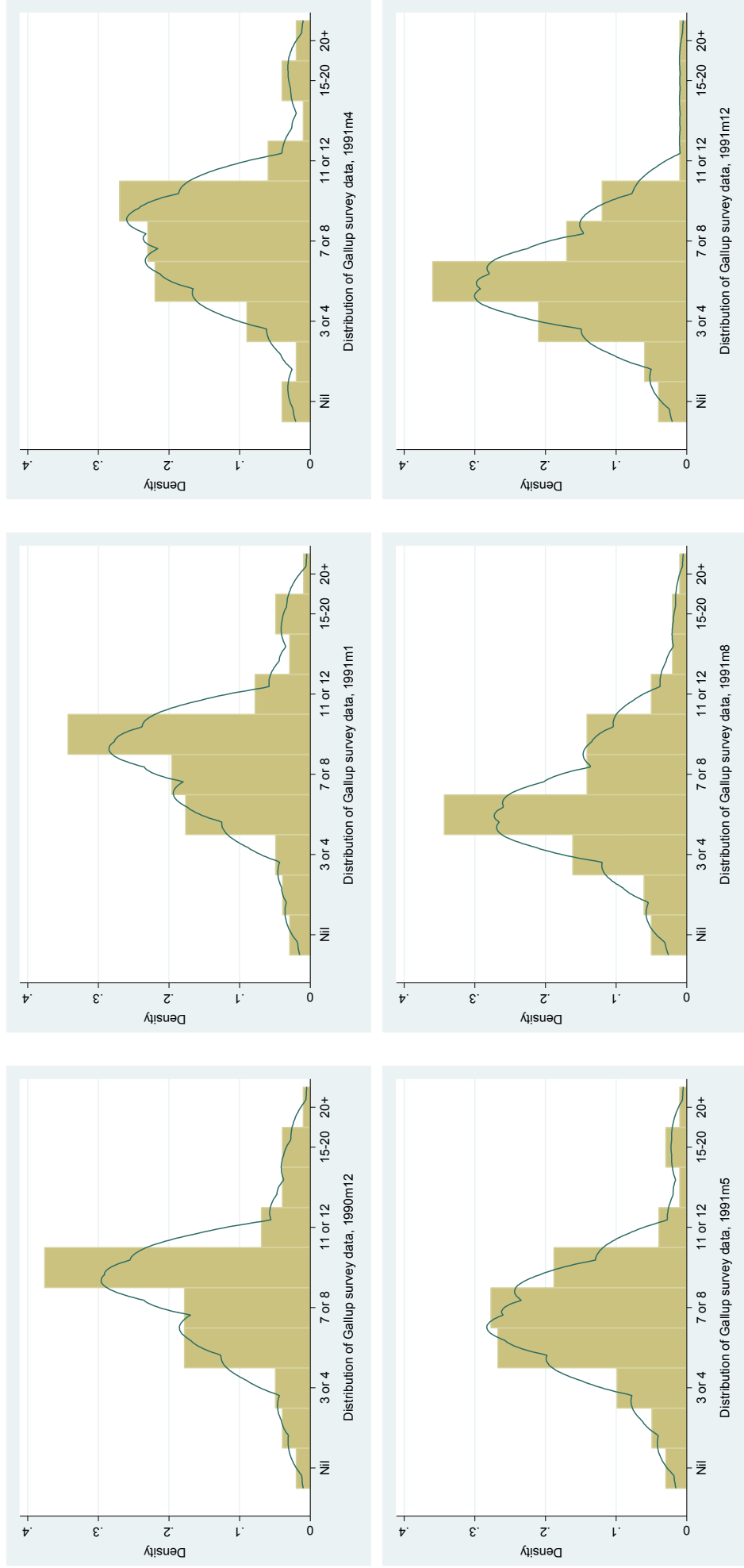
Figure 106 shows the distribution of expectations at the start (October 1990) and end (December 1992) of the disinflation process. Both distributions appear to be unimodal, but with very different modes; whereas at the start of the transition, the mode could be located around 9%, by the end this had shifted all the way down to between 3 and 4%. Figure 107 examines a number of key points between these two pictures that shed light on the issues at hand. We focus on the year 1991, as the sharpest part of the disinflation occurs during it, thereby yielding the clearest insights.

FIGURE 106: UK EXPECTATIONS, ERM DISINFLATION, ENDPOINTS



SOURCE: Gallup Political and Economic Index Reports, various issues, and own calculations

FIGURE 107: UK EXPECTATIONS DISTRIBUTION & THE ERM DISINFLATION, DETAIL



SOURCE: Gallup Political and Economic Index Reports, various issues, and own calculations

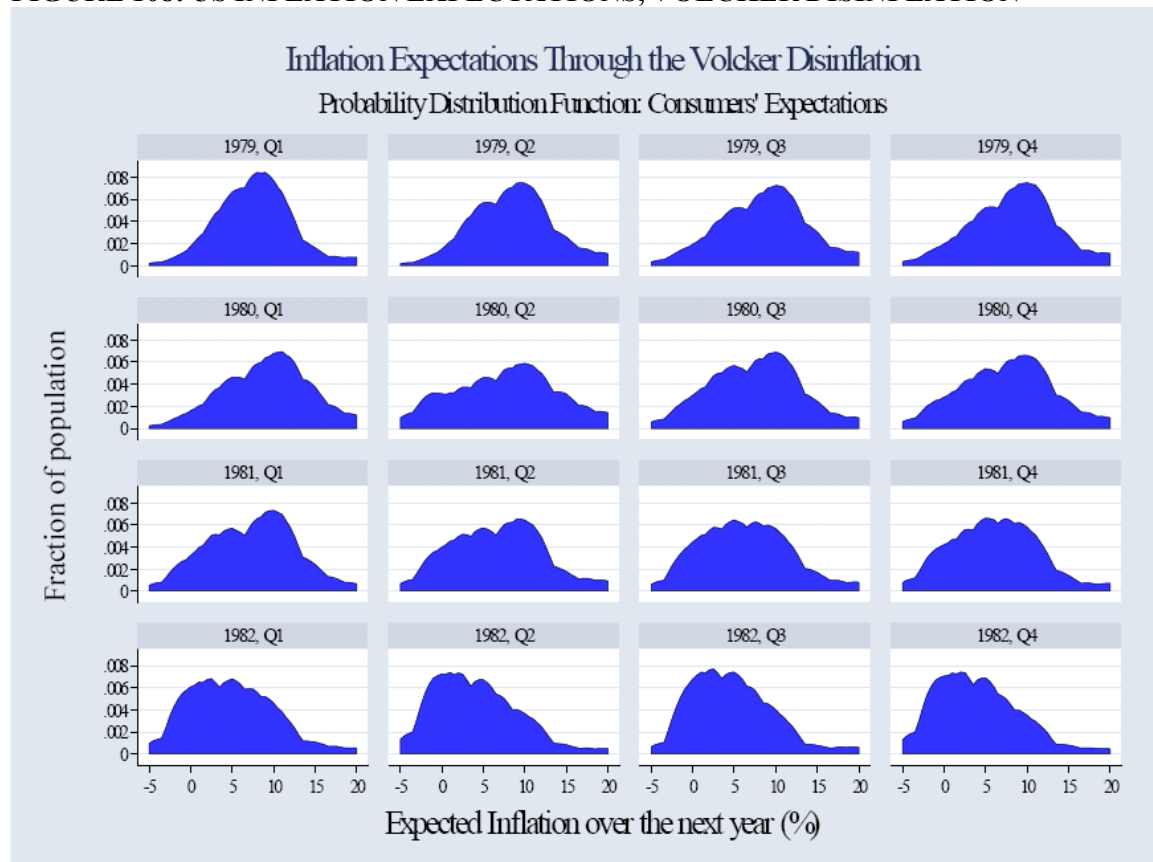
Contrasting the distributions for October and December 1990, it appears that the mode at 9-10% inflation has increased despite the start of the disinflation, a phenomenon which could perhaps reflect some momentum in inflation expectations. We note however that the categories “5-6%” and “7-8%” are also attracting a significant share of responses, and that the former category has indeed increased noticeably between these two dates. The evolution over the following three plots – concerning January, April and May 1991 – is of particular interest. Indeed, there appears to be a very clear and gradual shift of responses from the “9-10%” category to the two categories immediately below it, in particular. As a result, the mode of the distribution is decreasing, a point reinforced by the plots relating to August and December of that year. Thus, by the end of 1991 the mode of the distribution was located at about 5%.

A number of comments can be drawn from this episode regarding the three predictions of the sticky information model:

1. The shift towards the new distribution does indeed appear to be gradual, in accordance with the model.
2. The distribution of expectations does flatten somewhat during the transition process, a development that is also in accordance with the theory.
3. However, there is no evidence for the strong bimodality developments that are so clearly predicted by the theory.

Exactly the same conclusions can in fact be drawn from a consideration of US consumer inflation expectations from the Michigan survey during the Volcker disinflation. This is clearly visible from Figure 108.

FIGURE 108: US INFLATION EXPECTATIONS, VOLCKER DISINFLATION



SOURCE: Mankiw, Reis and Wolfers (2003), p.46.

Thus, in both episodes the data appear to seriously contradict the bimodality implications of the model. MRW do note that the predictions of the model are “too sharp”; however, they do not consider that such findings actually imply a different shape for the process of expectations updating.

The reason behind the bimodality prediction of the model is essentially the assumption of updating to a ‘rational forecast’ when an opportunity becomes available. Thus, at the beginning of a strong disinflation process, the fraction λ of agents that have an opportunity to update their expectations do so by dramatically reducing their forecast. This leads to the creation of a second, lower mode in the distribution. This new mode keeps gaining probability weight as more and more agents are able to update their forecasts, generating a very visible bimodal distribution with a clear trough between the modes. Eventually, when most agents have had a chance to accomplish an update, the distribution becomes unimodal once again, at the lower mode.

The actual distributional data are consistent with MRW’s intuition that different agents hold different expectations, which appear to be based on information that is more or less recently updated; there does indeed seem to be a gradual adjustment process to new conditions. However, the process seems to be more *adaptive* than *rational*. The probability mass in the plots shifts downwards, *accompanying* rather than *anticipating* the fall in inflation.

This view is fully consistent with the behaviour of the Carlson-Parkin quantified expectations, for which we have already emphasised the presence of a strongly adaptive component (e.g. Section 2.1.3). It is also consistent with the underlying data on response proportions from the EC Consumer Survey, notably the sample analysed in Section 2 of the present chapter. There, we noticed the smoothness of the adjustment of response proportions across the different categories, with the exception of the “expectations effect” associated with the introduction of physical currency in Eurozone countries,

discussed in detail earlier. The view that agents update gradually to less-than-rational forecasts would also address another criticism of Mankiw and Reis' (2002) work. Indeed, when considering their particular form of expectations updating in a Phillips curve framework the authors are forced to rely on serial correlation of money to explain inflation persistence (Mash, 2004), which is a somewhat ad hoc justification.

Section 6.4 will turn to another aspect of the MRW research project, namely seeking the fundamental causes behind the evidence of disagreement among agents in the data.

6.3 Double-peakedness through survey design: the van Lelyveld hypothesis

Whereas MRW (2003) offered an argument for temporary bimodality in the expectations distribution based on an idea of “inattentive agents” and rational updating, van Lelyveld (2000) argues in the context of the EC Consumer Survey that *“the particular way in which the questions for this survey have been posed forces double-peakedness in the data”* and that this therefore “[...] invalidates any approach based on fitting a single-peaked distribution invariant to current inflation rates”¹³¹. These arguments have been noticed and mentioned as a possible caveat to the measures of expected inflation obtained from the EC Consumer survey by Berk (2002), a point also highlighted more recently in Berk (2006).

This section aims to carefully examine van Lelyveld's hypothesis to determine whether this call for caution is warranted. Section 6.3.1 presents the arguments and empirical

¹³¹ Van Lelyveld (2000), page 61.

work put forward in van Lelyveld (2000). Section 6.3.2 provides a critique of both dimensions. Section 6.3.3 provides a conclusion on this matter.

6.3.1 The arguments and evidence presented in van Lelyveld (2000)

The author argues that the structure of question 6 of the EC Consumer Survey (presented in Table 1, page 17) forces the “middle category” to be empty at low rates of inflation (that is, the category corresponding to an expectation of prices increasing at a slower rate). This, argues the author, is due to the presence of two quantitative “anchors” in the question – an expectation of zero inflation (option 4) and one of inflation continuing to rise at the current perceived rate (option 2). As inflation drops, both “anchors” come closer together, with the implication that the *“range covered by the middle question becomes progressively smaller”*¹³² in potential distribution space. Furthermore, *“there is no reason to suspect that the distribution space covered by [answers 2 and 4] becomes smaller”*¹³³, leading the author to expect that the distribution should become double-peaked as inflation falls.

Van Lelyveld (2000) carries out an empirical exercise to check this hypothesis, consisting of the following steps:

- 1) The author marks on a horizontal line with the units of inflation an equidistant series of points representing each survey answer, from that representing deflation (denoted by C_1) to that representing an expectation of inflation rising at a faster rate (denoted by C_5).

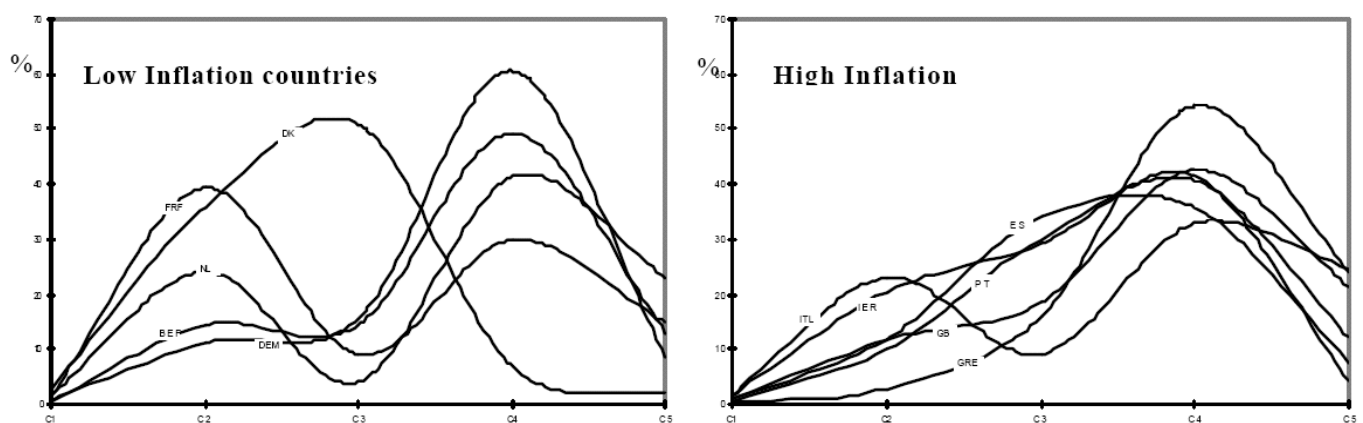
¹³² Van Lelyveld (2000), page 66.

¹³³ Van Lelyveld (2000), page 66.

- 2) For the data collected in a given month, the author plots vertically above each of these points the percentage of responses relating to the option represented. A spline function is then fitted to the data.
- 3) The results are plotted separately for low and high inflation countries, and the hypothesis is checked visually. Low-inflation countries should be characterised by more pronouncedly bimodal plots.

The plot resulting from steps (1) to (3) is reproduced as Figure 109 and concerns data for January 1995. The “low inflation” countries are France, Belgium, Germany, the Netherlands and Denmark; the “high inflation” group is constituted by Spain, Portugal, the UK, Greece, Ireland and Italy.

FIGURE 109: VAN LELYVELD’S DISTRIBUTION OF EXPECTATIONS (1995M1)



SOURCE: van Lelyveld (2000), Figure 6.3, p.67.

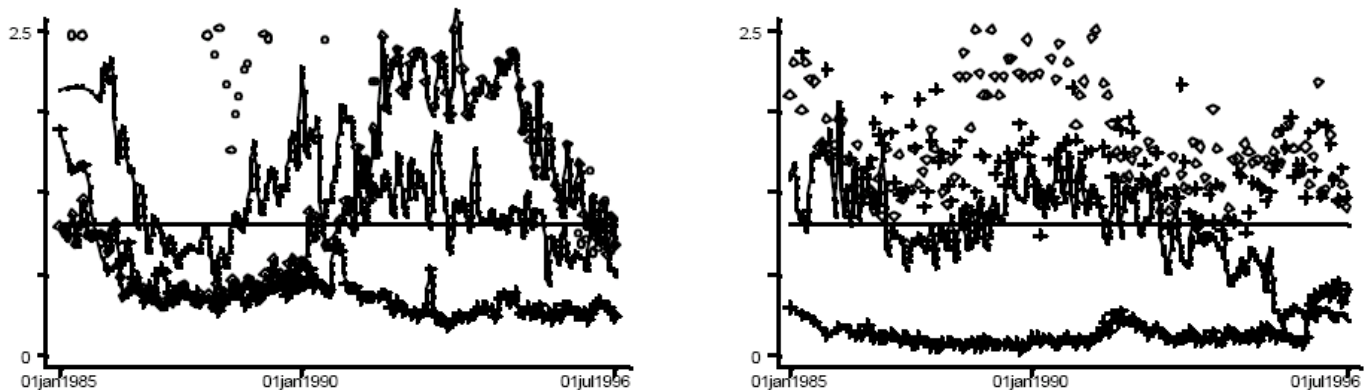
The author claims that the picture demonstrates the presence of double-peakedness for all countries catalogued as “low inflation” – except Denmark – and that this property also holds for Italy, despite it being a “high inflation” country.

As this exercise only concerned the data for a particular month, van Lelyveld constructs an index which indicates double-peakedness for values below 1:

$$DPINDEX = \left[\frac{answers3}{\min(answers2, answers4)} \right]$$

This index is simply the ratio of the proportion of responses in the middle category and the proportion of responses in the smaller of the two neighbouring categories. Figure 110 plots the time variation of this index for a number of countries, the bold horizontal line marking the location of 1.

FIGURE 110: VAN LELYVELD INDEX OF DOUBLE-PEAKEDNESS THROUGH TIME



SOURCE: van Lelyveld (2000), Figure 6.3, p.68. The left pane shows Belgium (line), France (line with + symbol), Germany (line with $\diamond$ symbol), and Greece (dots). The right pane shows Ireland (+ symbol), Italy (line), the Netherlands (line with + symbol), and the United Kingdom ($\diamond$ symbol). Denmark, Portugal, and Spain are not included since double-peakedness never occurred in the sample-period.

The author claims that these plots reveal widespread and persistent double-peakedness in the distributions of expectations for a number of countries. The author further regresses the index on a constant and ‘recent inflation’, proxied by a 3-month moving average. He finds that almost all of the slope coefficients are positive, showing that the index is low when inflation is low, which is argued to represent a confirmation of the working hypothesis. Finally, these exercises are reproduced for the EC Consumer Survey’s question on perceptions. This question is considered much less problematic by the author (as it includes only one clear quantitative “anchor”) and the findings regarding double-peakedness are indeed much milder. Accordingly, we focus in what follows on the expectations question, but everything translates straightforwardly to the case of perceptions.

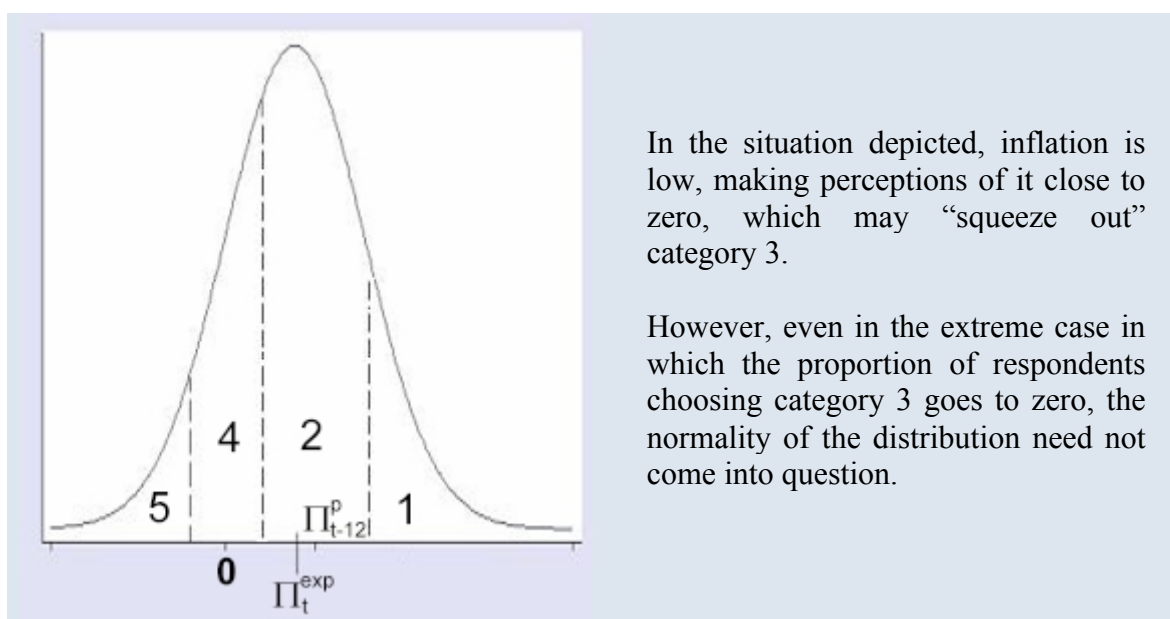
6.3.2 A critique of the van Lelyveld (2000) double-peakedness hypothesis

We will now argue that the theoretical arguments advanced above are unsound, and that the empirical methodology deployed fails to prove the case being made in any event.

On the theoretical front, the argument confuses the issue of the “space” occupied by the different categories with the issue of the distribution’s shape. Indeed, as remarked by van Lelyveld, at very low levels of inflation the answer category 3 becomes *irrelevant*. However, that is not the same thing as saying that the *distribution* is not normal anymore. It is perfectly possible to have a normal density function under which the probability space occupied by answer categories 2 and 4 is *adjacent*, with the “middle category” having been fully squeezed to zero probability mass, as it becomes undistinguishable

from its two neighbours; Figure 111 depicts such a situation. Indeed, it is straightforward to see that the derivation of the mean of the inflation expectations distribution which results in equation (10) on the basis of equations (1) to (4) ¹³⁴ is fully robust to such a development, provided of course that the neighbouring categories do move to occupy all the probability space freed by the collapse of the middle category.

FIGURE 111: EXAMPLE OF NORMAL EXPECTATIONS DISTRIBUTION



The arguments in favour of double-peakedness are damaged even further by the failure of the claim that “*there is no reason to suspect that the distribution space covered by [answers 2 and 4] becomes smaller*”¹³⁵. Quite on the contrary, signal extraction theory (Batchelor, 1986), the detailed analysis of the post-EMU data (Section 2.1.2 above) and the empirical results of Chapter 1 regarding the optimal quantification methodology all converge to show that indifference thresholds do indeed shift, and therefore displace the

¹³⁴ See Section 3.1 of Chapter 1.

¹³⁵ Van Lelyveld (2000), page 66.

range of inflation occupied by the probability mass associated with particular response categories. It can be expected that as inflation decreases, consumers become more ‘sensitive’ to inflation, in the sense that it takes a smaller absolute value of inflation (or deflation) for them to report it as such rather than claim it is undistinguishable from zero. This is indeed what has been found in the data of *all* the countries considered in this study. In any event, this argument is unnecessary to counter the case in van Lelyveld, which is fully answered by the previous paragraph.

As a direct consequence of these arguments, the empirical verification of the double-peakedness hypothesis is critically flawed. Indeed, plotting response proportions over a horizontal base which locates survey categories equidistantly can provide no information whatsoever as to the shape of the distribution. Indeed, it is precisely when inflation is low and when the “middle” category’s inflation range is being “squeezed” by the two neighbouring categories that it can stretch even quite a low probability mass relatively high vertically in density terms – in a way that is fully compatible with normality. In fact, the finding by van Lelyveld (2000) of an apparent “trough” for the “middle” category in low inflation countries is *exactly* what one would expect were the distribution actually normal. Thus it is no surprise that many low inflation countries seem to spuriously exhibit “double-peakedness”, both cross-sectionally and through time. As for the link interpreted by the author between the alleged double-peakedness and the rate of inflation – claimed to be demonstrated through the index regressions provided – it is, again, nothing that would not be expected in the canonical normal distribution case. Hence, the exercises carried out fail to demonstrate the hypothesis put forward, itself also flawed.

More fundamentally, it is simply not possible to make points about the shape of the actual distribution solely on the basis of this type of raw data – hence the need to rely on assumptions with the initial guidance of theory.

At most, given a set of assumptions, it is possible to:

- 1) Compare measures from the derived distribution (e.g. the mean) with other economic series (e.g. the variable being forecasted, in this case inflation) and assess whether the quantification assumptions are ‘sensible’.

This has been popular in the literature, which has contrasted a number of distributional approaches in terms of their predictive ability (Berk, 1999; Nielsen, 2003a). The normal, symmetric t, non-central scaled t (based on various asymmetry parameters) and the Chi squared have been among the distributions considered. However, as highlighted in Chapter 1, a strict application of a predictive ability criterion is an inadequate approach outside a world of rational expectations. Nevertheless, it is not uninteresting to note that such exercises have not revealed large differences between the distributional assumptions considered, suggesting a low degree of sensitivity to this dimension, within reason. Indeed, Nielsen (2003a) directly compared the derived mean expectations series, and concluded that there is essentially little to choose from.

Bimodal distributions have not been implemented in these robustness checks, with good reason: it seems impossible to sensibly fit such distributions on the available data. The

third point below presents a more practical alternative to examine whether bimodality arguments have any merit.

- 2) Comparing qualitative and quantitative measures under varying distributional assumptions.

A further, alternative approach, with more theoretical merit, is to compare the measures derived from the qualitative data with their quantitative data counterparts in countries where comparable data of both types are available. This would therefore be an exercise in the spirit of that carried out in Chapter 1 (Section 6). However, given the previous work just highlighted, it is highly unlikely that this would yield any major surprises.

- 3) Using quantitative data to settle “bimodality disputes”.

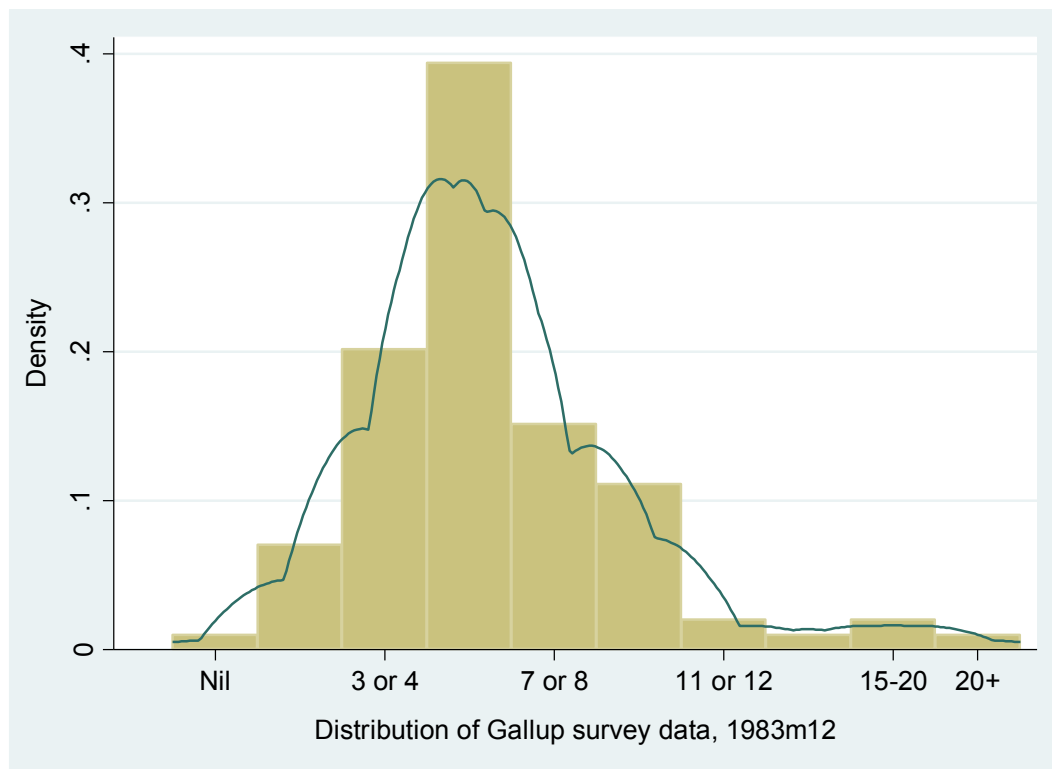
This is the approach taken in Section 6.2.2 regarding the claims by MRW (2003) of bimodality in the expectations distribution arising from inattentive agents.

In summary, the strategy recommended here to test distributional assumptions consists in (1) studying the distribution of quantitative data, (2) comparing qualitative measures quantified on varying distributional assumptions with the quantitative evidence, in case there are any particularities linked with the EC Consumer Survey and (3) carrying out more general ‘sensitivity’ checks relative to other economic variables. On the basis of both previous evidence (as suggested above) and theoretical arguments related to central

limit theory (developed in Chapter 1), the normal distribution appears to be a fully adequate choice. We now confirm this further by examining potential bimodality features in the UK Gallup data more generally than we did when considering the sticky information model of Mankiw and Reis (2002).

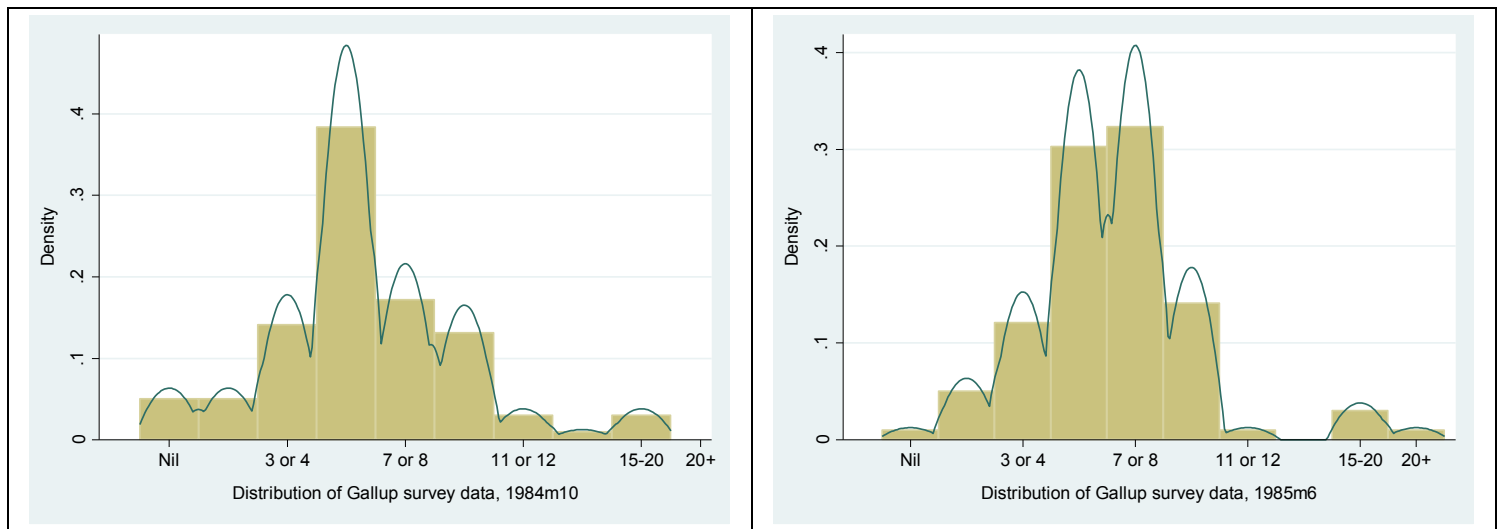
When considering the histogram plots of the Gallup expectations distribution for the 174 months of data available and their associated Epanechnikov kernel approximations, it appears very clearly that for over two thirds of the months the data are perfectly consistent with unimodal behaviour, and normality in particular. Figure 112 is a representative plot for these months:

FIGURE 112: GALLUP EXPECTATIONS DISTRIBUTION, 1983m12



The remaining months are, in fact, also consistent with unimodality and are either normal or can be well-approximated by a normal distribution. They present however certain features that make the Epanechnikov kernel approximation inappropriate. The following figure displays two representative plots.

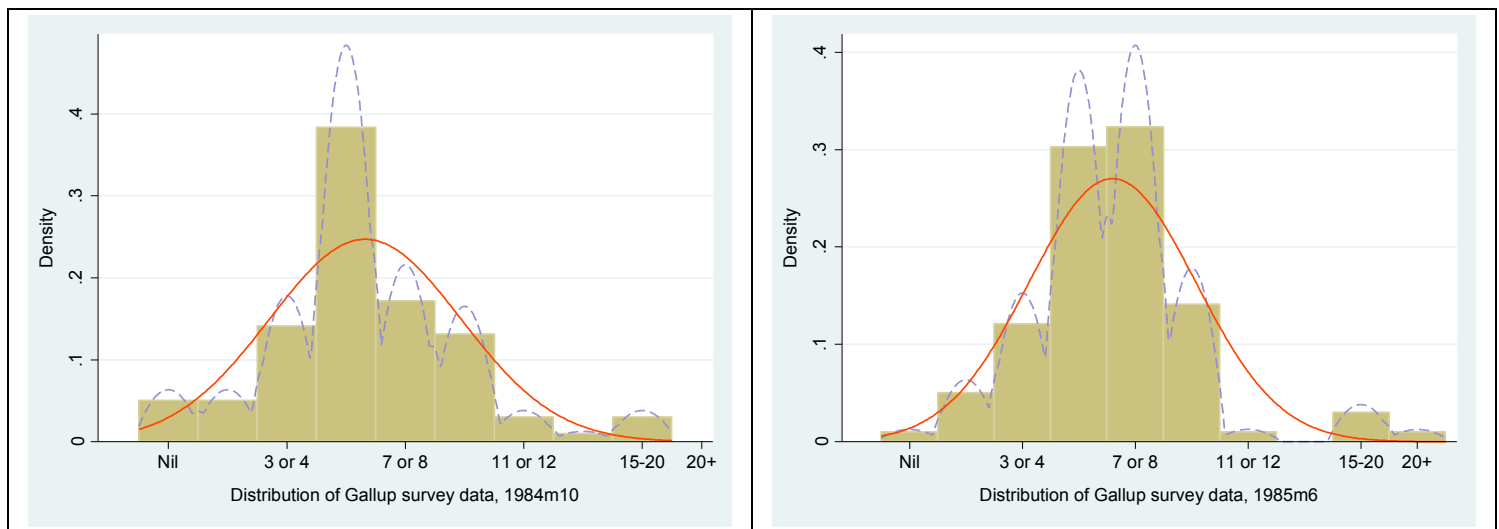
FIGURE 113: GALLUP EXPECTATIONS DISTRIBUTION, 1984m10 AND 1985m6



As can be seen, there is no real issue of bimodality in these plots, despite the behaviour displayed by the kernel. Consider the right hand side panel, for instance. Probability mass is quite densely concentrated around the centre of the distribution (options “5 or 6%” and “7 or 8%”). The overall shape seems to conform very closely to that of a normal distribution with relatively low variance. The reason the Epanechnikov kernel shows otherwise is that it is forced to take into account all the probability mass it encounters equally, forcing it to “fit” the extreme observations on the right of the distribution. However, it is clear that the *overall* distribution would be better approximated by a straightforward normal distribution. A very similar observation can be made for the

distribution displayed on the left hand side panel. In addition however, it appears as if the distribution may be somewhat skewed to the left in this plot. Nevertheless, the best fitting normal approximation still appears to be a very sensible option in this case. These points are reinforced by Figure 114.

FIGURE 114: GALLUP (1984M10 & 1985M6), NORMAL AND KERNEL DENSITY



Therefore, the normal distribution appears overall on the basis of the Gallup data to be an appropriate choice that can be safely used in quantification approaches. In cases where deviations from normality are detected, they appear to be negligible and certainly taking them into account (e.g. through a skewed distribution) would only modify the location of the mean in a very minor way. Furthermore, if anything, the distributional shape arising from qualitative data is likely to more closely resemble a normal distribution than that from quantitative survey data, given phenomena such as “digit preferences” and the like, which are often encountered in such data. Such issues are further discussed below (Section 6.6.4).

6.3.3 Conclusion

Whereas MRW (2003) offered an argument for *temporary* bimodality in the expectations distribution based on an idea of “inattentive agents” and rational updating, van Lelyveld (2000) argued that the formulation of the EC Consumer Survey question on expectations *itself* induced double-peakedness in the data at low levels of inflation. If true, this argument would have momentous implications for quantification methodologies, in that it would invalidate “[...] *any approach based on fitting a single-peaked distribution invariant to current inflation rates*”¹³⁶, in the words of the author.

We have shown here that the underlying theoretical argument behind this claim is flawed, as is the proposed empirical demonstration thereof. Furthermore, the quantitative data – the only type of data which, ultimately, can allow one to discern the shape of the expectations distribution – were shown to display no signs of bimodality, whether linked to MRW’s (2003) inattentive agents story (Section 6.2.2) or more generally.

The exercise undertaken in this section was particularly important given the implications of the argument in question, and the echo it was starting to generate – leading other authors to mention it as a possible caveat to the measures of expected inflation obtained from the EC survey (e.g. Berk, 2002, 2006).

Having refuted this claim, we now turn to other insights that can be gained from the distributional data at our disposal.

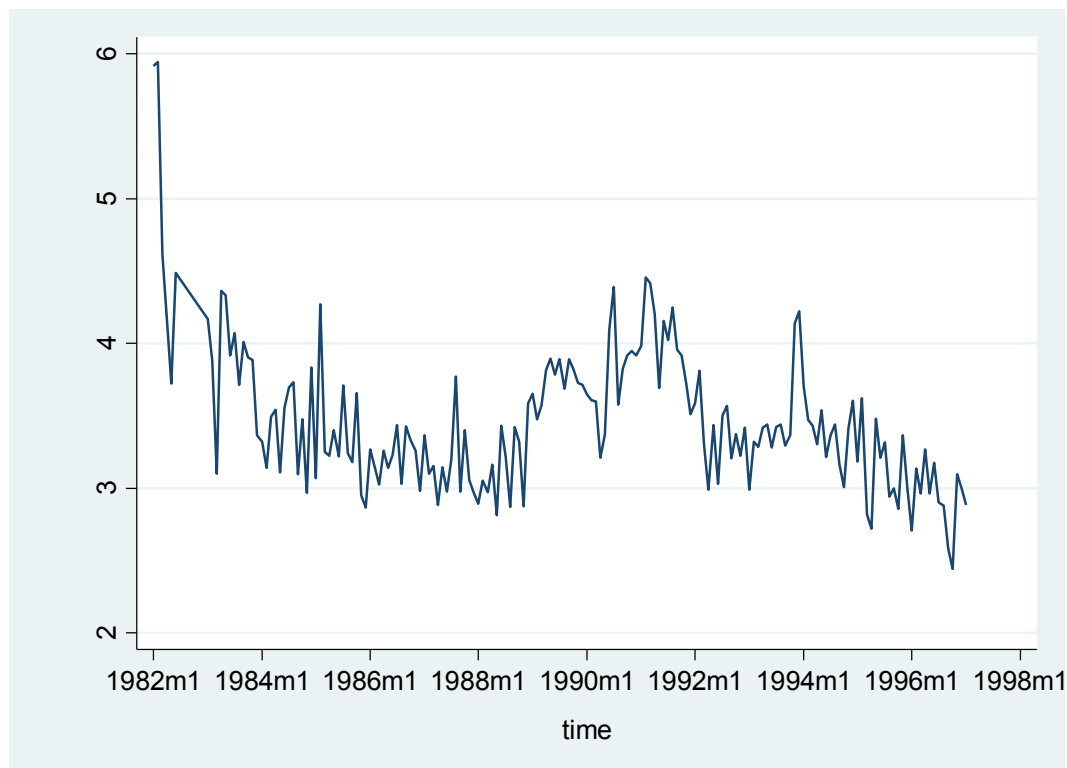
¹³⁶ Van Lelyveld (2000), page 61.

6.4 Analysing disagreement in UK quantitative Gallup data

Our aim in this section will be to analyse the impact of a few intuitively relevant variables highlighted by MRW (2003) on the level of disagreement about inflation expectations in the UK.

We follow MRW in choosing the interquantile range between the 75th and 25th percentiles as our measure of disagreement. We derived this series from the Gallup distribution and obtained the results plotted in Figure 115. Note that we have data for only five months of the year 1982, the remaining data points appearing through interpolation in the figure below¹³⁷.

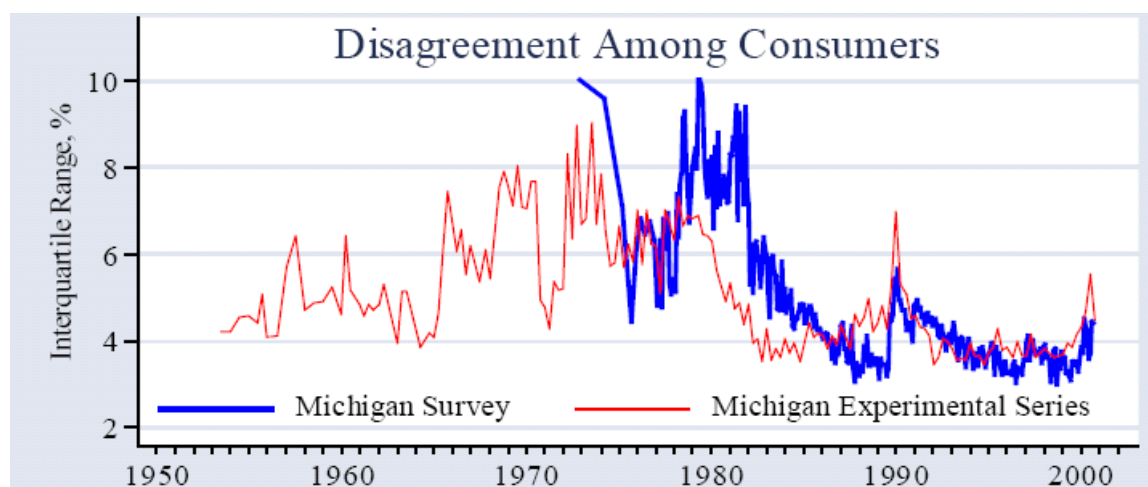
FIGURE 115: GALLUP EXPECTATIONS DATA, INTERQUANTILE RANGE



¹³⁷ The data are for January, February, March, May and June.

The most striking feature of this plot is how little the interquantile range of expectations appears to have varied over the sample period, the initial sharp drop excepted. The series seems to have persistently floated in the range 3-4%, with perhaps a permanent decrease at the end of the period. Figure 116 plots the interquantile range of the Michigan Survey data used by MRW for comparison. The US data appear to display considerably more variability, although most of it is focused early in the sample. The reason for this difference is not immediately straightforward; it may be that the regional inflation rates in the US are more dispersed than those applicable in the UK, leading to a greater implied disagreement in the US on account of this influence.

FIGURE 116: US, MICHIGAN SURVEY, INTERQUANTILE RANGE



SOURCE: Mankiw, Reis and Wolfers (2003), p.10.

When examining the Gallup interquantile range data it is possible to recognise the visual reflection of a number of important periods in the history of UK inflation: the disinflation of the early 1980s, the Lawson boom, the ERM disinflation and the 'quiet period' of the 1990s associated with Britain's inflation targeting regime.

This would immediately suggest the level of inflation as a potential driver of the extent of disagreement about inflation expectations among consumers. Other variables of interest would be the change in yearly inflation ($\pi_{t,t-12} - \pi_{t-12,t-24}$), the cyclical state of the economy (as proxied by the output gap) and, finally, the dispersion of inflation rates across different HICP sub-components (as proxied by the interquartile range for the distribution of inflation across twelve HICP subcategories). The relationships of these variables with disagreement about inflation expectations are explored graphically in detail in Appendix II, evidencing weak positive relationships.

A generalised finding when considering the previous bivariate relationships is that they are considerably less strong than those found to be applicable to the US by MRW (2003). This may be more reflective of a difference in the sample period than of any fundamental underlying difference in economic behaviour. Indeed, a careful consideration of the analogous plots in MRW reveals that data points corresponding to the 1970s seem to exert a crucial influence in steepening the slope of the bivariate regressions used to produce the best-fitting lines displayed. Those years are unfortunately unavailable in our sample.

We now turn to a multivariate consideration of the factors examined above in Table 74, in a regression seeking to explain the level of disagreement, as measured by the interquartile range of the Gallup expectations distribution.

The first set of results simply regresses the Gallup interquantile range on a constant, inflation, the squared change in yearly inflation, the output gap and the HICP interquantile range. As can be seen, the regression exhibits significant autocorrelation, autoregressive conditional heteroscedasticity and deviations for normality in the residuals. It is however possible to obtain a congruent specification by including an AR(1) process for the dependent variable, as in specification 2. Specification 3 drops all insignificant variables.

TABLE 74: REGRESSIONS OF GALLUP IQR, 1991m1-1997m1

	(1)		(2)		(3)	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
GLPIQR_1	-	-	0.492 ***	0.126	0.610 ***	0.093
Constant	2.624 ***	0.114	1.375 ***	0.337	1.113 ***	0.283
INFLA	0.064 *	0.033	0.053 *	0.030	0.061 **	0.027
DLTINFLA2	0.007	0.005	0.003	0.005	-	-
OUTGAP	-0.027	0.043	-0.027	0.039	-	-
GBRWIQR	0.171 ***	0.054	0.044	0.059	-	-
Diagnostics and Misspecification tests						
	Statistic	P-value	Statistic	P-value	Statistic	P-value
R2	0.48	-	0.58	-	0.56	-
Sigma	0.321	-	0.291	-	0.291	-
AR 1-5	3.6247	[0.0060]**	0.53271	[0.7507]	0.95998	[0.4489]
ARCH 1-5	3.38	[0.0095]**	0.97594	[0.4404]	0.59561	[0.7034]
Normality	6.0804	[0.0478]*	1.5702	[0.4561]	0.80557	[0.6685]
Hetero	0.75522	[0.6430]	0.36527	[0.9565]	0.10312	[0.9810]
Hetero-X	0.82202	[0.6426]	0.4951	[0.9551]	0.30454	[0.9084]
RESET	3.6989	[0.0587]	1.3453	[0.2503]	1.7276	[0.1931]

NOTE: *,**,*** respectively denote significance at the 10%, 5% and 1% levels. P-values provided in brackets [].

As can be seen from the set of results, only the inflation variable appears to be significant when the model is modified for congruence, with a coefficient of the expected sign but somewhat weak economically.

Given the lack of compelling evidence for the importance of the interquantile range of HICP inflation, we considered regressions which drop this variable, as this allows us to use a much larger sample starting in January 1983. Table 75 reports the results.

This time it was found that an AR(2) process was needed to address findings of autocorrelation in the residuals fully (specification 1). Despite the introduction of dummies to control for large residuals (about 3 standard errors in size), it appears that the model is unable to fully eliminate heteroscedasticity (specification 2) – possibly due to the limited information set or dynamic specification adopted. The parsimonious model (specification 3) is nevertheless very clear in attributing highly significant roles to both the level of inflation and its squared change in accounting for dispersion.

TABLE 75: REGRESSIONS OF GALLUP IQR, 1983m1-1997m1

	(1)		(2)		(3)	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
GLPIQR_1	0.278 ***	0.077	0.291 ***	0.072	0.305 ***	0.071
GLPIQR_2	0.226 ***	0.076	0.208 ***	0.071	0.222 ***	0.070
Constant	1.420 ***	0.257	1.408 ***	0.238	1.353 ***	0.234
INFLA	0.051 ***	0.015	0.053 ***	0.014	0.044 ***	0.012
DLTINFLA2	0.008 **	0.003	0.011 ***	0.003	0.011 ***	0.003
OUTGAP	-0.028	0.021	-0.024	0.019	-	-
I1993:11	-	-	0.942 ***	0.282	0.937 ***	0.283
I1985:2	-	-	0.879 ***	0.285	0.900 ***	0.285
I1983:3	-	-	-0.876 ***	0.289	-0.884 ***	0.290
Diagnostics and Misspecification tests						
	Statistic	P-value	Statistic	P-value	Statistic	P-value
R ²	0.47	-	0.55	-	0.55	-
Sigma	0.301	-	0.279	-	0.279	-
AR 1-5	1.5013	[0.1707]	1.3403	[0.2351]	1.3155	[0.2465]
ARCH 1-5	2.4971	[0.0188]*	1.0359	[0.4087]	1.2131	[0.2992]
Normality	3.6877	[0.1582]	1.5153	[0.4688]	1.9984	[0.3682]
Hetero	0.95213	[0.4877]	1.5042	[0.1224]	1.6342	[0.0947]
Hetero-X	1.8812	[0.0181]*	2.4585	[0.0007]**	2.9752	[0.0002]**
RESET	0.053044	[0.8181]	0.0087412	[0.9256]	0.023532	[0.8783]

NOTE: *, **, *** respectively denote significance at the 10%, 5% and 1% levels. P-values provided in brackets [].

Thus, although our bivariate plots earlier suggested possible relationships between disagreement in expectations and each of four variables – the inflation rate, its change, the output gap and the interquantile range of HICP inflation – a multivariate analysis was only able to provide support for the first two factors. This both contrasts with and reinforces the findings in MRW (2003). Indeed, for their Michigan “experimental series” (starting in 1947 with a quarterly frequency and becoming monthly in 1978) only the inflation rate itself appears to be significant; however, when the authors consider the monthly Michigan data exclusively, all variables are found to be strongly significant. The authors, however, do not provide evidence on the quality of their specification by publishing the relevant diagnostics; it can only be gathered that the authors found it necessary to implement Newey-West corrections in their regressions.

6.5 Is the Carlson-Parkin interquantile range a good measure of disagreement?

Given the vastly greater availability of qualitative survey data on expectations in Europe relative to the quantitative evidence available, it is apparent that an important question concerns the ability of dispersion measures derived from our quantification methods to replicate the ‘true’ underlying disagreement in the data. This question is all the more interesting as Batchelor (2006) has recently suggested that the CP quantification methodology may produce relatively robust results in what regards the *mean* of the distribution, but is to be less trusted in what regards the dispersion of the distribution, as measured by its standard deviation.

It is straightforward to derive the standard deviation of the Carlson-Parkin distribution from equations (1) to (4) ¹³⁸. This is a function of the perceived rate of inflation and of the survey response proportions (through the standardised points under the density function denoted by Zs):

$$\sigma_t^{\text{exp}} = \Pi_{t-12}^p * \left[\frac{2}{Z_t^1 + Z_t^2 - Z_t^3 - Z_t^4} \right] \quad (83)$$

Now, given that in the standardised normal distribution $P(z \leq 0.75) \approx 0.67449$, the interquantile range between the 75th and 25th percentiles can be expressed as:

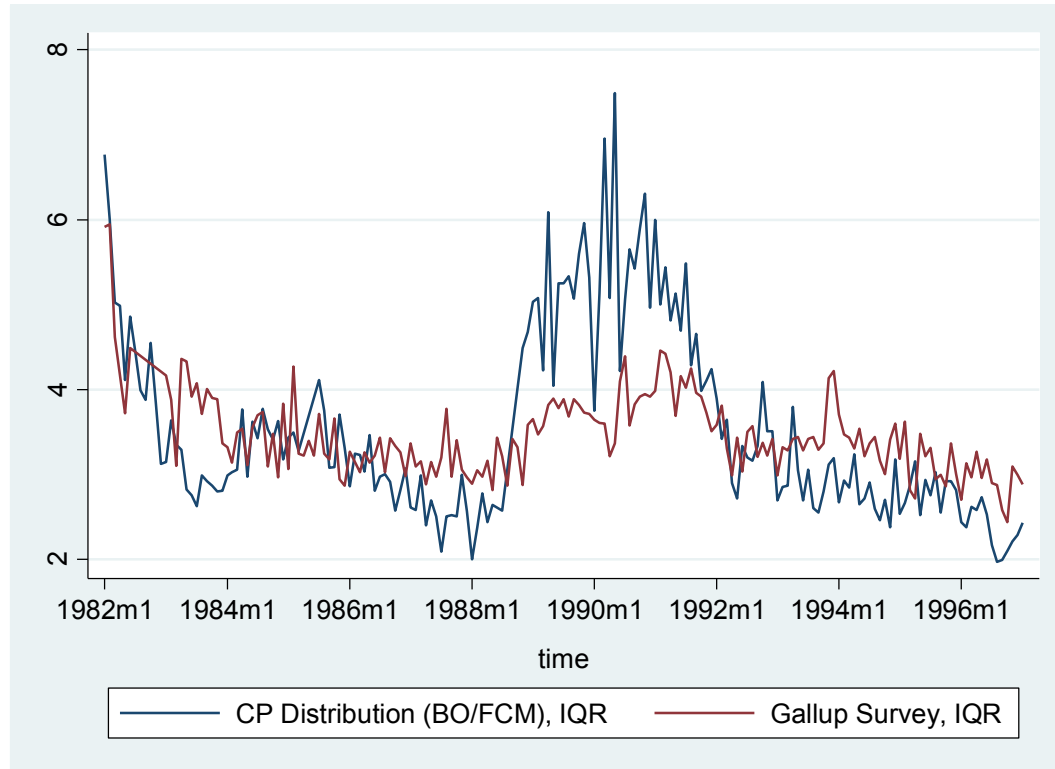
$$IQR_t^{\text{exp}} = 2 * 0.67449 * \sigma_t^{\text{exp}} \quad (84)$$

We take the interquantile range of the Gallup distribution as an approximation for the actual interquantile range of expectations in the population. Figure 123 plots both series for a visual comparison.

It appears graphically that the interquantile range from the CP approach provides a very reasonable guide to that from the Gallup distribution, both in terms of levels and overall time variation. The CP series displays greater variability, however, especially during the volatile boom-disinflation period of the late 1980s and early 1990s.

¹³⁸ Please refer to Section 3.1 of Chapter 1.

FIGURE 123: CP AND GALLUP INTERQUANTILE RANGES, UK



We provide some regression evidence in Table 76 to further explore the relationship between both variables. A simple regression without constant (specification 1) confirms the close relationship and similarity in level terms observed in the graph by a coefficient close to one. However, as is obvious from the specification diagnostics, this is very poor in modelling terms. An ADL(2,2) specification is found to be required to take into account the persistence in the data and thereby remove residual autocorrelation (specification 2). Specification 3 aims at parsimony and introduces indicator variables to control for particularly large residuals (around 3 standard errors in size), despite which some evidence of heteroscedasticity remains. Note that the dummies are (almost) the same as those already found in Section 6.4 above. It appears from this last specification

that the interquantile range from the CP method (BISIQR) is a highly significant explanatory variable, and enters lagged by two months. Thus, it is as if the CP data ‘anticipate’ movements in the Gallup data. This may be due to the sampling particularities of both surveys.

TABLE 76: REGRESSIONS OF GALLUP IQR ON CP IQR, 1983m1-1997m1

	(1)		(2)		(3)	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
GLPIQR_1	-	-	0.265 ***	0.076	0.323 ***	0.070
GLPIQR_2	-	-	0.231 ***	0.073	0.179 ***	0.068
Constant	-	-	1.272 ***	0.230	1.245 ***	0.210
BISIQR	0.924 ***	0.019	-0.021	0.045	-	-
BISIQR_1	-	-	0.054	0.043	-	-
BISIQR_2	-	-	0.096 **	0.046	0.125 ***	0.025
I1993:11	-	-	-	-	0.895 ***	0.273
I1985:2	-	-	-	-	0.951 ***	0.276
I1983:4	-	-	-	-	0.963 ***	0.276
Diagnostics and Misspecification tests						
	Statistic	P-value	Statistic	P-value	Statistic	P-value
R2	-	-	0.48	-	0.57	-
Sigma	0.884	-	0.297	-	0.272	-
AR 1-5	34.558	[0.0000]**	0.79934	[0.5890]	1.6817	[0.1173]
ARCH 1-5	7.771	[0.0000]**	1.4467	[0.1910]	1.094	[0.3701]
Normality	54.051	[0.0000]**	7.41	[0.0246]*	2.6012	[0.2724]
Hetero	261.37	[0.0000]**	0.91779	[0.5186]	1.5151	[0.1473]
Hetero-X	261.37	[0.0000]**	1.9145	[0.0156]*	2.9819	[0.0009]**
RESET	784.3	[0.0000]**	0.084663	[0.7715]	0.82795	[0.3642]

NOTES: *, **, *** denote respectively significance at the 10%, 5% and 1% levels.

Thus, the results here suggest that measures of the dispersion of expectations derived from qualitative surveys do appear to be informative to some extent, when the benchmark considered is the analogous degree of dispersion derived from quantitative data. These results contrast with those obtained by Batchelor (2006) in the US context and provide grounds for somewhat greater optimism on the issue.

6.6 Qualitative analysis of the expectations distribution

This section aims to give careful consideration to the structure of the distribution of inflation expectations in order to contribute to a few outstanding debates.

Section 6.6.1 considers the issue of whether it is more appropriate to use the mean or the median as a measure of inflation expectations derived from quantitative surveys. This issue has attracted particular attention in the context of the US, given the greater availability of such data for consumers there.

Section 6.6.2 introduces a new survey to be used in what follows, namely the Bank of England / NOP Inflation Attitudes Survey. To our knowledge, this is the first study to be making use of this recently developed survey for purposes of academic research.

Sections 6.6.3 and 6.6.4 provide new evidence on the extent to which consumers seem to be forming their expectations *qualitatively* as opposed to *quantitatively*. This involves stepping into a debate between Curtin (2006) and Bryan and Palmqvist (2006), the latter authors arguing that there are good indications that expectations are formed *qualitatively* to an important extent, and the former disagreeing. Section 6.6.4 will also consider the influence of policy changes upon the distribution of inflation expectations.

Finally, Section 6.6.5 brings together some of the insights gained earlier to provide observations on whether it is quantitative or qualitative surveys that are on the whole to be preferred.

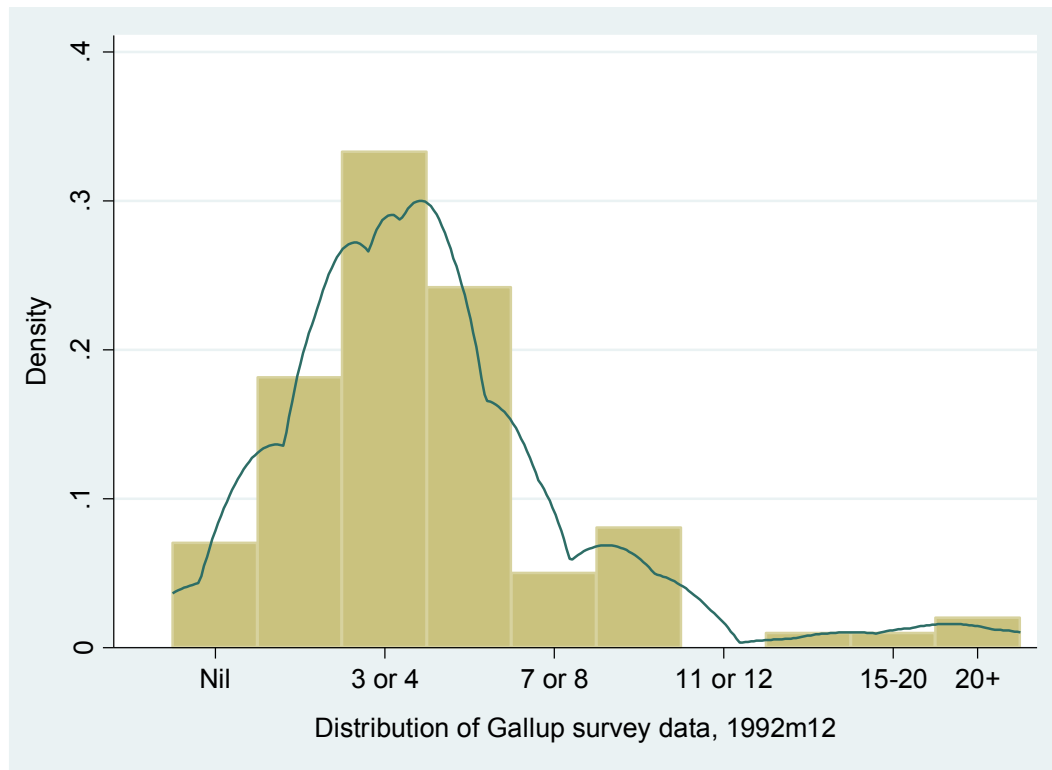
6.6.1 A note on the choice between the mean and median of the distribution

Most authors working on US data sourced from the Michigan Survey choose to use the *median* of the distribution rather than the *mean* as their preferred measure of inflation expectations. This is due to the finding across a large number of studies of a positive bias for the mean measure when comparing it to an aggregate price index, as noted by Bryan and Palmqvist (2006). For instance, Mehra (2002) reports that the median inflation forecasts of Michigan households are essentially unbiased over the period 1980q3-2000q3, unlike the mean forecasts which overestimate inflation. A possible culprit for this appears when considering the distribution of expectations in detail. Indeed, Curtin (2006) notes that the distribution of the Michigan data presents a long upper tail which is “*likely to represent measurement errors*” (see also Curtin, 1996); he further mentions the “*long standing debate*” as to whether these observations should be treated as missing data or considered to indicate a “*large*” expectation and truncated. Note that the resolution of this debate also has an impact for appropriate indicators of second moments – i.e. a ‘pessimistic’ view of these observations would make the interquantile range more appropriate than a standard deviation-based measure (Mankiw, Reis and Wolfers, 2003).

As has already been visible in some earlier plots, similarly “large” observations occur in the Gallup quantitative expectations data. Indeed, some respondents have reported expectations of over 20% when the maximum inflation rate attained over the Gallup sample period is 11.2% (in January 1982). Figure 124 provides an example of this phenomenon for December 1992. Whereas the underlying CPI inflation at the time was

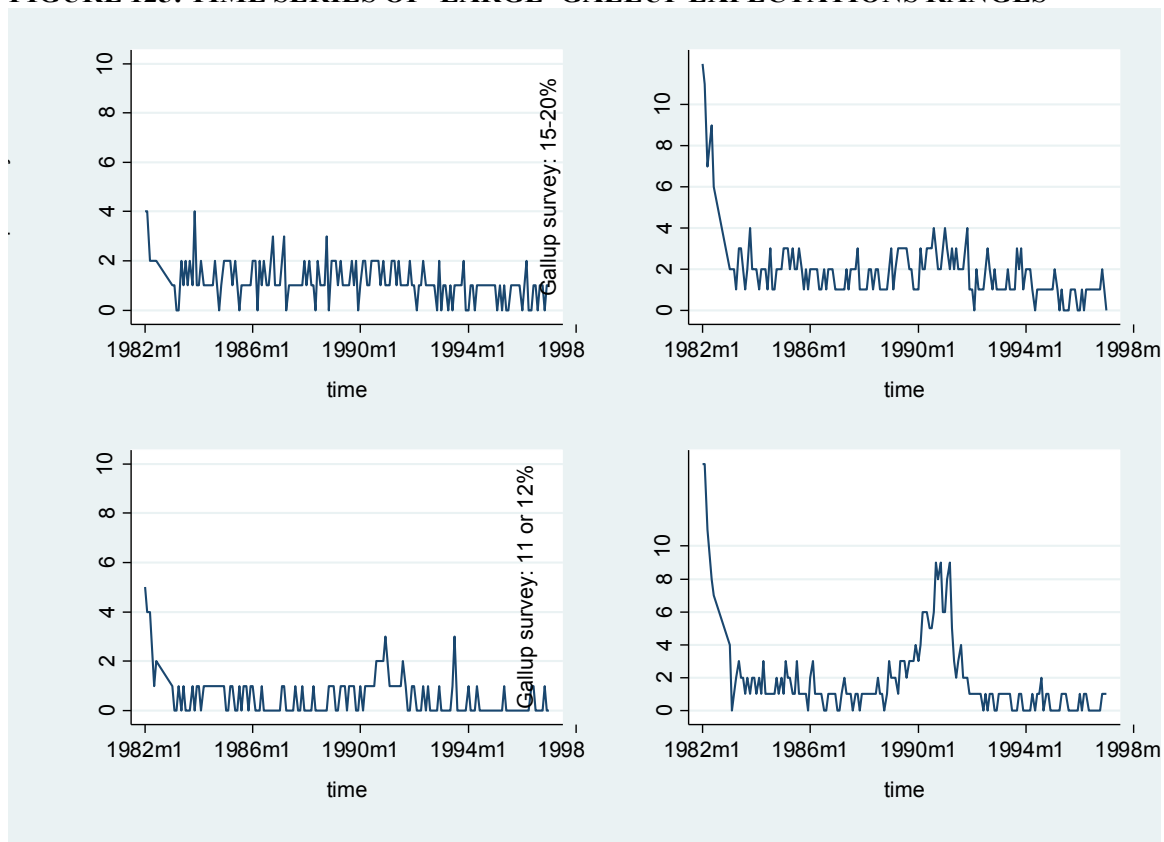
2.6%, a small proportion of respondents nonetheless reported an expectation of inflation over 20%.

FIGURE 124: GALLUP EXPECTATIONS DISTRIBUTION, 1992m12



Is this observation likely to be of consequence? To explore the question, we plot the time series of the proportion of respondents falling in the “large” bins of the distribution reported by Gallup in Figure 125.

FIGURE 125: TIME SERIES OF ‘LARGE’ GALLUP EXPECTATIONS RANGES



As can be seen, with the exception of the “11 or 12%” category, these large expectations attract a small (typically under 2%) and – crucially – almost constant share of the ‘vote’. The variation of the “11 or 12%” category is quite reasonable, as it becomes significant precisely when underlying inflationary behaviour justifies it (the early 1980s and the Lawson boom period). The relative constancy of these categories means that even in the extreme case that they represent pure noise, their impact on the mean would be mostly limited to a *level* difference relative to the median and both series would exhibit the same time series behaviour. This is in fact the finding highlighted by Curtin (2006) for the Michigan data: over the period 1978m1-2005m8 the mean was about one percentage point higher than the median, but both series displayed a time series correlation of 98%.

The extent to which these observations reflect noise rather than genuine economic behaviour is debatable. The relative constancy of response proportions in these categories despite large changes in inflation across the sample would seem to suggest that these data points are not particularly informative. At the same time, it is not inconceivable that the private consumption bundle of an individual is hit by a level of inflation well above that of the headline index (e.g. due to particular healthcare costs). It is difficult to make a more precise statement in the absence of individual-level data. In any event, this does not seem to be of consequence for our purposes.

6.6.2 Introducing the Bank of England / NOP Survey

In November 1999, the Bank of England initiated a survey of public attitudes to inflation in association with market research firm NOP, which included it in its regular Omnibus surveys. The survey concerns perceptions and expectations of inflation and interest rates, as well as the public's general understanding of the monetary policy framework in the UK. The frequency of the surveys is quarterly (February, May, August and November); some additional questions on dimensions unlikely to experience short-run shifts in opinion are asked on a yearly basis in the February survey.

The surveys are carried using a random location sample designed to be representative of all adults in Great Britain, and interviewing is carried out "face to face" in homes. The sample size of the quarterly surveys is about 2,000, whilst it is almost double that for the annual surveys in February. The raw data are weighted to match the demographic profile of Great Britain as a whole (see Bank of England, 2001-2005 for details).

The questions of most direct interest for our purposes are questions 1 and 2 of the survey, concerning inflation perceptions and expectations, respectively. These are listed in Table 77.

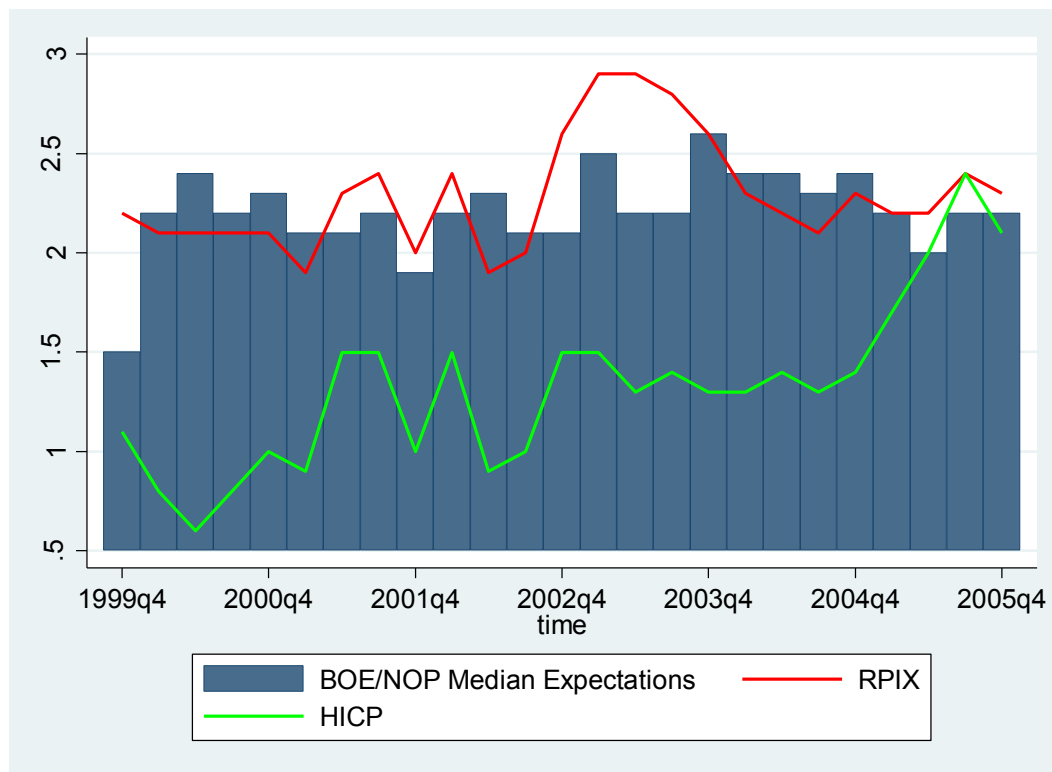
TABLE 77: BANK OF ENGLAND / NOP SURVEY, QUESTIONS 1 AND 2

Q.1 Which of these options best describes how prices have changed over the last 12 months?	Q.2 How much would you expect prices in the shops generally to change over the next 12 months?
Gone down	Go down
Not changed	Not change
Up by 1% or less	Up by 1% or less
Up by 1% but less than 2%	Up by 1% but less than 2%
Up by 2% but less than 3%	Up by 2% but less than 3%
Up by 3% but less than 4%	Up by 3% but less than 4%
Up by 4% but less than 5%	Up by 4% but less than 5%
Up by 5% or more	Up by 5% or more
No idea	No idea

The BOE/NOP survey thus offers respondents a set of *ranges* for inflation to choose from. The median is then computed on the basis of this information, assuming that responses are uniformly distributed within bands. In this respect, the survey is quite different from its Gallup counterpart, which used an open-ended question but reported distributional information in “range format”.

Figure 126 plots the median expectation *at the time the survey was conducted* against the actual inflation rate at that time, as measured alternatively by the RPIX (the retail price index excluding mortgage interest payments) and the HICP (harmonised index of consumer prices). The data were obtained from the Office of National Statistics (ONS).

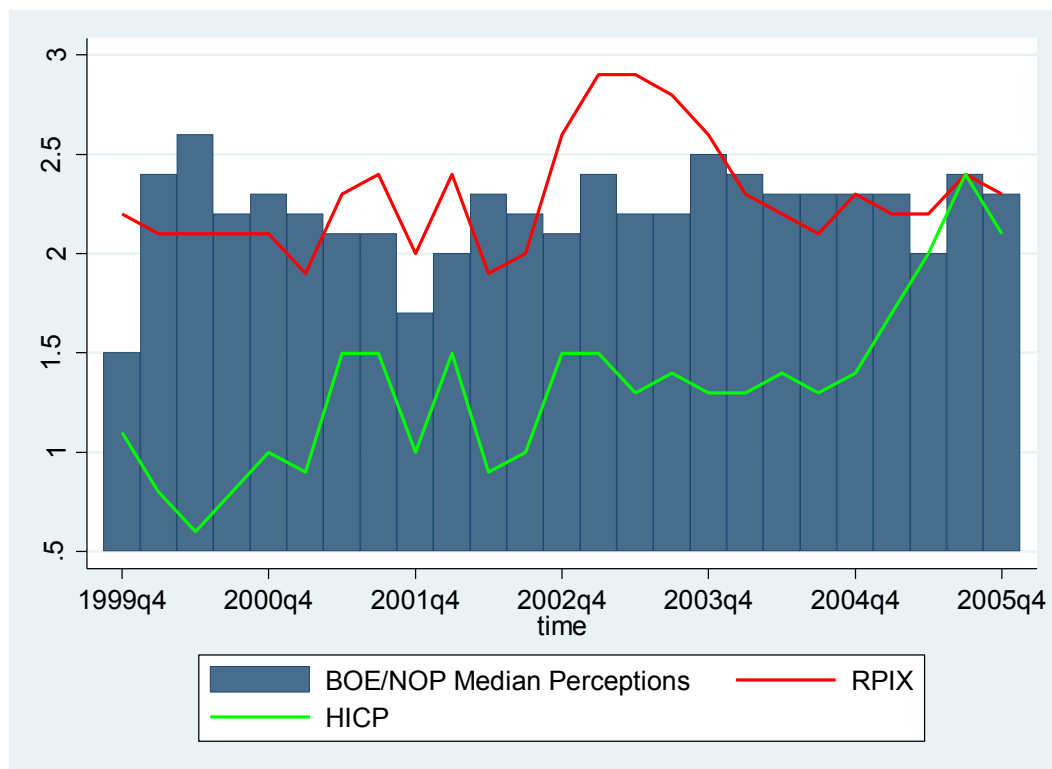
FIGURE 126: BOE/NOP EXPECTATIONS AT TIME SURVEYED VS PRICE INDICES



The interest of considering both of these measures of inflation is that they respectively represent the former and present measures used in the UK's inflation targeting framework. The impact of this framework on the distribution of inflation expectations is considered further below (Section 6.6.4). For now, we simply note that the median expectation has been much closer to the RPIX than to the HICP in the sample made available by these surveys.

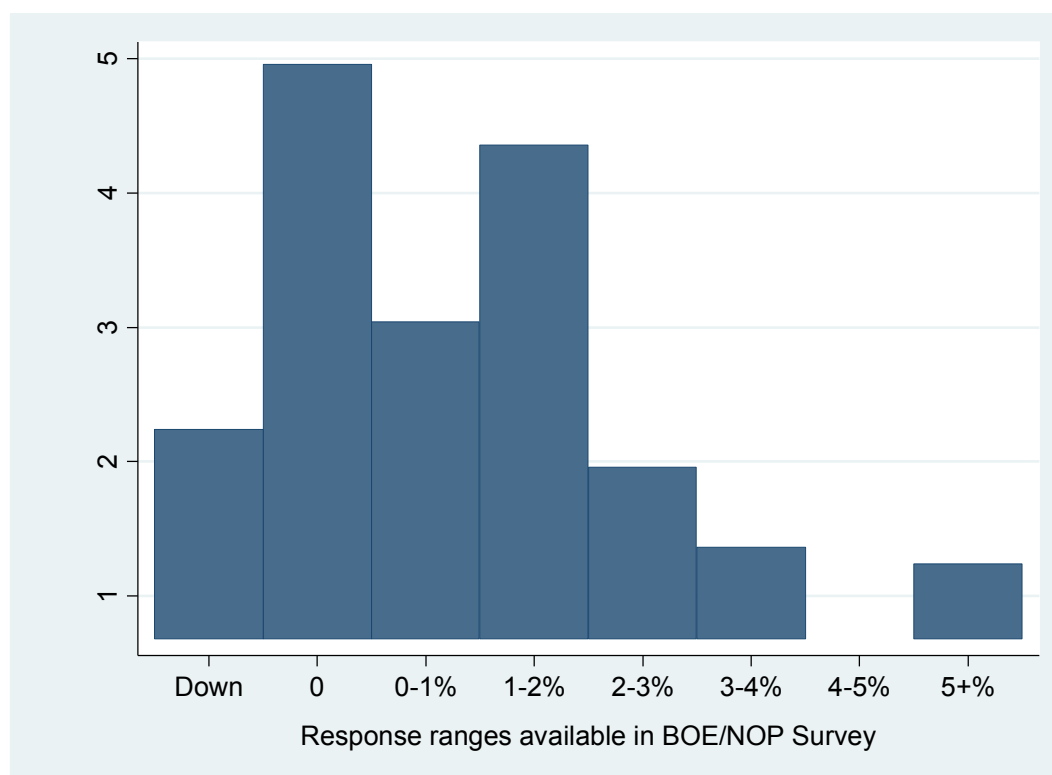
Figure 127 reproduces this exercise in the case of the BOE/NOP median *perceptions*. As can be seen, the plot is extremely similar to that obtained for expectations, and the same broad observations hold.

FIGURE 127: BOE/NOP MEDIAN PERCEPTIONS VS PRICE INDICES



The similarity between both plots is not a surprise. Bank of England (2005) already noted that both median series moved closely together, and that indeed the *distributions* underlying both questions were themselves similar. Figure 128 explores this by plotting for each survey range, the absolute difference between the proportion of respondents expressing an *expectation* in that range and the proportion expressing a *perception* in that range, averaged over the full sample 1999q4-2005q4.

FIGURE 128: AVG. ABSOLUTE DIFFERENCE, EXPECTATIONS - PERCEPTIONS



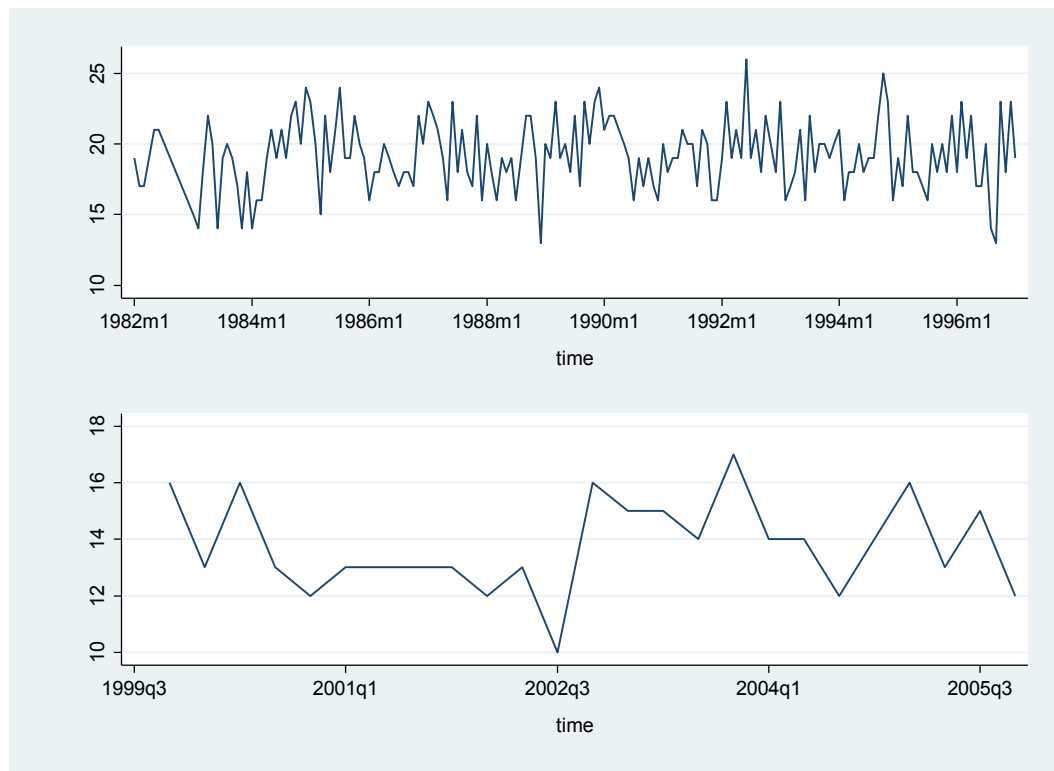
NOTE : only actual answers were considered when computing these percentages; in other words, 'don't know' responses were excluded and actual answers reweighted to sum to 100.

As can be seen, the difference between perceptions and expectations in a common range is very small on average – 2% points of difference or less for most options. The categories in the range 0-2% constitute an exception by exhibiting slightly larger differences. Further comments will be made on this matter below (Section 6.6.4). For now, suffice it to say that this finding is consistent with that already highlighted for the EC Consumer Survey data (Section 2.2 in this chapter) and suggests a rather adaptive behaviour on the part of consumers – who seem to project perceptions into the future when forming expectations, to a large extent.

6.6.3 Studying the pattern of “don’t know” responses across surveys

The literature has so far paid no attention to the significance of the proportion of respondents choosing “don’t know” answers when asked for their inflation expectations. *Prima facie*, this lack of attention is difficult to justify given the “popularity” of this response over the last twenty years of data, as recorded by *both* the Gallup and the Bank of England / NOP quantitative surveys and displayed in Figure 129.

FIGURE 129: ‘DON’T KNOW’ ANSWERS, GALLUP AND BOE/NOP SURVEYS

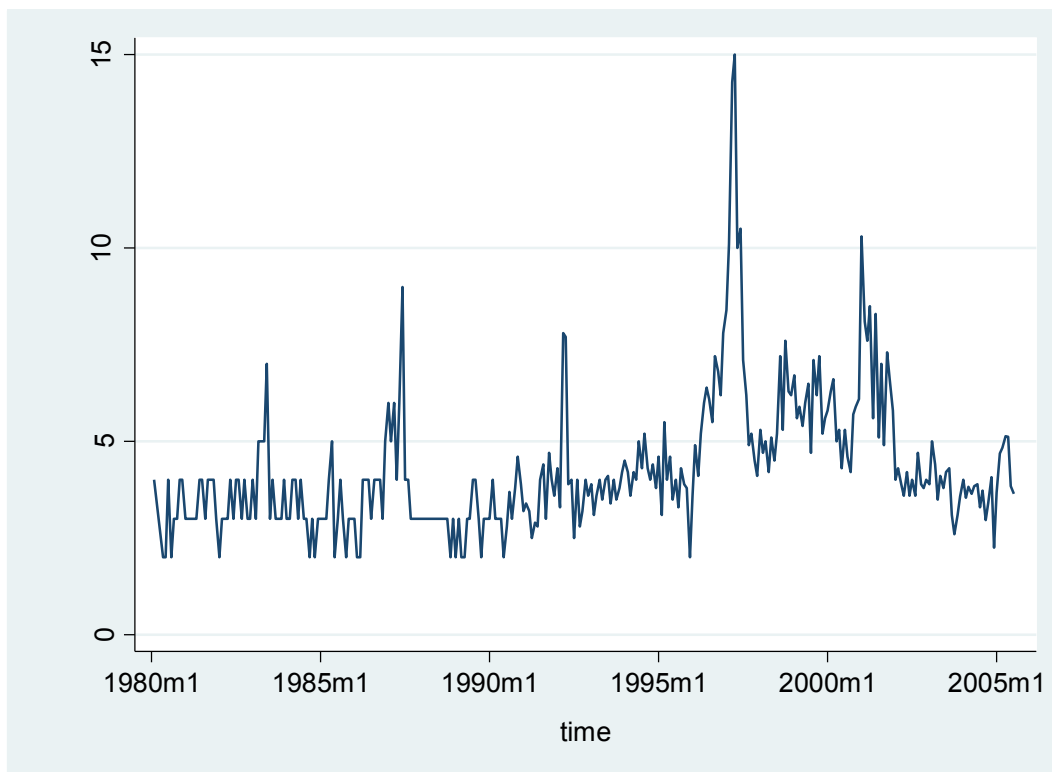


As can be seen, the proportion of “don’t know” responses is very significant in both surveys – with a minimum of 10% in the BOE/NOP survey and a maximum of over 25% in the Gallup survey. The mean value of the proportion of respondents choosing this

category is 19.2% for the Gallup survey and 13.8% for the BOE/NOP survey over their respective full samples.

The real surprise however comes when contrasting these time series with the EC Consumer Survey's "don't know" category. The latter is plotted in Figure 130.

FIGURE 130: 'DON'T KNOW' ANSWERS, EC CONSUMER SURVEY



Two things are immediately obvious when considering the plot:

1. There are considerably fewer 'don't know' responses to the expectations question of the EC Consumer Survey relative to those in the Gallup and BOE/NOP surveys

2. There are a series of important peaks in the data that require explanation

We now deal with both issues in turn.

Table 78 presents the means of “don’t know” response proportions for the full sample of each of the surveys. We also present the results for two sub-samples of the EC Consumer Survey in order to match as closely as possible the Gallup and BOE/NOP samples.

TABLE 78: ‘DON’T KNOW’ ANSWERS, BY SURVEY AND SAMPLE PERIOD

Survey	Sample	Mean of “Don’t Know” responses (%)
Gallup	1982m1-1997m1 ¹³⁹	19.2%
Bank of England / NOP	1999q4-2005q4	13.8%
EC Consumer Survey	1982m1-1997m1 ¹⁴⁰	3.79%
	1999m11-2005m7	4.78%
	1980m2-2005m7 ¹⁴¹	4.28%

As can be seen, ‘don’t know’ responses to the EC Consumer Survey represent less than 20% of the proportion collected in the Gallup survey and less than 35% of the proportion collected in the BOE/NOP survey over comparable samples. This is clearly a very significant difference.

Are these observations robust to the choice of country? Sweden is the only other country for which we have both quantitative and qualitative survey data on consumer expectations. However, we were only able to obtain data on the “don’t know / don’t want to give an answer” category of the SCB/GfK survey for the months of October 2005-January 2006. The proportion of responses in this category ranged from 6.7% to 9.9%

¹³⁹ With gaps: 1982m4, 1982m7-m12

¹⁴⁰ With gaps: 1982m4, 1982m7-m12

¹⁴¹ With gaps: 1980m3, 1980m4, 1981m10.

and averaged 8.7% over those four months. This is in stark contrast with the proportion of “don’t know” responses in the EC Consumer survey which averaged 0.88% in the year to July 2005 (our last data point). This represents only about 10% of the proportion collected in the quantitative survey, thereby confirming our results for the UK.

In light of this, we formulate the hypothesis that the general public (or at least a proportion thereof) forms expectations of a *qualitative* nature to a certain extent. Thus, the public may have a “rough view” of inflation – perhaps thoughts of the type:

- “*I expect inflation to be in the range 0-4%*”; or
- “*Inflation over the coming year should be roughly the same as it is now*”, without necessarily having a very clear *numerical* idea of what the inflation numbers currently are.

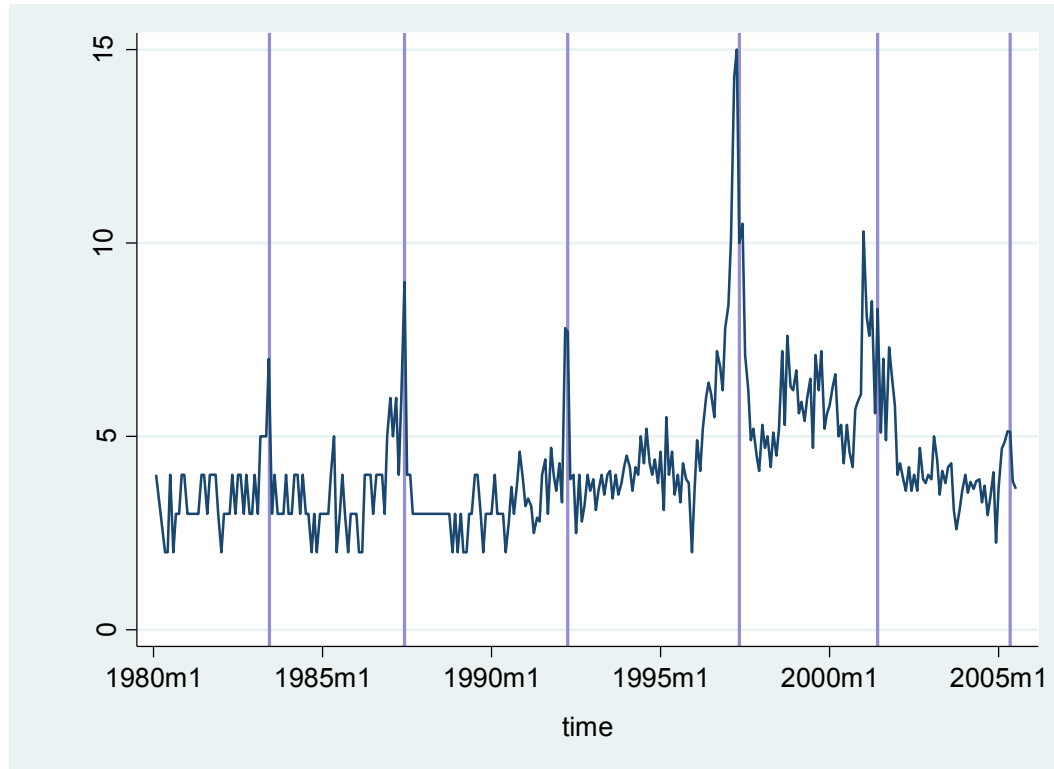
Such individuals may be able to distinguish the rough ‘steps’ of an ‘inflation staircase’ (e.g. low, medium, high or very high inflation) and associate certain numbers or ranges of numbers with these concepts; indeed, they may even adapt their behaviour to the broad “regime” they perceive to be in. However, if a survey seeks to elicit information that is more “defined” or “precise” than the way it is cognitively conceptualised by these sections of the public, it may lead them to turn to the “don’t know” option as a form of escaping the question. This is not implausible for either the Gallup survey (which takes

an open-ended question approach) or the BOE/NOP survey (which provides rather narrow sub-ranges to choose from what is also a relatively narrow range overall).

Of course, the proportion of “don’t know” answers in surveys depends on the instructions given to pollsters regarding how much probing they should apply in an attempt to get a defined answer. The size of the phenomenon observed nonetheless suggests its robustness, as well as the fact that it was observed both for the UK and Sweden. Furthermore, Bryan and Palmqvist (2006) had already highlighted certain qualitative features of the Swedish expectations distribution, although they did not examine the proportion of “don’t know” responses. We will consider their arguments and this hypothesis further below.

We now turn to the six peaks that are so clearly visible in Figure 130 for the proportion of “don’t know” responses in the EC Consumer Survey. They correspond to the months of June 1983, May/June 1987, March/April 1992, January/June 1997, January/June 2001 and April/May 2005. In fact, these peaks identify clearly and precisely each of the six general election periods that the United Kingdom has known over the sample period, since 1980. Figure 131 makes this point by indicating the actual months in which each of these elections were held.

FIGURE 131: EC SURVEY ‘DON’T KNOW’ ANSWERS & UK GENERAL ELECTIONS



This is clear evidence that, in the mind of the public, inflation uncertainty is to some extent linked to the political cycle. It is intuitive that the greatest uncertainty was faced when Labour won the 1997 election – a switch from the series of Conservative governments that had held power since 1979. Accordingly, this is associated with the largest increase in the “don’t know” response proportion by far. Table 79 presents the data in greater detail, highlighting the months in which general elections were held.

TABLE 79: GENERAL ELECTION EFFECTS, ‘DON’T KNOW’ RESPONSES

Election	Time	% Don’t Know	Election	Time	% Don’t Know
9th June 1983	1982m11	3	1st May 1997	1996m11	6.2
	1982m12	3		1996m12	7.8
	1983m1	4		1997m1	8.4
	1983m2	3		1997m2	10.1
	1983m3	5		1997m3	14.3
	1983m4	5		1997m4	15
	1983m5	5		1997m5	10
	1983m6	7		1997m6	10.5
	1983m7	3		1997m7	7.1
	1983m8	4		1997m8	6.2
	1983m9	3		1997m9	4.9
11th June 1987	1987m1	6	7th June 2001	2001m1	10.3
	1987m2	5		2001m2	8.1
	1987m3	6		2001m3	7.6
	1987m4	4		2001m4	8.5
	1987m5	6		2001m5	5.6
	1987m6	9		2001m6	8.3
	1987m7	4		2001m7	5.1
	1987m8	4		2001m8	7
	1987m9	3		2001m9	4.9
9th April 1992	1992m1	4.3	5th May 2005	2005m1	3.68
	1992m2	3.3		2005m2	4.69
	1992m3	7.8		2005m3	4.84
	1992m4	7.7		2005m4	5.13
	1992m5	3.9		2005m5	5.12
	1992m6	4		2005m6	3.84

A number of hypotheses remain open as to what the public may be reading in the relationship between inflation and the possibility of change in the governing party. Does the associated increase in agnostic responses reflect hesitation as to what type of fiscal policy will be pursued? Or rather, does the public fear a change in monetary policy? It is also possible that the figures simply reflect ‘noise’ around the election rather than an actual thought process. This uncertainty makes it difficult to provide additional interpretation.

An example where such distinctions matter for interpretation arises when considering the last election. The peak in ‘don’t know’ responses associated with it was the weakest of this series of six. A proponent of the current monetary policy system of operational independence and inflation targeting may wish to claim that these institutional innovations have reduced the role of ‘politics’ in the determination of inflation, and that this is perceived by the public. However, the weak peak may also plainly be reflecting a low degree of uncertainty surrounding the last election.

The answer to such questions depends on the understanding of the current monetary framework by the public; this issue is considered in the next section.

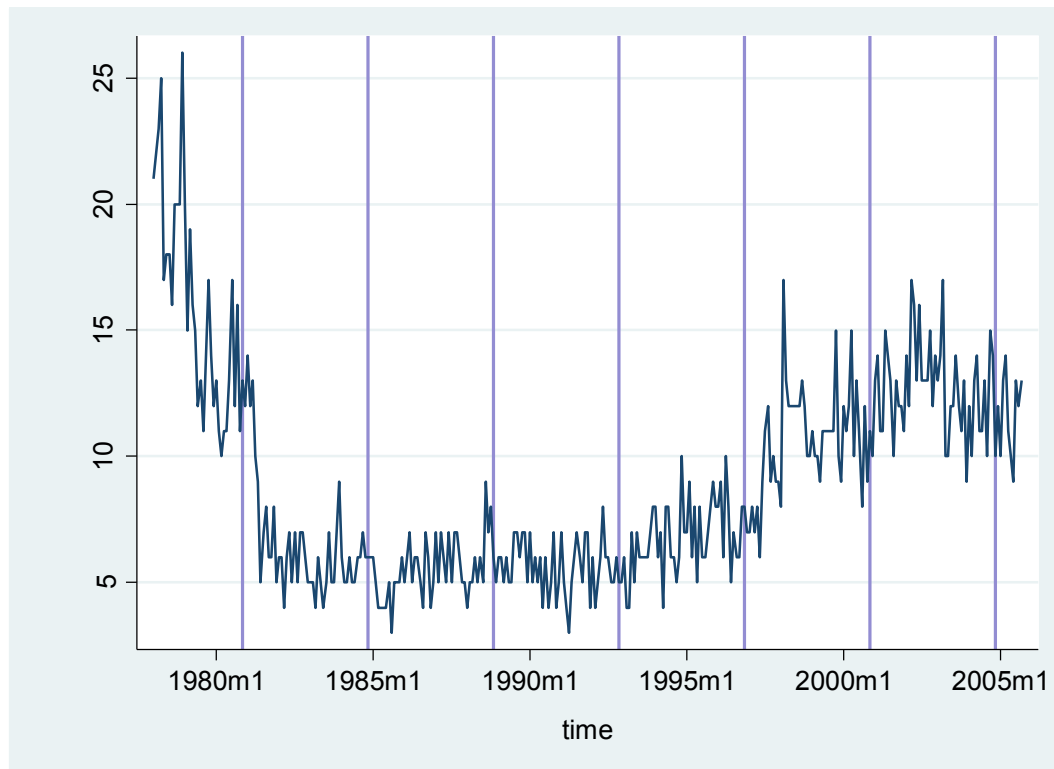
Before turning to other aspects, however, we propose to check whether such a relationship between the “don’t know” category and elections cannot also be found in the other surveys available to us. Figure 132 explores this in the case of the Gallup and Bank of England / NOP data; the dates of general elections are marked by vertical lines.

FIGURE 132: QUANTITATIVE ‘DON’T KNOW’ ANSWERS & UK ELECTIONS



The results are very different from those found using the EC Consumer Survey data for the UK, as no relationship between elections and the “don’t know” proportions can be found in these alternative British surveys. As a further check, we carry out the same exercise for quantitative data collected by the Michigan Survey Research Centre on a sample of US consumers. Figure 133 plots the relevant response proportion and highlights the months of US Presidential Elections.

FIGURE 133: US SRC “DON’T KNOW” ANSWERS & US PRESIDENTIAL ELECTIONS



Again, it does not appear possible to make a strong case on the basis of these data for the relationship between elections and uncertainty in inflation expectations, as reflected by the “don’t know” category. The data make two interesting points nonetheless.

Firstly, the “don’t know” answers collected in the Michigan Survey and presented in Figure 133 are really of two types: first, a “Don’t know, NA” category is present, translating the absence of any indication regarding future inflation and presumably complete uncertainty on the part of the respondent; second, the respondent can answer that he/she expects prices to go up, but is unable to tell by how much. This second category constitutes by far the overwhelming majority of don’t know answers. Indeed, over the period between August 1987 and January 2006, the first category has only

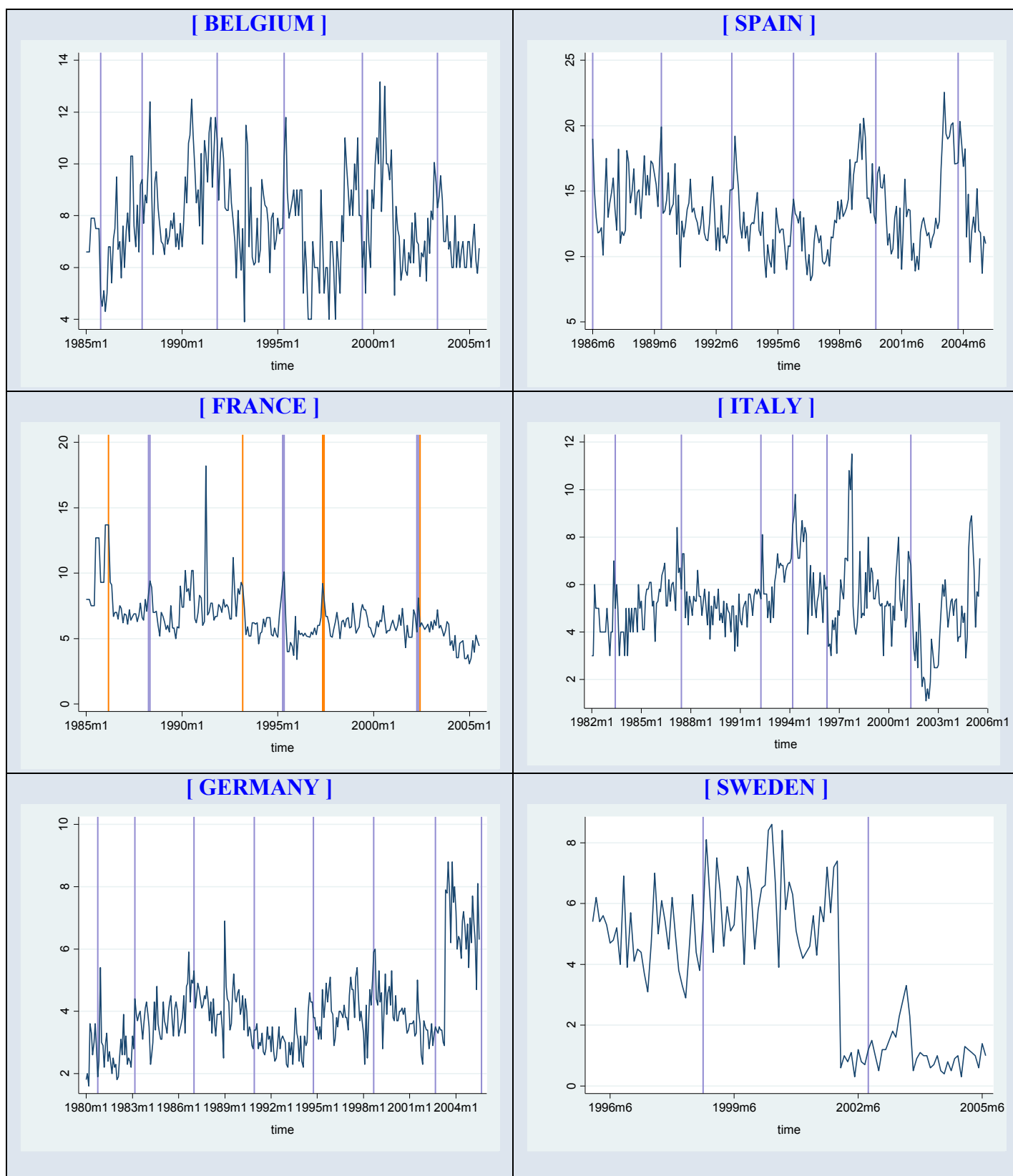
collected an average of 1% of survey responses, whereas the second has collected 7.8% on average. This second category of “don’t know” answers may well concern respondents who are able to express a *qualitative* view about inflation, and would quite possibly be able to give further qualitative precisions (such as “up by a lot” or “not very much”) but are uncomfortable with providing a specific number. Such an interpretation is consistent with the argument made above and the finding that qualitative surveys seem to have a much lower incidence of “don’t know” answers than their quantitative counterparts.

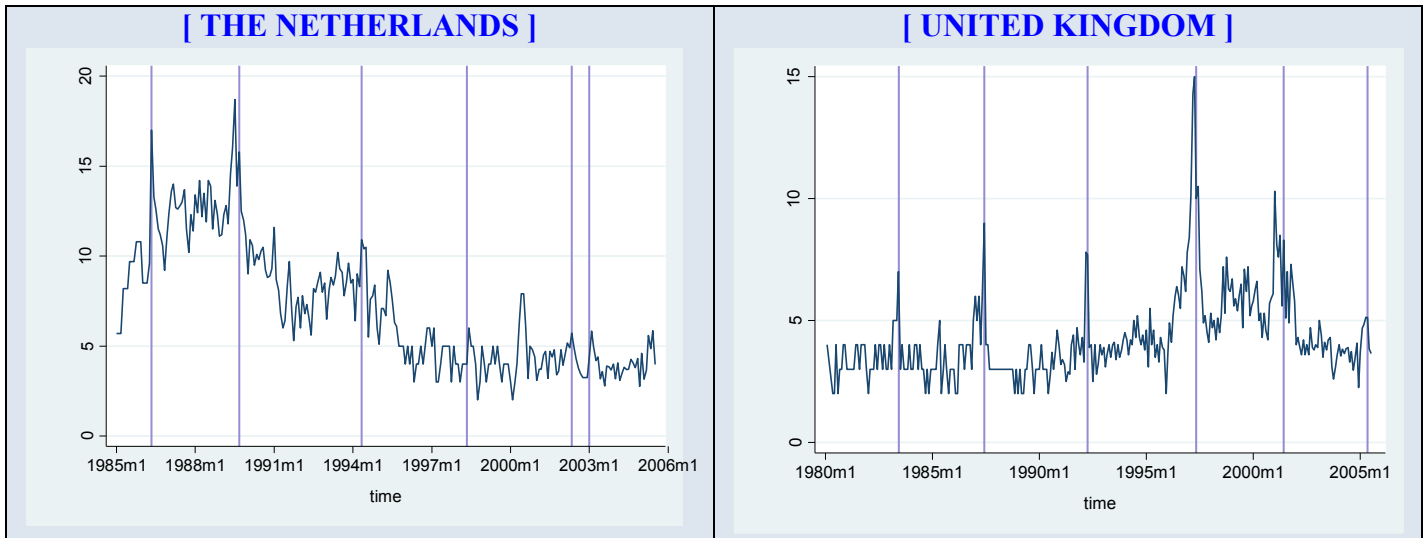
Secondly, Figure 133 shows that the proportion of “don’t know” responses collected by the Michigan Survey Research Centre appears to have increased steadily over the Greenspan era (August 1987 – January 2006). The data for the category “Up, Don’t Know By How Much” display nearly identical time series features as the aggregate presented in the figure. This may perhaps be a consequence of the period of relative price stability presided over by Alan Greenspan, and would seem to fit with his view that *“Price stability means that expected changes in the average price level... do not materially enter business and household financial decisions”*¹⁴².

Finally, we notice that the election effect has so far been invisible only on data arising from quantitative surveys. To check whether one would be justified in generalising this observation, Figure 134 considers the “don’t know” response proportions of the qualitative EC Consumer Survey data in all the other countries in our sample, marking with vertical lavender lines the months in which general elections were held.

¹⁴² Monetary Policy Report to Congress, February 21, 1989.

FIGURE 134: EC CONSUMER SURVEY, “DON’T KNOW” ANSWERS & ELECTIONS





NOTE: Vertical lavender lines mark the months in which general elections were held. France is an exception due to the peculiarities of its political system: lavender lines mark *Presidential* elections, whilst orange lines mark *legislative* elections.

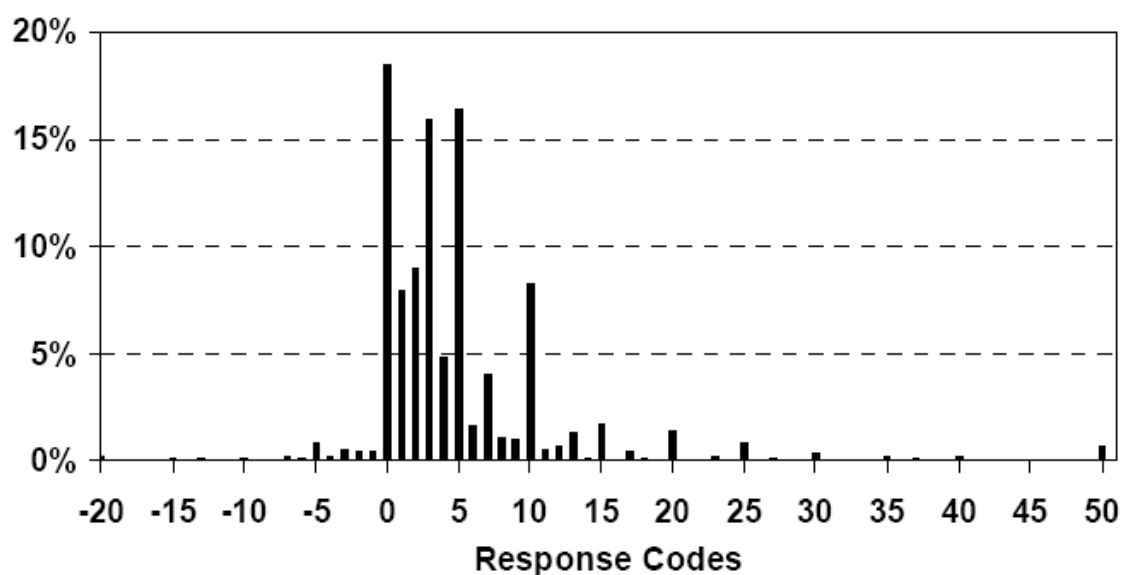
As can be seen, it is possible to find some association between elections and inflation uncertainty as measured by the proportion of “don’t know” responses for all of these countries (except perhaps on the limited Swedish sample). It is unclear why this finding is not replicated in the quantitative data samples we have considered. It could be related to a difference in the cognitive processes deployed when answering quantitative questions as opposed to the qualitative questions. It may be also due to a specificity of the EC qualitative question, which is expressed *relative* to the past 12 months, whereas the quantitative data employed here are elicited in *absolute* terms. The divergence could also be related to the way in which the Gallup and BOE/NOP data were collected (e.g. probing procedures) and to a perception in the US that the inflation process is largely out of the politicians’ hands. In any event, this constitutes a new and puzzling result. Its resolution would certainly benefit from the consideration of similar data for other countries.

6.6.4 “Zero barrier”, “digit preferences” and the impact of inflation targeting

A review of previous empirical work and definition of our hypotheses

In a recent paper, Curtin (2006) highlights two phenomena that can be observed in the US distribution of consumer inflation expectations, derived from the survey data collected by the University of Michigan’s Survey Research Centre (SRC). This distribution is computed from all the individual responses to the monthly surveys carried out from January 1978 to August 2005 and is reproduced as Figure 135, covering around 200,000 observations.

FIGURE 135: US, MICHIGAN SURVEY; EXPECTATIONS DISTRIBUTION, 1978-2004



Source: Curtin (2006), Chart 1, p. 25.

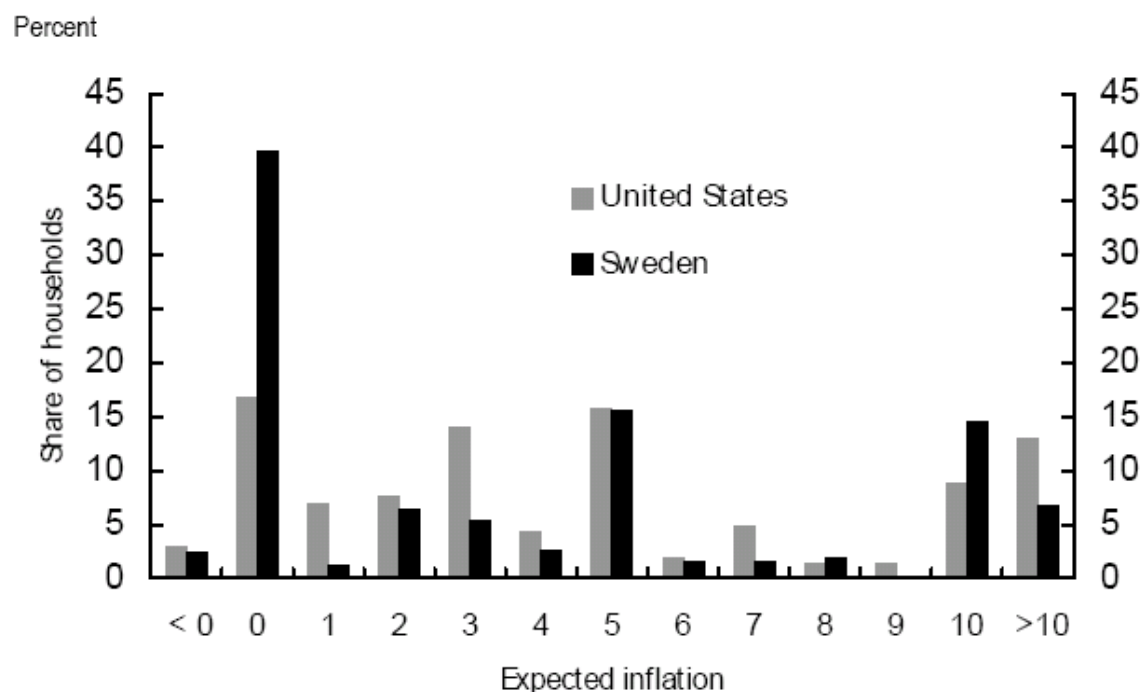
First, Curtin (2006) notes a concentration of responses around zero. Over the sample period, only 3% of observations related to an expectation of deflation, whilst 18% of all respondents chose zero inflation. This truncation effect is not a product of averaging across surveys, as Curtin shows that it holds true both for 1980 (when inflation expectations reached their peak) and for 2001 (their lowest point). Curtin suggests a

number of potential explanations for this “zero barrier” including: psychological aspects; a belief on the part of consumers that the underlying probability distribution is discontinuous at zero; or simply a rounding effect when consumers expect a very low rate of inflation.

Second, Curtin highlights a phenomenon of “digit preference” for certain numerical values, such as 0, 5, 10, 15 and so on until 50. This has frequently been observed in the results of open-ended numerical questions, irrespective of the subject (Baker, 1992; Edouard and Senthilselvan, 1997). On the other hand, the prevalence of digits such as 3, 7, 13 and 17 may be to some extent an artefact of the coding of range responses (e.g. “0 to 5%”) implemented by the SRC.

Interestingly, Bryan and Palmqvist (2006) make similar findings when considering the distribution of the Swedish SCB survey over the period 1979q1-2001m12 and report concentrations of responses around 0, 5 and 10% especially, and to a lesser extent around other points. Figure 136 reproduces their comparison of the Swedish and US distributions over the sample periods considered.

FIGURE 136: DISTRIBUTION OF INFLATION EXPECTATIONS, US & SWEDEN



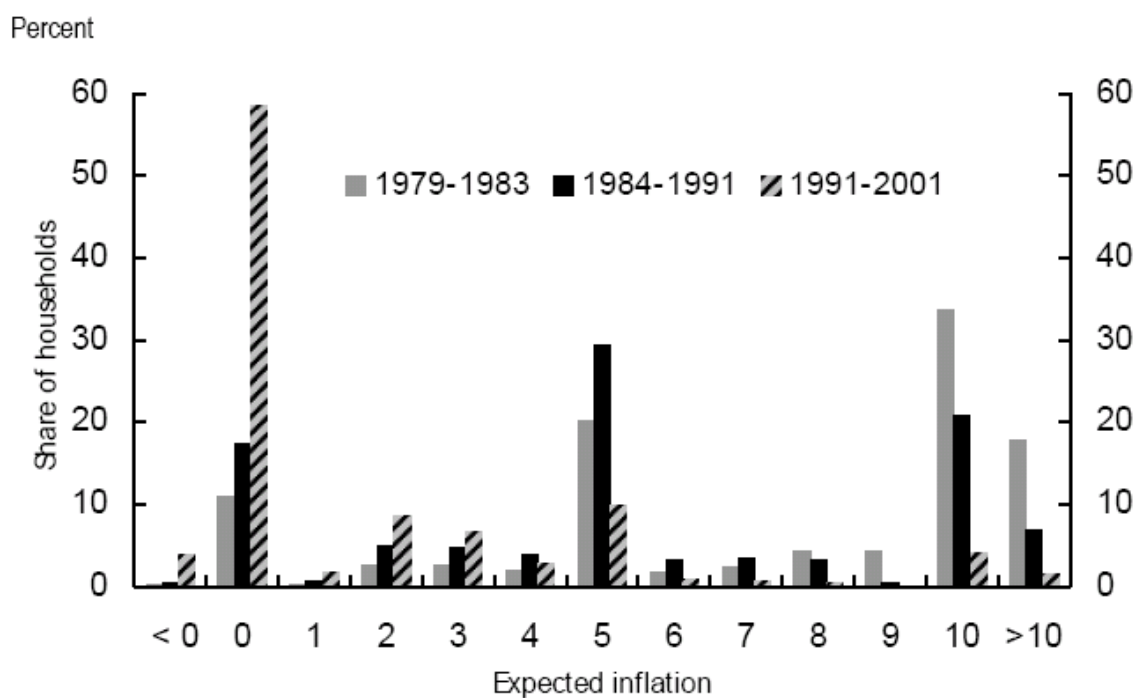
Source and notes: Bryan and Palmqvist (2006), Figure 10, p.25. The US distribution is based on the Michigan SRC survey over the period 1978-1999.

Bryan and Palmqvist interpret these findings in terms of the existence of “focal points” in the distribution, defined as points at which the proportion of responses is greater than that for the integer responses immediately above and below them. The authors hypothesise that this is a consequence of respondents forming “qualitative expectations”, an idea that was examined in a different context above, in Section 6.6.3. We note however that Curtin (2006) disagrees strongly with this idea, seemingly on the grounds that it would imply that qualitative expectations hold wherever digit preferences have been observed – for income, assets, prices etc. We reconsider this debate further below.

When examining the stability of the focal points identified month by month, the authors find that those associated with the values 0%, 5% and 10% appear to be stable, whereas

others shift as inflation regimes change (e.g. from a focal point at 3% early in the sample to one at 2% in the last inflation regime). Even if a particular point can be described as a focal point throughout the sample, the proportion of responses it attracts can change dramatically across inflation regimes, as illustrated in Figure 137.

FIGURE 137: SWEDISH INFLATION EXPECTATIONS ACROSS REGIMES



Source: Bryan and Palmqvist (2006), Figure 12, p.30.

The enormous jump recorded for response proportions relating to zero inflation expectations is not mirrored in the US data despite a similar inflation history, and is explained by the authors as a consequence of policy. Concretely, the authors credit the inflation targeting objective of the Riksbank centred at 2% for this development, and explain the popularity of the “zero” response relative to the “2%” option as a

consequence of a greater understanding by the public of the objective of “price stability” than of the specific inflation target.

This review of past empirical results therefore suggests a number of hypotheses that can be analysed in the context of the UK using the Gallup and BOE/NOP surveys. These concern:

1. The presence of a “zero barrier” in the distribution of expectations.
2. The presence of “digit preferences” and the interpretation of the latter in terms of the formation of “qualitative expectations”.
3. The influence of the “inflation regime” on the distribution of expectations, notably in what regards the location and popularity of potential focal points.
4. The influence of policy – and in particular inflation targeting – on the distribution of inflation expectations.

Examining the hypotheses in light of the Gallup and BOE/NOP surveys

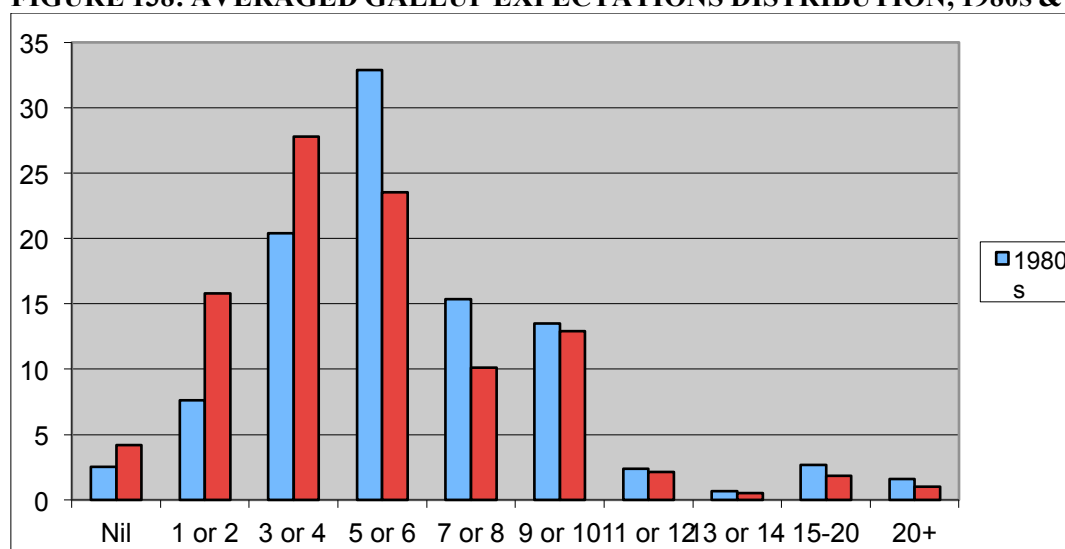
We first consider the results relating to the Gallup survey (introduced in Chapter 1), which runs monthly from January 1982 until January 1997, with some gaps in 1982 ¹⁴³. The survey elicits inflation expectations for the coming 12 months on the basis of an open-ended question. In addition to providing the ‘true mean’ obtained from these individual data points, Gallup summarised the distributional dimension of the results in its publication by classifying the answers in 10 categories, which we detailed in Table 73 earlier (page 391). It is notable that those categories do not include a ‘deflation’ option. Thus, the *presentation* of the results chosen by Gallup – no doubt motivated by the high

¹⁴³ The gaps concern the following months: 1982m4, 1982m7-m12

inflation environment in which the survey was started – implies *by itself* a truncation at zero, and if the results discussed above for the US and Sweden are robust, we should expect to see an even more significant “zero barrier”, if anything.

Figure 138 explores this issue by averaging the proportions of responses falling in each bar chart bin, and splitting the sample in two to present results for the sample years in the 1980s and the 1990s separately.

FIGURE 138: AVERAGED GALLUP EXPECTATIONS DISTRIBUTION, 1980s & 1990s



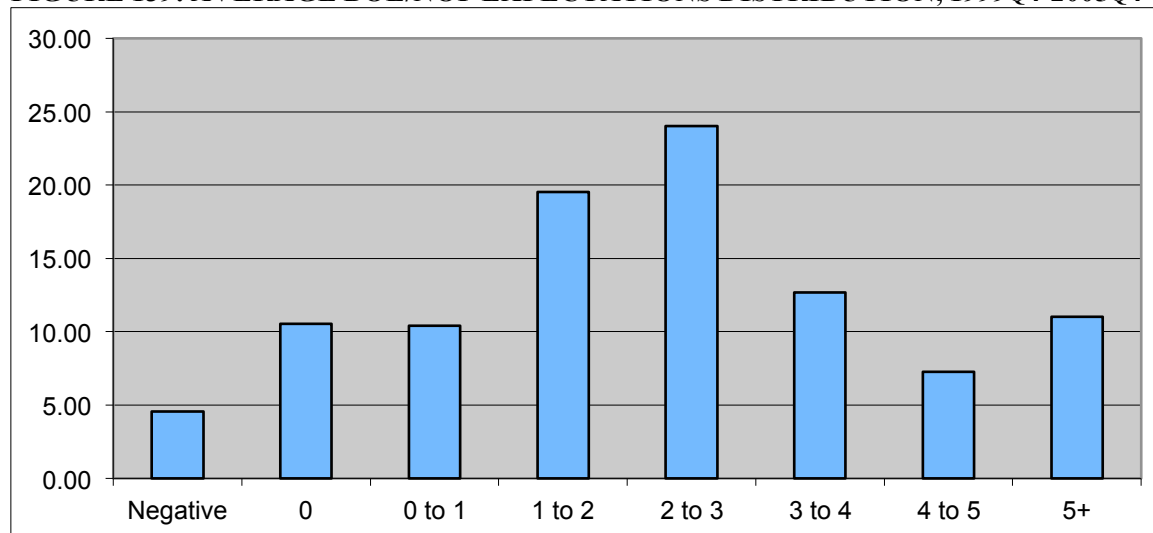
NOTE: The sample concerned is 1982m1-1997m1 (see footnote below).

As would be expected, the inflation distribution shifts leftwards when considering the inflation “regime” of the 1990s. However, as can be seen, there is no sign of a “zero barrier” in this survey despite the truncation implied by the presentation, not even in the relatively low inflation period of the 1990s. Neither is there very compelling evidence of focal points. Only the category “15-20%” would correspond to Bryan and Palmqvist’s definition of the concept over both periods, whilst the category “9-10%” emerges as a possible “focal point” in the 1990s. A reason for the notable difficulty in detecting such points here is the presentation of the data, grouping two percentage points in inflation per

bin. However, the focal point findings of the previous authors are so extreme that even this presentational difference should not mask such a phenomenon, were it to exist in the underlying data.

We now perform a similar exercise by presenting the average response proportion in each of the bins of the BOE/NOP survey in Figure 139. Unlike the other surveys considered here, the BOE/NOP survey is not based on an open-ended question, but instead asks respondents to choose from a discrete set of numerical categories to state their expectations of inflation (see Section 6.6.2).

FIGURE 139: AVERAGE BOE/NOP EXPECTATIONS DISTRIBUTION, 1999Q4-2005Q4



Here again, it is not possible to find compelling evidence of a “zero barrier”, although the category would be marginally classified as a focal point in Bryan and Palmqvist’s definition. The category “5+ %” also has the characteristics of a focal point, but given that it represents a *de facto* truncation this may not have any particular significance.

Thus, there appears to be precious little evidence in the UK for either a zero barrier or the focal points which feature so prominently in the US and Swedish surveys previously analysed in the literature.

One candidate explanation for such a remarkable difference relates to survey methodology. It might be thought that “probing” procedures may be more or less strict for different surveys. Thus, in “lax” surveys, respondents would be able to choose a “don’t know” option relatively easily, whereas in “strict” surveys respondents would be probed further until a defined answer is obtained. For instance, it is known that in the case of the Michigan Survey *“whenever a respondent would give a range response, it was always probed for a more exact point estimate”* (Curtin, 2006). Thus, it may be thought that particularly intense probing would result in a squeeze of the “don’t know” category and that respondents that may have wished to resort to that option choose instead a “round” or “focal” number to reflect their uncertainty or lack of precise knowledge of the situation. Of course, there may be other reasons behind different “don’t know”/focal point answer trade-offs; “cultural” factors may even be at play (e.g. in terms of the degree of preference between an “imprecise” focal answer and the “don’t know” category). However, irrespective of the exact causality, the question of the relationship between the “don’t know” category and the presence of focal points remains of relevance. Table 80 presents the proportion of “don’t know” responses collected in the various surveys considered. For Sweden, only a very short and recent sample was available.

TABLE 80: US, UK, SWEDEN % OF “DON’T KNOW” RESPONSES

Country	Survey	Sample	Mean of “Don’t Know” responses (%)
UK	Gallup	1982m1-1997m1 ¹⁴⁴	19.2%
	Bank of England / NOP	1999q4-2005q4	13.8%
US	Michigan SRC ¹⁴⁵	1982m1-1997m1	5.9%
		1997m1-2005m9	11.7%
Sweden	SCB / GfK	2005m10-2006m1	8.7%

As can be seen, the UK quantitative surveys do tend to have relatively large numbers of respondents in the “don’t know” category. This is despite the fact that the BOE/NOP survey “helps” respondents by offering them a number of pre-defined inflation ranges to choose from and covers a period of low inflation uncertainty. The contrast between the Gallup survey and the Michigan survey over the period 1982m1-1997m1 is particularly striking, with the Gallup “don’t know” proportion being over three times that of its US counterpart (despite the fact that the Gallup sample concerned employees exclusively, which might have been thought to be a “better informed” group than “consumers” as a whole). Thus, there may be some merit to the argument that the clear presence of focal points in the US and Sweden is not unrelated to smaller “don’t know” categories, which may in turn be reflecting differing survey procedures. However, the amplitude of the phenomenon in the US and Sweden suggests that other forces are at work, and that this constitutes at best a small explanatory contribution.

We now turn to an examination of the impact of policy on the distribution of inflation expectations.

¹⁴⁴ With gaps: 1982m4, 1982m7-m12

¹⁴⁵ The average figures provided concern *both* the “Up, Don’t know how much” and “Don’t Know, NA” categories

Monetary policy and the distribution of inflation expectations

As already noted earlier, Bryan and Palmqvist (2006) interpreted their findings for the Swedish expectations distribution as a result of the *policy* framework in place. The authors argued in particular that this may be the reason behind the large difference between the US and Swedish surveys regarding the proportion of people expecting zero inflation. Indeed, this proportion is over three times as large in Sweden as it is in the US over the last sub-sample considered by the authors (1991-1999) whereas it was broadly comparable in earlier periods. The key difference in policy emphasised by the authors is the explicit inflation targeting system in place in Sweden. This framework can be traced back to a decision of the Governing Board of the Riksbank on 14 January 1993 establishing a CPI inflation target of 2 per cent a year with a tolerance interval of ± 1 percentage point, from 1995 onwards (Berg, 1999).

When considering this conjecture, it appears especially interesting to examine the British experience in some detail, given the status of the country as an early and notable inflation targeter. Table 81 provides an overview of key dates in recent British monetary policy history.

TABLE 81: KEY DATES IN UK MONETARY HISTORY

Date	Description
08/10/1992	First inflation target – rate of annual RPIX inflation of 1-4%. Bank asked to assess inflation prospects every quarter in independent report. Government objective for inflation to be in lower half of range by end of that Parliament.
14/06/1995	Chancellor of the Exchequer in Mansion House speech announced an inflation target of 2.5% or less as measured by RPIX.
06/05/1997	Bank of England given operational independence and creation of MPC announced.
12/06/1997	Chancellor of the Exchequer in Mansion House speech outlines remit for the MPC's symmetrical target of 2.5% annual RPIX inflation and system of accountability.
01/06/1998	Bank of England Act comes into force.
09/04/2003	Chancellor of the Exchequer's Budget speech announces Government's intention to switch to an inflation target measured by the Consumer Price Index (also known as HICP).
10/12/2003	Chancellor announces switch of inflation target to 2% annual change as measured by Consumer Price Index (also known as HICP).

SOURCE: Bank of England, <http://www.bankofengland.co.uk/monetarypolicy/history.htm>

We will now consider whether there is any visible effect of these policy changes on the distribution of expectations. Figure 140 (Appendix II) presents the distribution of expectations for the Gallup data for the month of January in the years 1992-1997; Figure 141 (Appendix II) carries out the same exercise for the BOE/NOP responses in the first quarter of the years 2000-2005.

The most striking conclusion from these figures is the lack of any special importance attached to the “zero inflation” category, much unlike the Swedish experience. This is despite the UK arguably constituting a model student in the implementation of an inflation targeting system with price stability at its heart and the attached communication strategy. Thus, if policy is having any “direct” influence on the UK distribution of

expectations, it is certainly not by making the public equate “price stability” with zero inflation. We now consider the plots in greater detail.

The only notable evolution in the expectations distributions presented in Figure 140 is the switch in the mode from “5 or 6%” to “3 or 4%” that takes place in the second half of 1992; the rest of the plots essentially show fluctuations in the salience of the mode but no major changes otherwise. In particular, there appears to be no significant reaction in the switchover from a range inflation target to a symmetrical point target of 2.5% in June 1995; indeed, the mode of the distribution overshoots this in both the 1996 and 1997 plots. Similar qualitative findings apply for the expectations distributions shown in Figure 141 relating to the BOE/NOP survey. The mode is located at “2 to 3%” and the distribution appears to be stable around it over the period. This is consistent with the RPIX inflation target of 2.5%. However, there is no sign that the switch to a 2% HICP target had any effect on the distribution.

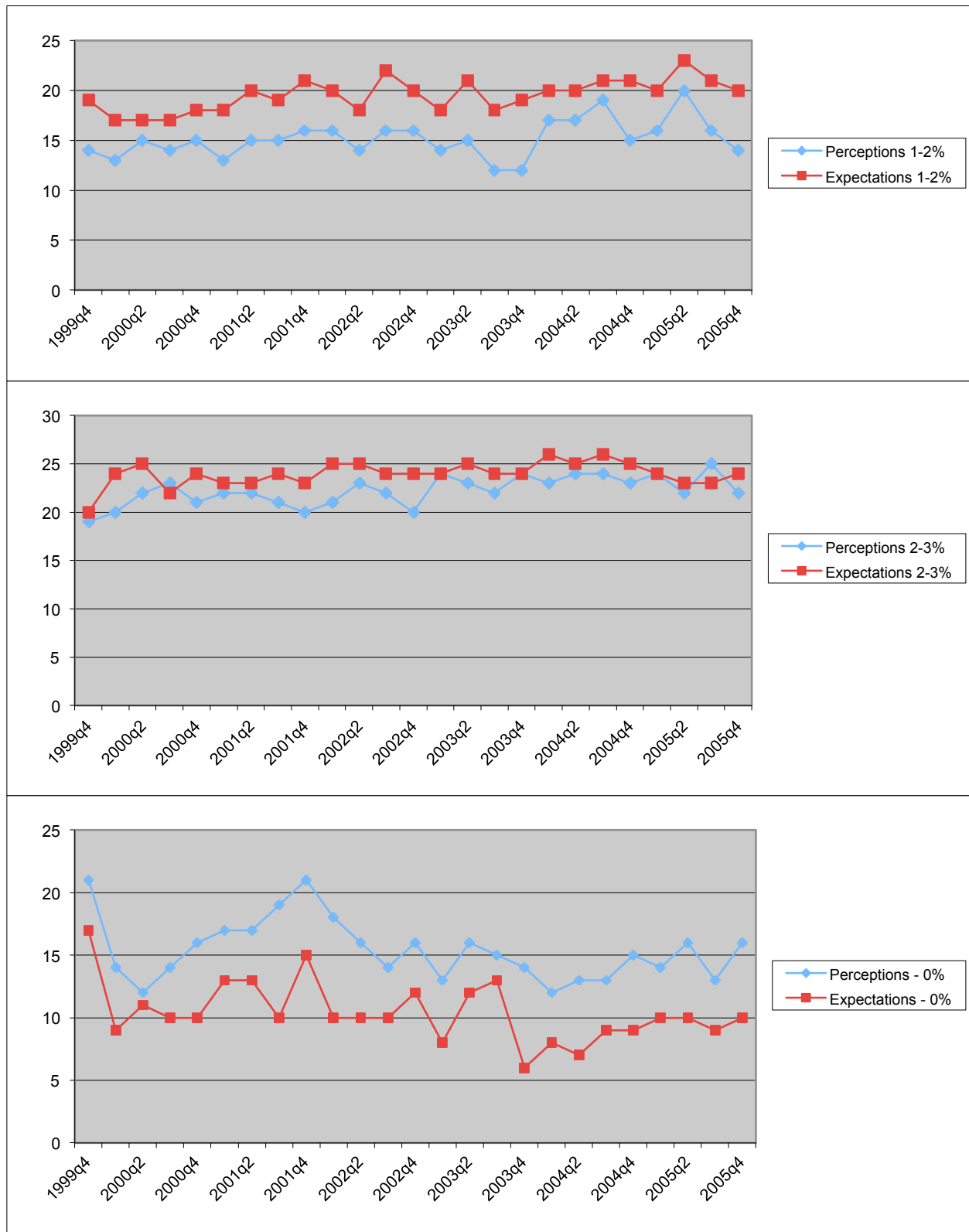
It now seems appropriate to reconsider a finding made in Section 6.6.2. In Figure 128 (page 427) we highlighted that for any given inflation range offered as a possible choice to respondents of the BOE/NOP survey, there was little difference on average between the proportion of respondents choosing that range as reflecting their *perception* and the proportion choosing it as reflecting their *expectation*. This was consistent with an earlier finding made in Section 2.2 on EC Consumer Survey data, and seems to suggest a largely adaptive/extrapolative behaviour on the part of consumers in the formation of expectations. However, Figure 128 showed that the largest differences between

perceptions and expectations could be found in the range of 0-2%. Could this be perhaps interpreted as a manifestation of the anchoring effect of the inflation target on expectations towards price stability (e.g. perceiving 4% inflation, but believing the inflation target and therefore forming an expectation in its range)?

Figure 142 studies this hypothesis in greater detail, focusing on the inflation ranges most relevant given an inflation target of 2.5% for most of this period (and 2% since December 2003). As can be seen in the uppermost panel, the proportion of respondents expecting inflation to be in the range 1-2% has consistently exceeded that perceiving inflation in this zone. However, the same is only marginally true of the range 2-3%, where the anchoring might have been expected to be the strongest. The real surprise, however, concerns the proportion of respondents perceiving or expecting zero inflation. Whereas Bryan and Palmqvist's (2006) suggested that people took inflation targeting to mean "price stability" in the very literal sense of "zero inflation", the bottom panel of the figure shows that there has consistently been a higher proportion of people currently perceiving zero inflation than expecting it.

All in all, the evidence of an anchoring effect going beyond the inertia delivered by adaptive/extrapolative expectations is not compelling.

FIGURE 142: BOE/NOP PERCEPTIONS/EXPECTATION BY INFLATION RANGE



The UK data are therefore consistent with a situation in which the public has come to expect inflation to be in a relatively low range – about 80% of the views expressed fall in the 0-4% range. There does not seem to be any direct “magic” about the operation of the inflation targeting framework other than the fact that it creates low and stable inflation, to which the public “gets used to” in a predominantly adaptive manner; this is of course not unexpected given our earlier findings and is consistent with the understanding of inflation expectations reached earlier (for example, in Section 2.1.3 or Section 6.2.2).

This contradicts the conclusion of Bryan and Palmqvist (2006) and motivates a re-examination of the logic leading to it:

1. The authors sought to explain the much greater popularity of the “zero inflation” focal point over the period 1991-1999 in Sweden relative to the US.
2. They noted that this difference arose “*despite the fact that the nations followed the same inflation trend over roughly the same period*”¹⁴⁶;
3. thus, this would “*suggest that the findings are a consequence of a changed policy*”¹⁴⁷ and in particular that “*inflation targeting in Sweden has substantially increased the proportion of Swedes who ignore inflation*”¹⁴⁸.

¹⁴⁶ Bryan and Palmqvist (2005), p.29.

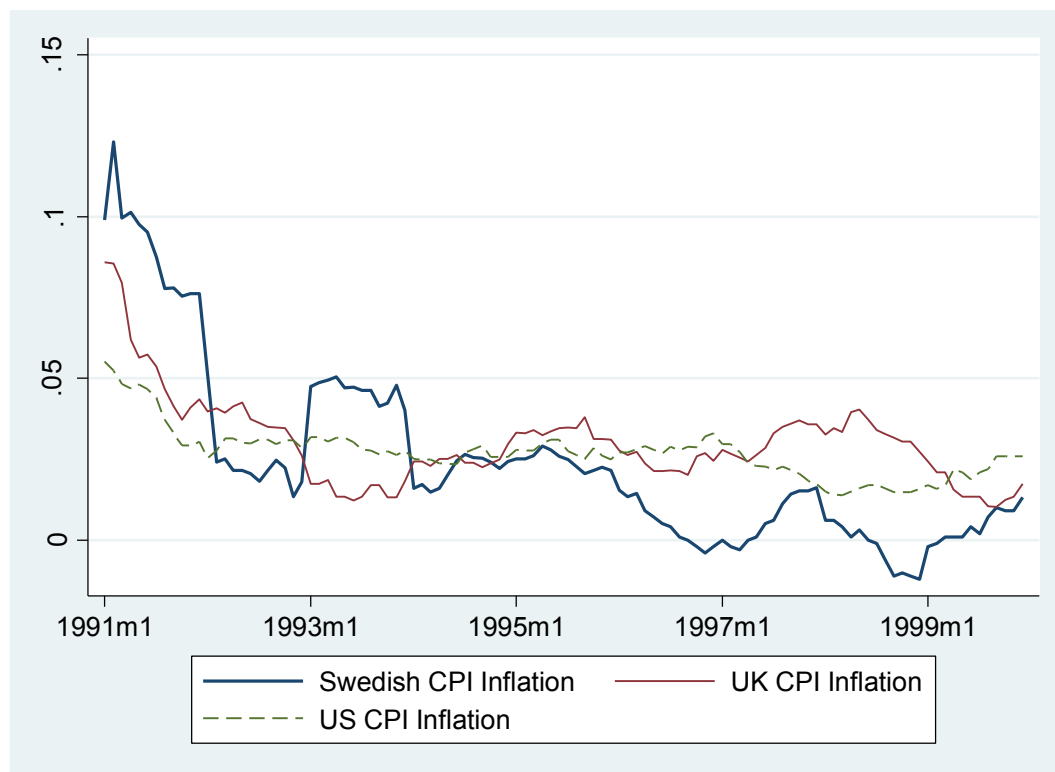
¹⁴⁷ Bryan and Palmqvist (2005), p.29.

¹⁴⁸ Bryan and Palmqvist (2005), p.31.

Essentially, the second step of their reasoning is correct but masks an important development in Swedish inflation over the period, which in turn implies that their third step does not follow.

Figure 143 plots CPI inflation (all items) in Sweden, the US and the UK over the period of concern (1991-1999):

FIGURE 143: US, SWEDEN AND UK, 1991-1999 CPI INFLATION



As is clear from the plot, Swedish CPI inflation has been inferior to that in the other two countries from the end of 1994 onwards. It has taken values close to zero and has even moved into the area of deflation at several points. Such a period of low inflation/deflation is an important development; indeed, actual deflation is a rather unusual event, and leads

to the suspicion that expectations should be affected proportionately. This is not highlighted in Bryan and Palmqvist (2006), as the authors are instead considering *core* CPI inflation for Sweden, which shows broadly the same qualitative characteristics but does not actually move into negative territory. The core CPI measure in Sweden (UND1X) excludes mortgage rates, as well as the direct effects of changes in indirect taxes and subsidies. However, an important reason for the behaviour of actual CPI at the end of 1998 and the beginning of 1999 was an easing of monetary policy, with the consequence of lower interest expenditure for households; this took place in the context of a number of “*eye-catching media headlines on the threat of deflation*” (Sveriges Riksbank, 2003). This may well be the reason for the dramatic increase in “zero inflation” responses observed, rather than any *direct* linkage to the policy regime. This interpretation would then make the US, Swedish and British findings consistent.

The existence of focal points is evident in the data for the US and Sweden. It remains a puzzle why this is not also the case of the United Kingdom. A partial explanation may be provided by the survey methods employed or more generally the posited trade-off with the “don’t know” category, as already discussed. A more complete explanation may perhaps involve the notion of British consumers being *relatively* better informed about inflation – at least those respondents who provide an answer other than “don’t know”. Indeed, some arguments in this direction relating to specificities in labour and housing markets were already presented in Section 5.6. Further distributional evidence for the UK and other countries would be invaluable in shedding light on this peculiarity.

6.6.5 On the comparative advantages of quantitative and qualitative surveys

This study has examined many types of survey data, which can be broadly categorised as follows:

1. Qualitative survey data (EC Consumer Survey)
2. Quantitative survey data
 - a. Collected on the basis of open-ended questions (Gallup data for the UK; Michigan SRC data for the US; SCB/GfK data for Sweden)
 - b. Collected on the basis of “range” options offered to respondents (BOE/NOP data for the UK)

Very few studies have considered qualitative and quantitative survey data on inflation expectations in a comparative light. The evidence collected as part of this research allows us to make some new remarks regarding the relative desirability of both types of data.

The appeal of quantitative data is obvious. The data sought are obtained directly from survey respondents, thereby by-passing the potentially hazardous process of quantification of qualitative survey data (see Chapter 1), which may well introduce measurement error. However, two caveats to this argument need to be emphasised.

First, not all types of quantitative data can be considered ideal. McFadden et al. (2004) notably emphasise the importance of the cognitive processes taking place when respondents answer a survey. Indeed, it appears that respondents extract information *from the survey questionnaire itself* when providing their answers. This is particularly visible in the “bracketing effect”, a well-established phenomenon in the psychology and

survey research literature (e.g. Schwarz et al., 1985; Schwartz et al., 1998; Menon et al., 1997; Winter, 2002). The “bracketing effect” occurs in situations where respondents are asked to provide their perceptions/expectations by choosing among predefined ranges (e.g. inflation will be “0-1%”, “1-2%” etc.). Such an elicitation procedure is deployed in the Bank of England / NOP survey, for instance (see Section 6.6.2). The problem with such an approach is well summarised by Sudman et al. (1996):

“Respondents assume that the range of response alternatives reflects the researcher’s knowledge of or expectations about the distribution of behaviour in the ‘real world’. Thus, response alternatives in the middle of the range would reflect ‘average’ or ‘typical’ behavioural frequency whereas the extremes of the scale would correspond to the extremes of the distribution”.

McFadden et al (2004) experimented with a number of treatments to elicit perceptions on the proportion of Americans that would be admitted to a nursing home (low-bracket range, high-bracket range and open-ended question) in the context of the AARP Retirement Perspectives Survey; their results confirm the significance of the bracketing effect. The extent to which such an effect would impact the BOE/NOP survey is unclear *a priori*, but these arguments certainly suggest treating its results with some caution.

Second, caveats apply even in the case of quantitative surveys relying on open-ended questions. A counter-argument that is often made in certain statistical institutes is that consumers are too unsophisticated a group to answer quantitative questions¹⁴⁹. This position is much too strong – the Gallup data, the BOE/NOP survey and the SCB/GfK survey have all provided useful information in this study. Notwithstanding this, as we have seen in previous sub-sections, a strong case can be made to support the hypothesis

¹⁴⁹ Indeed, we understand that this argument has currently resurfaced in the debate between the European Commission and national statistical agencies as to whether to continue the pilot project of collecting quantitative data on inflation perceptions and expectations.

that consumers form *qualitative* rather than *quantitative* expectations to a large extent. Indeed, Section 6.6.3 emphasised the novel result that “don’t know” responses are considerably more widespread in *quantitative* surveys than they are in *qualitative* surveys on inflation expectations; furthermore, the previous section documented the “focal point” results obtained for US and Swedish quantitative data in the literature, which are also suggestive of qualitative expectations formation (although such findings could not be replicated on UK data). Thus, qualitative surveys may be able to collect some information that is not adequately captured by quantitative surveys, as they provide an elicitation procedure closer to the conceptualisation of inflation held by many in the general public.

Overall, when taking all elements into account, it seems inappropriate to favour *a priori* one type of data over another, as it would require judgements that are not possible to make (e.g. weighing the measurement errors from quantification procedures against those caused by the elicitation procedure of quantitative methods). In fact, both types of data may in many ways provide complementary information. Chapter 1 in this study found the availability of quantitative data useful as a benchmark for quantified qualitative data, in the spirit of Batchelor (1986). Furthermore, Section 2.3.2 above found that taking into account the failure of the rational perceptions hypothesis in the quantification process brought quantified qualitative expectations more in line with the quantitative expectations data in Sweden. Thus, in the current state of knowledge, efforts in collecting quantitative data in parallel with qualitative data – such as those currently sponsored by the European Commission – seem to be fully appropriate and useful investments. If continued, they ought to be exploited by future research.

7 CONCLUSION

This chapter has covered a great variety of empirical applications of inflation expectations data. As most of the detailed results have been summarised and discussed in the concluding comments of the individual sections, the aim here is instead to draw a few common threads from them and highlight a few key messages.

First, as was discussed in Section 2, the inflation expectations of consumers are complicated processes. They show many adaptive characteristics, yet are not purely adaptive. In any event, they are certainly not “rational”. They are also subject to a touch of “animal spirits” as was well-illustrated by the episode surrounding the introduction of EMU. In fact, the results provided here highlighted that even *perceptions* of inflation could not be sensibly described as “rational”, with over- and under-weighting of HICP components being the norm. Even controlling for these distortions in perceptions does not increase the success of “rationality” tests for expectations. Nevertheless, furthering our understanding of these distortions appears to be an important research avenue for the future – not least because it may help improve the quantification of qualitative survey data, as our evidence suggested.

Second, attempts to move beyond rational expectations by espousing simple “composite” models of expectations are broadly rejected by the data, notwithstanding their popularity in the literature. This was shown in Section 3 through tests of the aggregate “near-rationality” hypothesis of Akerlof et al. (2000), according to which some agents in the population are rational and others nearly so, in varying proportions. A similar point was

made in Section 4 by showing the lack of robustness of the “rule-of-thumb” hypothesis. Such findings cast doubt over vast areas of theoretical and empirical work which treat expectations as an assumption. Clearly, a new paradigm for them is needed.

Third, some recent contributions to the literature have precisely been encouraging in that respect. Section 5 studied the interaction between expert and consumer expectations. This was met with limited success in most countries, although small sample problems may have been an issue. In contrast, the results for the UK were surprisingly successful showing an influence of expert forecasts on consumer expectations, and supporting the role of the news media in the transmission mechanism, as posited by Carroll (2003). Institutional and forecasting reasons for the British exception were suggested; greater availability of data will no doubt prove invaluable for shedding light on this issue.

Fourth, Section 6 added another dimension to our understanding of expectations by considering their *distribution*. An examination of the Mankiw et al. (2003) view of expectations confirmed their intuition that these are *gradually adjusting*; however, this updating process does not involve rational expectations and looks rather adaptive, confirming our earlier findings. Overall, the expectations distributions of the quantitative data considered here showed arguments implying bimodal shapes to be unconvincing. On the other hand, a case was made highlighting the *qualitative* features of expectations, both by considering evidence on the behaviour of “don’t know” responses across qualitative and quantitative surveys, as well as evidence on “focal points” - although once again, the UK appeared to be an exception on the latter point. Thus, overall, the

objections raised by Curtin (2006) and mentioned earlier appear to be unconvincing in light of the data. Indeed, unlike him, we would not be surprised to see that qualitative perceptions or expectations extend to other areas beyond consumer price inflation (e.g. income, assets). Furthermore, such behaviour would be consistent with the types of “bracketing” effects mentioned in Section 6.6.5. Finally, the findings here emphasise the importance of neither over- nor underestimating the influence of policy on the minds of the general public. Indeed, the British evidence does not reveal the image of a nation of consumers paying daily attention to the evolution of its inflation targeting system or institutions. On the other hand, the distribution of expectations does evidently respond to inflation regimes. Thus, consumers acknowledge good policy in their own way – *gradually*.

CONCLUSION

This thesis has been concerned with the *quantification* of inflation expectations (Chapter 1), their *application* in a wage equation setting (Chapter 2) and finally, with the *properties* of inflation expectations and perceptions in their own right (Chapter 3). The main focus was placed on consumer inflation expectations in the context of eight countries of the European Union (henceforth “our sample countries”).

CHAPTER 1: Finding the Optimal Method of Quantifying Inflation Expectations on the Basis of Qualitative Survey Data

Chapter 1 carried out an empirical contrast between a number of quantification methodologies for qualitative survey data on inflation expectations, belonging to three distinct traditions in the literature: the Carlson-Parkin, Pesaran and Seitz approaches. The predictive ability of the alternative quantified expectations series was considered, although this was argued to be an inappropriate criterion outside a world of rational expectations. In any event, generalised underperformance relative to a “naïve expectations” benchmark was found. Instead, the empirical assessment in the chapter rested on two pillars. First, the quantified expectations series were contrasted with comparable *quantitative* expectations data available in the case of the United Kingdom and Sweden. Second, those series were used as part of an exercise in modelling wage equations for the eight countries in our sample; this was intended to discriminate between quantification methodologies on the basis of their ability to generate results that were sensible and interpretable, both statistically and economically.

Overall, the best results were achieved by a variation of the Carlson-Parkin methodology introduced in Curto Millet (2004). This method uses survey data on inflation *perceptions* as part of the process of quantification of expectations. We argued that a key distinctive feature in its success was that it allowed for the expression of long-run trend changes in the “indifference thresholds” underlying the distribution of perceptions.

Furthermore, two important additional findings were made in the wage equation modelling exercise. First, expectations were found to play a significant role in all the countries in our sample *except* France. It was argued that this was the result of a critical measurement error generated by the formulation of the French questionnaire. Second, it was found that the presence of expectations was important in reaching wage equations that were richer and more economically interpretable.

The database on consumer inflation expectations *constructed* in this chapter following an important data recovery exercise is the largest available in European Union countries to date. It was further leveraged in the following chapters of this study.

CHAPTER 2:

The Impact of European Monetary Unification on Labour Market Flexibility

Chapter 2 investigated the hypothesis that the formation of the European Monetary Union might have led to an increase in the ‘flexibility’ of its participating labour markets.

The theoretical basis for such a claim was established, and a survey of the relevant empirical literature was presented; this concerned both the EMS experience and more

recent work focusing on EMU. It was argued that these past studies presented important shortcomings, one of the most critical being their failure to control adequately for the effects of inflation expectations and institutional variables. Indeed, accounting for these variables would be necessary to lend robustness to the analysis, especially in a period of such marked changes as that surrounding the introduction of the euro.

This study proposed to address these and other issues by developing a wage equation specification with an equilibrium correction mechanism. The theoretical roots of this equation were double: the shape of the long-run solution was informed by a bargaining model à la Layard et al. (1991), whilst the dynamics were designed to be consistent with a staggered contracts approach. Inflation expectations data were obtained according to the preferred methodology selected in Chapter 1, and a database of institutional variables was compiled and extended appropriately for this study – the most up-to-date to be used on many of these dimensions yet. These included union density, bargaining coverage and coordination, employment protection legislation, benefit duration and replacement rates.

Our wage equations were implemented for the six Eurozone countries in our sample, as well as for two non-members (the United Kingdom and Sweden). Model selection followed the ‘General-to-Specific’ methodology. The null hypothesis was examined through the stability properties of our wage equation – with a battery of stability tests, the forecasts of the model in the post-1999 period and the behaviour of recursive coefficients. The behaviour of the speed of adjustment, productivity and unemployment coefficients were studied with particular care.

Overall, no evidence of an “EMU effect” on wage flexibility could be found, nor in fact of an “EMS effect” in the two countries whose samples enabled the study of this earlier phenomenon (Italy and the Netherlands). However, the prospect of EMU membership may have been helpful in other respects; these include recent wage moderation in Belgium and the overall strong productivity orientation of the Eurozone bargaining equilibrium, notably in the cases of Italy and Spain. Nevertheless, our equations also highlight a cost of EMU. The finding of highly significant and stable coefficients for inflation expectations across the Eurozone (except France) combined with the knowledge from Chapter 3 of the important rise in inflation expectations associated with EMU means that the latter translated into higher wage inflation in recent years. This therefore ought to be counted as a cost of EMU, albeit possibly a short-term one.

Finally, our attempt to better identify institutional effects by imposing pooled mean group restrictions made clear that this would have to be at the cost of bias. Institutional heterogeneity across countries is thus too important to ignore.

CHAPTER 3: **Properties and Applications of Data on Inflation Perceptions and Expectations**

Chapter 3 was concerned with the empirical properties of inflation expectations and perceptions *themselves*. These were examined through a set of distinct applications, from which a number of important findings emerge.

First, expectations and perceptions were shown to have been pushed upwards and influenced in a complex way by the events surrounding monetary unification, notably

generating a historic divergence between both concepts. The behaviour of the actual expectations data turns the conventional wisdom of ‘credibility theorists’ on its head. There was no obvious ‘credibility’ gain from all the elaborate building of independence of the ECB in the economy at large. In fact, this is hardly surprising given the deviations from rational expectations exhibited by the data. The evidence thus suggests that the case for central bank independence ought to be based on considerations other than credibility theory (Forder, 1998).

The failure of the rational expectations hypothesis itself is hardly surprising given the failure of the rational *perceptions* hypothesis found by this study. The chapter demonstrated systematic deviations between what people conceptualise as ‘prices in general’ and CPI-type indices. The pattern of these deviations was explored for the first time in the literature. Taking such effects into account was shown not to render expectations more rational relative to an adjusted price benchmark, but proved promising in terms of improving the accuracy of quantification methods.

Second, the chapter showed that commonly used models of expectations conceptualising these as a simple combination between “rational” and “non-rational” components were misleading. In the aggregate, no evidence was found to support the Akerlof et al. (2000) hypothesis according to which the population is divided between a fraction of “rational” agents and another of “near-rational” ones. Similarly, the popular “rule-of-thumb” expectations model – according to which these are a combination of adaptive and rational

expectations behaviour – was found to be intrinsically unstable, being misspecified and both sample- and measure-dependent.

Third, the relationship between expert forecasts and consumer inflation expectations was considered. The null hypothesis explored was the view that expert forecasts had a higher predictive ability than those of consumers and fed into the latter through an updating mechanism intermediated by the news media. The evidence here proved to be mixed for all countries considered – possible due in part to the small samples used – except, remarkably, for the United Kingdom. Indeed, we showed that Carroll's (2003) model of inflation expectations is congruent in the case of the UK, and that his predictions regarding the influence of news coverage of inflation on the gap between consumer and expert forecasts were verified.

Fourth, we considered the distributional aspects of inflation expectations and perceptions. We showed in the context of the Volcker and ERM disinflations that the updating of expectations by consumers is indeed gradual, as posited by Mankiw and Reis (2002), but is certainly not rational – displaying instead rather adaptive features. Overall, the expectations distributions of the quantitative data considered here showed arguments implying bimodal shapes to be unconvincing.

Finally, a case was made in support of the view that consumer expectations are formed qualitatively to some extent. Evidence of “focal points” was discussed (although the United Kingdom once again stood out by failing to display these) and new evidence regarding the behaviour of the “don't know” answer category was presented in support of

this case. Accordingly, we showed that the distribution of consumer expectations was clearly influenced by the *inflation regime* in force, but bore no visible link with the minutiae of inflation targeting implementation.

CONCLUDING REMARKS

This study has made three important contributions to the debate on the direction future macroeconomic research ought to take.

First, Chapter 1 showed that vast data resources on inflation expectations and fully satisfactory quantification methodologies are available for empirical research. There is in the present day no valid justification for a divergence between theoretical and applied work in this respect.

Second, Chapter 2 indeed confirmed that inflation expectations were an important variable to be used in support of empirical work where expectations are not the main focus. The use of expectations in applications other than rational expectations testing is well overdue.

Third, Chapter 3 showed in addition that there is a veritable “gold mine” of research to be explored in what regards the properties of expectations and perceptions *themselves*.

The prospecting of this site is still in its very early stages. Yet the conclusions already reached have been successful in overturning much received knowledge in macroeconomics and stimulating further research.

Expectations of future yields run high.

REFERENCES

- Agell, J. and P. Lundborg (1995). "Theories of Pay and Unemployment: Survey Evidence from Swedish Manufacturing Firms." *Scandinavian Journal of Economics*, 97.
- Agell, J. and P. Lundborg (1999). "Survey Evidence on Wage Rigidity and Unemployment: Sweden in the 1990s." *IFAU Working Paper 1999:2*.
- Akerlof, G. (1970). "The Market for Lemons: Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*, 89, pp. 488-500.
- Akerlof, G. (1982). "Labour Contracts as Partial Gift Exchange." *Quarterly Journal of Economics*, November.
- Akerlof, G. and J. L. Yellen (1990). "The Fair Wage-Effort Hypothesis and Unemployment." *Quarterly Journal of Economics*, May.
- Akerlof, G. A., W. T. Dickens, and G. L. Perry (2000). "Near-Rational Wage and Price Setting and the Long-Run Phillips Curve." *Brookings Papers on Economic Activity*, 1, pp. 1-60.
- Alesina, A. (1987). "Macroeconomic Policy in a Two-Party System as a Repeated Game." *Quarterly Journal of Economics*, 102, pp. 651-78.
- Alesina, A., O. Blanchard, J. Galí, F. Giavazzi, and H. Uhlig (2001). "Defining a Macroeconomic Framework for the Euro Area. ." *Monitoring the European Central Bank No. 3, CEPR*.
- Alho, K. E. (2000). "The Finnish EMU Buffers and the Labour Market Under Asymmetric Shocks." *ENEPRI Workshop "EMU and Asymmetries in Labour Markets"*: Paris.
- Anderson, O. (1952). "The Business Test of the IFO-Institute for Economic Research, Munich, and its Theoretical Model." *Revue de l'Institut International de Statistique*, 20:1-17.
- Anderton, R. and R. Barrell (1995). "The ERM and Structural Change in European Labour Markets: A Study of Ten Countries." *Weltwirtschaftliches Archiv*, 131/1, pp. 47-66.
- Anderton, R., R. Barrell, J. W. In't Veld, and N. Pittis (1992). "Forward-Looking Wages and Nominal Inertia in the ERM." *National Institute Economic Review*: August, pp. 94-105.
- Aron, J., J. Muellbauer, and B. W. Smit (2004). "Modelling the Inflation Process in South Africa." *Keynote Address, Eighth Annual Conference on Econometric Modelling for Africa*: Stellenbosch University, South Africa.
- Aron, J., J. Muellbauer, and B. W. Smit (2004). "A Structural Model of the Inflation Process in South Africa." *CSAE WPS/2004-08*.
- Artis, M. J. and P. Ormerod (1991). "Is there an "EMS effect" in European Labour Markets?" *repr. with addenda in C. Johnson and S. Collignon (eds.), The Monetary Economics of Europe, Pinter, 1994* (repr. with addenda in C. Johnson and S. Collignon (eds.), *The Monetary Economics of Europe*, Pinter, 1994).
- Artis, M. J. and P. Ormerod (1996). "Another look at the "EMS Effect" in European Labour Markets," in *Inflation and Wage Behaviour in Europe*. P. de Grauwe, S. Micossi and G. Tullio eds: Oxford University Press, 1996.
- Artus, P. and R. Salomon (1996). "The EMS, Credibility and Disinflation: The French Case," in *Inflation and Wage Behaviour in Europe*. P. de Grauwe, S. Micossi and G. Tullio eds: Oxford University Press, 1996.

- Baghestani, H. (1992). "Survey Evidence on the Muthian Rationality of the Inflation Forecasts of US Consumers." *Oxford Bulletin of Economics and Statistics*, 54:2, pp. 173-86.
- Baghestani, H. (1992). "Survey Evidence on the Muthian Rationality of the Inflation Forecasts of US Consumers." *Oxford Bulletin of Economics and Statistics*, 54:2, pp. 173-86.
- Bai, J. and P. Perron (1998). "Estimating and Testing Linear Models With Multiple Structural Changes." *Econometrica*, 66:1, pp. 47-78.
- Bai, J. and P. Perron (2003a). "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics*, 18, pp. 1-22.
- Bai, J. and P. Perron (2003b). "Critical Values for Multiple Structural Change Tests." *The Econometrics Journal*, 6:1, pp. 72-78.
- Baker, Michael (1992). "Digit Preferences in CPS Unemployment Data." *Economic Letters*, 39:1, pp. 117-21.
- Bakhshi, H. and A. Yates (1998). "Are UK Inflation Expectations Rational?" *Bank of England, Working Paper No81*.
- Ball, L. (1994). "Credible Disinflation with Staggered Price Setting." *American Economic Review*, LXXXIV, pp. 282-89.
- Ball, L. and D. Romer (1991). "Sticky Prices As Coordination Failure." *American Economic Review*, 81.
- Baltagi, B. H. (2001). *Econometric Analysis of Panel Data*: Wiley.
- Bank of Canada (1998). "Information in Financial Asset Prices." Bank of Canada: Ottawa.
- Bank of England (2001-2005). "The Bank of England Inflation Attitudes Survey." *Bank of England Quarterly Bulletin*: Summer.
- Bank of England, . (2005). "Inflation report, November 2005."
- Bardsen, G., E.S. Jansen, and R. Nymoen (2002). "Testing the New Keynesian Phillips Curve." Norges Bank.
- Barran, F., V. Coudert, and B. Mojon (1997). "The Transmission of Monetary Policy in the European Countries," in *European Monetary Policy*. S. Collignon ed: Pinter Press.
- Barrell, R. and K. Dury (2001). "Asymmetric Labour Markets in a Converging Europe: Do Differences Matter?" *ENEPRI Working Paper No 2*.
- Barro, R. and D. Gordon (1983). "Rules, Discretion and Reputation in a Model of Monetary Policy." *Journal of Monetary Economics*, 12, pp. 101-21.
- Batchelor, R. A. (1982). "Expectations, Output and Inflation: The European Experience." *European Economic Review*, 17, pp. 1-25.
- Batchelor, R. A. (1986). "Quantitative v. Qualitative Measures of Inflation Expectations." *Oxford Bulletin of Economics and Statistics*, 48:2, pp. 99-120.
- Batchelor, R. A. (2006). "How robust are quantified survey data? Evidence from the United

States." *The Role of Inflation Expectations in Modelling and Monetary Policy Making*. National Bank of Poland: Warsaw.

Batchelor, R. A. and P. Dua (1987). "The Accuracy and Rationality of UK Inflation Expectations: Some Quantitative Evidence." *Applied Economics*, 19, pp. 819-28.

Batchelor, R. A. and A. B. Orr (1988). "Inflation Expectations Revisited." *Economica*, 55, pp. 317-31.

Batini, N., B. Jackson, and S. Nickell (2000). "Inflation Dynamics and the Labour Share in the UK." *External MPC Unit Discussion Paper No. 2*.

Baum, C. F. (2004). "SUGUK 2004 invited lecture: Topics in time series modeling with Stata."

Bazen, S. and J. P. Martin (1991). "The Impact of the Minimum Wage on Earnings and Employment in France." *OECD Economic Studies*, 16:Spring, pp. 199-221.

Bell, B. and N. Dryden (1996). "The CEP-OECD Data Set (1950-1992)." Centre for Economic Performance (CEP), London School of Economics: London.

Berg, C. (1999). "Inflation Forecast Targeting: the Swedish Experience." Sveriges Riksbank.

Berk, J. M. (1999). "Measuring Inflation Expectations: A Survey Data Approach." *Applied Economics*, 31, pp. 1467-80.

Berk, J. M. (2000). "Consumers' Inflation Expectations and Monetary Policy in Europe." *Research Memorandum, Vrije Universiteit, Amsterdam, 2000-20, June*.

Berk, J. M. (2002). "Consumers' Inflation Expectations and Monetary Policy in Europe." *Contemporary Economic Policy*, 20:2, pp. 122-32.

Berk, J. M. (2006). "Consumers' inflation expectations and monetary policy in Europe." *The Role of Inflation Expectations in Modelling and Monetary Policy Making*. National Bank of Poland: Warsaw.

Berthold, N., R. Fehn, and E. Thode (1999). "Real Wage Rigidities, Fiscal Policy and the Stability of EMU in the Transition Phase." *IMF Working Papers*, 99:83.

Bewley, T. F. (1999). *Why Wages Don't Fall during a Recession*. Cambridge, MA: Harvard University Press.

Bewley, T. F. (1999). *Why Wages Don't Fall during a Recession*. Cambridge, MA: Harvard University Press.

Blanchard, O. and R. Perotti (1999). "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *NBER Working Paper No. 7269*.

Blanchard, O. and D. Quah (1989). "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review*, 79, pp. 655-73.

Boeri, T., A. Brugiavini, and L. Calmfors (2001). *The Role of Unions in the Twenty-first Century - A Report for the Fondazione Rodolfo De Benedetti*. Oxford: Oxford University Press.

Borghijs, Alain, Sjeff Ederveen, and Ruud de Mooij (2003). "European Wage Coordination: Nightmare Or Dream To Come True? An Economic Analysis Of Wage Bargaining Institutions In The EU." *CPB Discussion Paper No. 17*.

- Borio, C (1995). "The Structure of Credit to the Non-Government Sector and the Transmission Mechanism of Monetary Policy: A Cross-Country Comparison." *BIS Working Papers No. 24*.
- Bowdler, C. and E.S. Jansen (2004). "Testing for a Time-Varying Price-Cost Markup in the Euro Area Inflation Process." *Nuffield College Economics Working Papers Ref: 2004-W10*.
- Branch, W. A. (2004a). "The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations." *The Economic Journal*, 114:July, pp. 593-621.
- Branch, W. A. (2004b). "Restricted Perceptions Equilibria and Learning in Macroeconomics." *Mimeo*.
- Breen, R. and C. García-Peñalosa (1999). "Income Inequality and Macroeconomic Volatility: An Empirical Investigation." *Review of Development Economics*, 9:3, pp. 380-98.
- Britton, E. and J. Whitley (1997). "Comparing the Monetary Transmission Mechanism in France, Germany, and the United Kingdom: some Issues and Results." *Bank of England Quarterly Bulletin*.
- Bryan, M. F. and S. Cecchetti (1994). "Measuring Core Inflation," in *Monetary Policy*. N. G. Mankiw ed. Chicago: University of Chicago Press for NBER.
- Bryan, M. F. and S. Palmqvist (2005). "Testing Near-Rationality Using Detailed Survey Data." *European Economy*:228.
- Bryan, M. F. and S. Palmqvist (2006). "Testing Near-Rationality Using Detailed Survey Data." *The Role of Inflation Expectations in Modelling and Monetary Policy Making*. National Bank of Poland: Warsaw.
- Bryan, M. F. and G. Venkatu (2001). "The Demographics of Household Inflation Surveys," in *Economic Commentary*: Federal Reserve Bank of Cleveland.
- Buiter, W. (1995). "Macroeconomic Policy During a Transition to Monetary Union." *CEPR Discussion Paper No. 1222*.
- Burda, M. (2001). "European Labour Markets and the Euro: How Much Flexibility Do We Really Need?" *ENEPRI Working Paper No 3*.
- Cadiou, L. and S. Guichard (1999a). "La Diversité des Marchés du Travail en Europe: Quelles Conséquences pour l'Union Monétaire? (Partie I)." *CEPII Document de Travail N°99-10*.
- Cadiou, L., S. Guichard, and M. Maurel (1999b). "La Diversité des Marchés du Travail en Europe: Quelles Conséquences pour l'Union Monétaire? (Partie II)." *CEPII Document de Travail N°99-11*.
- Cagan, P. (1956). "The Monetary Dynamics of Hyperinflation," in *Studies in the Quantity Theory of Money*. M. Friedman ed. Chicago: University of Chicago Press, pp. 25-117.
- Calmfors, L. (1998a). "Macroeconomic Policy, Wage Setting, and Employment - What Difference Does The EMU Make?" *Oxford Review of Economic Policy*, 14:3, pp. 125-51.
- Calmfors, L. (1998b). "Unemployment, Labour-Market Reform and Monetary Union." *Stockholm University, Institute for International Economic Studies:Seminar Papers*, No. 639.
- Calmfors, L. (2001a). "Nominal Wage Flexibility and Fiscal Policy - How Much Can They Reduce

Macroeconomic Variability in the EMU?," in *Submissions on EMU from Leading Academics*. H.M. Treasury ed.

Calmfors, L. (2001b). "Wages and Wage-Bargaining Institutions in the EMU - A Survey of the Issues." *Empirica*, 28, pp. 325-51.

Calmfors, L. and J. Driffill (1988). "Bargaining Structure, Corporatism and Macroeconomic Performance." *Economic Policy*, 6, pp. 14-47.

Calmfors, L. and A. Johansson (2002). "Nominal Wage Flexibility, Wage Indexation and Monetary Union." *Stockholm University, Institute for International Economic Studies:Seminar Papers*, No. 716.

Calvo, G. A. (1983). "Staggered Prices in a Utility Maximizing Framework." *Journal of Monetary Economics*, XII, pp. 383-98.

Cameron, G. and C. Wallace (2002). "Macroeconomic Performance in the Bretton Woods Era and After " *Oxford Review of Economic Policy*, 18, pp. 479-94.

Card, D. and D. Hyslop (1997). "Does inflation "grease the wheels of the labor market"?," in *Reducing inflation: Motivation and strategy, number 356*. C. D. Romer and D. H. Romer eds. Chicago: University of Chicago Press, pp. 71-114.

Carlin, W. and D. Soskice (1990). *Macroeconomics and the Wage Bargain*.

Carlson, J. A. and M. Parkin (1975). "Inflation Expectations." *Economica*, 42:May, pp. 123-38.

Carroll, C. D. (2003). "Macroeconomic Expectations of Households and Professional Forecasters." *Quarterly Journal of Economics*:February.

CCE (1998). "Avis Concernant l'Incidence du Fonctionnement de l'Union Économique et Monétaire sur la Conduite de la Politique Économique Belge." Conseil Central de L'Economie: Bruxelles.

CfDT (2005). "Le SMIC: Mode d'Emploi."

Chan-Lee, J. H. (1980). "A Review of Recent Work in the Area of Inflationary Expectations." *Weltwirtschaftliches Archiv*, 116, pp. 45-86.

Clarida, R., J. Galí, and M. Gertler (1998). "Monetary Policy Rules and Macroeconomic Stability: Evidence and some Theory." *NBER Working Paper No.6442*.

Clavería González, O. (2003). "Cuantificación de las Expectativas de Precios a Partir de la Encuesta Industrial de la UE." *Tesis Doctoral, Departamento de Econometría, Estadística y Economía Española de la Universidad de Barcelona*.

Clemente, J., A. Montañés, and M. Reyes (1998). "Testing for a Unit Root in Variables with a Double Change in the Mean." *Economics Letters*, 59, pp. 175-82.

Clements, M. P. and D. F. Hendry (1999). *Forecasting Non-stationary Economic Time Series*.

Consensus Economics (1994-2006). "Consensus Forecasts."

Cooley, T. F. and E. C. Prescott (1973). "Varying Parameter Regression: A Theory and Some Applications." *Annals of Economic and Social Measurement*, 2, pp. 463-73.

Cooley, T. F. and E. C. Prescott (1976). "Estimation in the Presence of Stochastic Parameter Variation." *Econometrica*, 44, pp. 167-84.

Crawford, I. and Z. Smith (2002). "Distributional Aspects of Inflation." *Institute for Fiscal Studies (IFS) Commentary*, No 90.

CSERC (1999). "Le SMIC : Salaire Minimum de Croissance." *La Documentation française*.

Cukierman, A. (1986). "Measuring Inflation Expectations" *Journal of Monetary Economics*, 17, pp. 315-324

Curtin, R. (1996). "Procedure to Estimate Price Expectations, Mimeo," University of Michigan Survey Research Centre.

Curtin, R. (2006). "Inflation Expectations: Theoretical Models and Empirical Tests " *The Role of Inflation Expectations in Modelling and Monetary Policy Making*. National Bank of Poland: Warsaw.

Curto Millet, F. (2004). "The Impact of EMU on the Flexibility of European Labour Markets - Structural Break or 'Business as usual'?" *University of Oxford MPhil Thesis*.

Curto Millet, F. (2006). "Discussion of "Survey Expectations, Rationality and the Dynamics of Euro Area Inflation" by M. Forsells and G. Kenny." *The Role of Inflation Expectations in Modelling and Monetary Policy Making*. National Bank of Poland: Warsaw.

Dasgupta, S. and K. Lahiri (1992). "A Comparative Study of Alternative Methods of Quantifying Qualitative Survey Responses Using NAPM Data." *Journal of Business & Economic Statistics*, 10:4, pp. 391-400.

Debelle, Guy (2004). "Household Debt and the Macroeconomy." *BIS Quarterly Review*: March, pp. 51-64.

Defris, L. V. and R. A. Williams (1979). "Quantitative Versus Qualitative Measures of Price Expectations." *Economics Letters*, 2, pp. 169-73.

Dessy, O. (2002). "Nominal Wage Rigidity in the European Countries: Evidence from the Europanel." *10th International Conference on Panel Data, July 5-6, 2002*: Berlin.

DG ECFIN (2002). "Germany's Growth Performance in the 1990s." *European Economy*, 170: May.

DG ECFIN (European Commission) (2004). "The Joint Harmonised EU Programme of Business and Consumer Surveys, User Guide."

Dolado, J., F. Kramarz, S. Machin, A. Manning, D. Margolis, and C. Teulings (1996). "The Economic Impact of Minimum Wages in Europe." *Economic Policy*, 23.

Ebbinghaus, B. (2002). "Trade Unions' Changing Role: Membership Erosion, Organisational Reform, and Social Partnership in Europe." *Industrial Relations Journal*, 33:5, pp. 465-83.

Ebbinghaus, B. and J. Visser (2000). *Trade Unions in Western Europe since 1945*. London: Macmillan.

Edouard, L. and A. Senthilselvan (1997). "Observer Error and Birthweight: Digit Preference in Recording." *Public Health*, 111:2, pp. 77-79.

Eichengreen, B. (1998). "European Monetary Unification: A Tour d'Horizon." *Oxford Review of Economic Policy*.

Eichengreen, B. and C. Wyplocz (1993). "The Unstable EMS." *Brookings Papers on Economic Activity*, 1, pp. 51-124.

EIRO (1997-2004). "Annual Review, various issues."

EIRO (1997a). "Germany: Recent Debates on Pay Policy in Germany ".

EIRO (1997b). "Spain: Industrial Action in Public Administration ".

EIRO (1998a). "1998 Annual Review for Germany ".

EIRO (1998b). "Germany: Trade unions call for a U-turn in pay policy in 1999 ".

EIRO (1998c). "1998 Annual Review for Italy."

EIRO (1998d). "1998 Annual Review for France."

EIRO (1999a). "1999 Annual Review for France."

EIRO (1999b). "1999 Annual Review for Spain."

EIRO (1999c). "1999 Annual Review for Sweden."

EIRO (2000a). "Wage Policy and the EMU."

EIRO (2000b). "France: Le Retour des Revendications Salariales."

EIRO (2000c). "Industrial Relations in the EU, Japan and USA, 2000."

EIRO (2001a). "2001 Annual Review for Belgium."

EIRO (2001b). "2001 Annual Review for Spain."

EIRO (2001c). "Industrial Relations in the EU, Japan and USA, 2001."

EIRO (2001d). "Netherlands: 2001 bargaining round likely to be difficult ".

EIRO (2001e). "Spain: National Court Rules Against Freezing of Civil Servants' Pay ".

EIRO (2001f). "Netherlands: First Collective Agreements Reached for 2001 ".

EIRO (2002a). "Collective Bargaining Coverage and Extension Procedures."

EIRO (2002b). "Debate Resurfaces on Automatic Pay Indexation ".

EIRO (2003a). "Netherlands: Deadlock on Occupational Pensions in Company Bargaining ".

EIRO (2003b). "Netherlands: Occupational Pensions Are Prominent Issue in Bargaining ".

EIRO (2004a). "2004 Annual Review for the Netherlands ".

EIRO (2004b). "Netherlands: Bargaining under the central wage freeze ".

- EIRO (2004c). "Occupational Pensions and Industrial Relations ".
- EIRO (2004d). "Netherlands: Occupational Pension Issues Place Increasing Pressure On Industrial Relations."
- EIRO (2004e). "2004 Annual Review for Sweden ".
- EIRO (2005a). "Minimum wages in Europe ".
- EIRO (2005b). "Spain: Minimum Wage Agreement Introduces Guarantee Clause ".
- EIRO (2005c). "Spain: Minimum Pensions Increased ".
- EMIRE (2005). "Sliding-Scale Mechanism."
- Engsted, T. (1991). "A Note on the Rationality of Survey Inflation Expectations in the United Kingdom." *Applied Economics*, 23, pp. 1269-76.
- Erickson, C. and A. Ichino (1993). "Wage Differentials in Italy: Market Forces, Institutions and Inflation " *Working paper, n. 34, Innocenzo Gasparini Institute for Economic Research*.
- Ericsson, N. R. and J. S. Irons (1995). "The Lucas Critique in Practice: Theory Without Measurement," in *Macroeconometrics: Developments, Tensions and Prospects*. K. D. Hoover ed. Dordrecht: Kluwer Academic Press.
- European Central Bank (2003). "Structural Factors in the EU Housing Markets." March.
- European Commission (2004). "Industrial Relations in Europe 2004."
- European Commission (2005a). "VAT Rates Applied in the Member States of the European Community [DOC/1636/2005 - EN]."
- European Commission (2005b). *Quarterly Report on the Euro Area (4:4)*.
- Evans, G. W. and S. Honkapohja (2001). *Learning and Expectations in Macroeconomics*: Princeton University Press.
- Figlewski, S. and P. Wachtel (1981). "The Formation of Inflationary Expectations." *Review of Economics and Statistics*, 63:1, pp. 1-10.
- Figlewski, S. and P. Wachtel (1983). "Rational Expectations, Informational Efficiency, and Tests Using Survey Data: A Reply." *Review of Economics and Statistics*, 65, pp. 529-31.
- Flanagan, R. J. (1988). "Unemployment as a Hiring Problem." *OECD Economic Studies*, 11:Autumn, pp. 123-54.
- Forder, J. (1996). "On the assessment and implementation of 'institutional remedies'." *Oxford Economic Papers*, 48, pp. 39-51.
- Forder, J. (1998). "Central bank independence conceptual clarifications and interim assessment." *Oxford Economic Papers*, 50:307-334.
- Forder, J. (2000). "Could Reputation-Bias be a Bigger Problem than Inflation-Bias?" *Oxford University, Department of Economics Working Paper No22*.
- Forder, J. and C. Huhne (2001). *Both Sides of the Coin*: Profile Books.

Forsells, M. and G. Kenny (2006). "Survey Expectations, Rationality and the Dynamics of Euro Area Inflation." *The Role of Inflation Expectations in Modelling and Monetary Policy Making*. National Bank of Poland: Warsaw.

Forsells, M. and G. Kenny (2006). "Survey Expectations, Rationality and the Dynamics of Euro Area Inflation." *The Role of Inflation Expectations in Modelling and Monetary Policy Making*. National Bank of Poland: Warsaw.

Foster, J. and M. Gregory (1977). "Inflation Expectations: The Use of Qualitative Survey Data." *Applied Economics*, 9, pp. 319-29.

Frankel, J. A. and A. K. Rose (1998). "The Endogeneity of the Optimum Currency Area Criteria." *The Economic Journal*, 108:449, pp. 1009-25.

Fratianni, Michele and Jurgen von Hagen (1992). *The European Monetary System and the European Monetary Union*. Boulder: Westview Press.

Fregert, K. and L. Jonung (1998). "Monetary Regimes and Endogenous Wage Contracts: Sweden 1908-1995."

Friedman, M. (1957). *Theory of the Consumption Function*. Princeton: Princeton University Press.

Friedman, M. (1968). "The Role of Monetary Policy." *American Economic Review*, LVIII:1, pp. 1-17.

Fuhrer, J. C. and G. R. Moore (1995). "Inflation Persistence." *Quarterly Journal of Economics*, 110:February, pp. 127-59.

Galí, J. and M. Gertler (1999). "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics*, 44, pp. 195-222.

Gallup (1981-1990). *Gallup Political Index*.

Gallup (1991-1997). "Gallup Political and Economic Index."

Gerberding, C. (2001). "The Information Content of Survey Data on Expected Price Developments for Monetary Policy." *Deutsche Bundesbank ERC Discussion Paper 9/01*.

Gerberding, C. (2006). "Household versus expert forecasts of inflation: New evidence from European survey data." *The Role of Inflation Expectations in Modelling and Monetary Policy Making*. National Bank of Poland: Warsaw.

Gerlach, S. and F. Smets (1995). "The Monetary Transmission Mechanism: Evidence from the G-7 Countries." *BIS Discussion Paper*.

GfK (2003). *Consumer Survey 2003*: GfK Sverige AB.

Giavazzi, F. and M. Pagano (1988). "The Advantage of Tying One's Hands." *European Economic Review*, 32, pp. 1055-82.

Golden, Miriam, Peter Lange, and Michael Wallerstein (2002). "Union Centralization among Advanced Industrial Societies: An Empirical Study."

Goubert, L. and E. Omy (1996). "An Alternative Measure of Wage Flexibility." *International*

Advances in Economic Research, 2:3, pp. 199-221.

Gressani, D., L. Guiso, and I. Visco (1988). "Disinflation in Italy: An Analysis with the Econometric Model of the Bank of Italy." *Journal of Policy Modeling*, 10, pp. 163-203.

Guichard, S. and J-P. Laffargue (2000). "The Wage Curve : The Lessons of an Estimation over a Panel of Countries." *CEPII Document de Travail N°2000-21*.

Hallett, A. H. and N. Viegi (2000). "Labour Market Institutions and Monetary Policy in EMU: Do Asymmetries Matter?" *ENEPRI Workshop "EMU and Asymmetries in Labour Markets"*: Paris.

Hanssens, M. and P. M. Vanden Abeele (1987). "A Time-Series Study of the Formation and Predictive Performance of EEC Production Survey Expectations." *Journal of Business & Economic Statistics*, 5:4, pp. 507-19.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Heinemann, F. and K. Ullrich (2004). "The Impact of EMU on Inflation Expectations." *ZEW Discussion Paper No. 04-01*.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, D. F. (2001). "Modelling UK Inflation 1875-1991." *Journal of Applied Econometrics*, 16, pp. 255-75.

Hendry, D. F. (2002). "Macro-econometrics." *Lecture Notes, Oxford University, Mimeo*.

Hendry, D. F. (2004). "Economic Forecasting." *Lecture Notes, Oxford University, Mimeo*.

Hendry, D. F. and J. A. Doornik (2001). *Empirical Econometric Modelling Using PcGive*. London: Timberlake Consultants Press.

Hendry, D. F. and K. Juselius (2000). "Explaining Cointegration Analysis: Part 1." *The Energy Journal*, 21:1, pp. 1-42.

Hendry, D. F. and K. Juselius (2001). "Explaining Cointegration Analysis: Part 2." *The Energy Journal*, 22:1, pp. 75-120.

Hendry, D. F. and H.-M. Krolzig (2001). *Automatic Econometric Model Selection Using PcGets*: Timberlake Consultants Press.

Hicks, J. (1974). *The Crisis in Keynesian Economics*. Oxford: Basil Blackwell.

HM Treasury (2002). "Towards Full Employment in the European Union."

HM Treasury (2003). "EMU and Labour Market Flexibility." HM Stationary Office.

Hochreiter, E. and G. Winckler (1995). "The Advantages of Tying Austria's Hands: The Success of the Hard Currency Strategy." *European Journal of Political Economy*, 11:1.

Hodrick, R. J. and E. C. Prescott (1980). "Post-war U.S. Business Cycles: An Empirical Investigation." *Discussion Paper No. 451, Carnegie-Mellon University*.

Hodrick, R. J. and E. C. Prescott (1997). "Post-war U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit and Banking*, 29:1, pp. 1-16.

- Holden, S. (1994). "Wage Bargaining and Nominal Rigidities." *European Economic Review*, 38.
- ILO (1997). "World Labour Report 1997-98: Industrial Relations, Democracy and Social Stability."
- ILO (2005a). "Social Pacts in Italy: Pact on the Termination of the Cost of Living Index Mechanism."
- ILO (2005b). "Social Pacts in the Netherlands."
- INSEE (2004). "Enquête Mensuelle de Conjoncture Auprès des Ménages (Note Méthodologique)."
- Jonung, L. and F. Sjöholm (1999). "Should Finland and Sweden Form a Monetary Union?" *The World Economy*, 22:5, pp. 683-700.
- Kaitz, H. (1970). "Experience of the Past: The National Minimum Wage," in *Youth Unemployment and Minimum Wages, Bulletin 1657*. US Department of Labor (Bureau of Labor Statistics) ed.
- Keynes, J. M. (1936). *The General Theory of Employment, Interest, and Money*. London: Macmillan.
- Kokoszcyński, Ryszard, Tomasz Łyziak, and Ewa Stanisławska (2006). "Consumer Inflation Expectations. Usefulness of Survey-Based Measures - A Cross-Country Study." *The Role of Inflation Expectations in Modelling and Monetary Policy Making*. National Bank of Poland: Warsaw.
- Konjunkturinstitutet "Hushållens Inköpsplaner (Consumer Survey, User Manual)."
- Koopman, S. J., A. C. Harvey, J. A. Doornik, and N. Shephard (1999). *STAMP Structural Time Series Analyser, Modeller and Predictor*. London: Timberlake Consultants Press.
- Kose, A., E. S. Prasad, and M.E. Terrones (2003). "Financial Integration and Macroeconomic Volatility." *IMF Staff Papers*, 50:Special Issue.
- Koubi, Malik, Stéphane Mussard, Françoise Seyte, and Michel Terraza (2003). "La hiérarchie des salaires dans le secteur privé: permanences et évolutions." *Working Paper 2003-21, LAMETA, Université de Montpellier 1*.
- Kremers, J. J. M. (1990). "Gaining Policy Credibility for a Disinflation." *IMF Staff Papers*, 37 (Mar.), pp. 116-45.
- Krugman, P. (1993). "'Lessons of Massachusetts for EMU'," in *Adjustment and growth in the European Monetary Union*. F. Torres and F. Giavazzi eds: Cambridge University Press (for CEPR, London).
- Kydland, F. and E. Prescott (1977). "Rules Rather Than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy*, 85, pp. 473-91.
- Labhard, V. and C. Wyplocz (1996). "The New EMS: Narrow Bands inside Deep Bands." *American Economic Review*, 86:2, pp. 143-46.
- Lagos, R. A. (1994). "Labour Market Flexibility: What Does It Really Mean?" *CEPAL Review*, 54:December, pp. 81-95.
- Lauer, C. (1999). "The Effects of European Economic and Monetary Union on Wage Behaviour."

Layard, R., S. Nickell, and R. Jackman (1991). *Unemployment Macroeconomic Performance and the Labour Market*.

Lazear, E. (1990). "Job Security Provision and Employment." *Quarterly Journal of Economics*.

Lindén, S. (2006). "400 000 Observations on Inflation Perceptions and Expectations in the EU. What Will They Tell Us?" *The Role of Inflation Expectations in Modelling and Monetary Policy Making*. National Bank of Poland: Warsaw.

Liu, J., S. Wu, and J.V. Zidek (1997). "On Segmented Multivariate Regressions." *Statistica Sinica*, 7, pp. 497-525.

Lo Cascio, I. (2001). "Do Labour Markets Really Matter? Monetary Shocks and Asymmetric Effects Across Europe." *Ente per gli Studi Monetari, Bancari e Finanziari "Luigi Einaudi"*.

Lucas, R. (1972). "Expectations and the Neutrality of Money." *Journal of Economic Theory*, 4, pp. 103-24.

Lucas, R. (1973). "Some International Evidence on Inflation-Output TradeOffs." *American Economic Review*, LXIII, pp. 326-34.

Lucas, R. (1976). "Econometric Policy Evaluation: A Critique," in *The Phillips Curve and Labour Markets*. K. Brunner and A. H. Meltzer eds: Carnegie-Rochester Conference Series on Public Policy, No1.

Lundborg, P. and H. Sacklén (2001). "Is There a Long Run Unemployment-Inflation Trade-off in Sweden?" *FIEF Working Paper No. 173*.

Łyziak, Tomasz (2003). "Consumer Inflation Expectations in Poland." *ECB Working Paper No. 287*.

MacLennan, D., J. Muellbauer, and M. Stephens (1998). "Asymmetries in Housing and Financial Market Institutions and EMU." *Oxford Review of Economic Policy*, 14:3.

Mankiw, N. G. (2000). "The Inexorable and Mysterious Tradeoff Between Inflation and Unemployment." *NBER Working Paper N°7884*.

Mankiw, N. G. and R. Reis (2002). "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics*: November.

Mankiw, N. G., R. Reis, and J. Wolfers (2003). "Disagreement about Inflation Expectations." *NBER Working Paper No. 9796*.

Mash, R. (2004). "Optimising Microfoundations for Inflation Persistence." *Oxford University, Department of Economics Working Paper No183*.

Mavroeidis, S. (2002). *Econometric Issues in Forward-Looking Monetary Models*: DPhil Thesis, Oxford University.

McFadden, D., N. Schwarz, and J. Winter (2004). "Measuring Perceptions and Behavior in Household Surveys."

Mehra, Y. P. (2002). "Survey Measures of Expected Inflation: Revisiting the Issues of Predictive Content and Rationality." *Federal Reserve Bank of Richmond Economic Quarterly*, 88:3, pp. 17-

Menon, G., O. Raghurir, and N. Schwarz (1997). "How Much Will I Spend? Factors Affecting Consumers' Estimates of Future Expense." *Journal of Consumer Psychology*, 6:2, pp. 141-64.

Mermet, E. (2001). *Wage Formation in Europe*. Brussels: ETUI.

Miles, David (2004). "The Miles Review . The UK Mortgage Market: Taking a longer-term view." HM Treasury.

Ministère de l'Emploi et de la Solidarité (2000). *La Négociation Collective en 1999, Tome I: La Tendence, Tome II: Chiffres et Documents, Tome III: Les Dossiers*. Paris: Éditions législatives.

Moggridge, D. E. (1992). *Maynard Keynes. An Economist's Biography*. London and New York: Routledge.

Moghadam, R. and S. Wren-Lewis (1994). "Are Wages Forward-Looking?" *Oxford Economic Papers*, 46, pp. 403-24.

Mundell, R. A. (1961). "A Theory of Optimum Currency Areas." *American Economic Review*, 51, pp. 657-65.

Muth, J. F. (1961). "Rational Expectations and the Theory of Price Movements." *Econometrica*.

Nardo, M. (2003). "The Quantification of Qualitative Survey Data: A Critical Assessment." *Journal of Economic Surveys*, 17:5, pp. 645-68.

Nerlove, M. (1958). "Adaptive Expectations and Cobweb Phenomena." *Quarterly Journal of Economics*, 72, pp. 227-40.

Nickell, S. (1997). "Unemployment and Labor Market Rigidities: Europe versus North America." *Journal of Economic Perspectives*, 11:3, pp. 55-74.

Nickell, S. (2003). "Employment and Taxes." *Tax Policy and Employment Conference, CESifo*.

Nickell, S. and L. Nunziata (2001). "Labour Market Institutions Database."

Nickell, S., L. Nunziata, and W. Ochel (2005). "Unemployment in the OECD Since the 1960s. What Do We Know?" *The Economic Journal*, 115:January, pp. 1-27.

Nicoletti, G., S. Scarpetta, and O. Boylaud (2001). "Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation." *OECD Economics Department Working Paper No. 226*.

Nielsen, H. (2003a). "Inflation Expectations in the EU - Results from Survey Data." *Discussion Paper No. 13/2003: Sonderforschungsbereich 373, Humboldt Universität zu Berlin*.

Nielsen, H. (2003b). "Inflation Expectations in the EU - rational?" *Unpublished Working Paper (VIIIth Spring Meeting of Young Economists, Katholieke Universiteit Leuven)*.

Nolte, I. and W. Pohlmeier (2004). "Using Forecasts of Forecasters to Forecast."

Nordhaus, W. (1975). "The Political Business Cycle." *Review of Economic Studies*, 42, pp. 169-90.

- Nunziata, L. (2001a). "Institutions and Wage Determination: A Multi-Country Approach." *Nuffield College Working Papers in Economics*:2001-W29.
- Nunziata, L. (2001b). "Labour Market Institutions and the Cyclical Dynamics of Employment." *Nuffield College Economics Working Papers Ref: 1999-W26*.
- Obstfeld, M. and G. Peri (1998). "Regional Nonadjustment and Fiscal Policy: Lessons for EMU." *NBER Working Paper No. W6431*.
- Ochel, W. (2000). "Collective Bargaining (Centralization/Co-ordination)." *Ifo Institute*.
- Ochel, W. (2001). "Collective Bargaining Coverage in the OECD from the 1960s to the 1990s." *CESifo Forum*, 2:4, pp. 62-65.
- OECD (1997). "Employment Outlook."
- OECD (2002a). "Economic Survey of the Euro Area." OECD: Paris.
- OECD (2002b). "Benefits and Wages." OECD: Paris.
- OECD (2004a). "OECD Employment Outlook 2004."
- OECD (2004b). "Benefits and Wages." OECD: Paris.
- OECD (2005). "The OECD Labour Market Statistics Database."
- Osgood, C. E. (1953). *Method and Theory in Experimental Psychology*. New York: Oxford University Press.
- Oswald, A. J. (1996). "A Conjecture on the Explanation for High Unemployment in the Industrialised Nations. ." *Warwick Economic Research Papers*, 475.
- Oswald, A. J. and P. Turnbull (1985). "Pay and Employment Determination in Britain: What Are Labour Contracts Really Like?" *Oxford Review of Economic Policy*, 1:Summer, pp. 80-97.
- Pagan, A. (1984). "Econometric Issues in the Analysis of Regressions with Generated Regressors." *International Economic Review*, 25, pp. 65-76.
- Palmqvist, S. and L. Strömberg (2004). "Households' Inflation Opinions - a Tale of Two Surveys." *Sveriges Riksbank Economic Review*, 4, pp. 23-42.
- Papadia, F. (1983). "Rationality of Inflationary Expectations in the European Economic Communities Countries." *Empirical Economics*, 8, pp. 187-202.
- Papadia, F. and V. Basano (1981). "EEC-DG II Inflationary Expectations. Survey Based Inflationary Expectations for the EEC Countries." *DG ECFIN Economic Papers*.
- Pérez, Sofía A. (2000). "Collective Bargaining Under EMU: Lessons From the Italian and Spanish Experiences." *Working paper No 72, Harvard Center for European Studies*.
- Perron, P. (1997). "L'Estimation de Modèles avec Changements Structurels Multiples." *Actualité Economique*, 73, pp. 457-505.
- Pesaran, M. H. (1984). "Expectations Formation and Macroeconometric Modelling," in *Contemporary Macroeconomic Modelling*. P. Malgrange and P. A. Muet eds. Oxford: Blackwell.

- Pesaran, M. H. (1985). "Formation of Inflation Expectations in British Manufacturing Industries." *Economic Journal*, 95:December 1985, pp. 948-75.
- Pesaran, M. H. (1987). *The limits to rational expectations*. Oxford, UK: Basic Blackwell.
- Pesaran, M. H., Y. Shin, and R. Smith (1999). "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels." *Journal of the American Statistical Association*, 94:446, pp. 621-34.
- Pesaran, M. H. and R. Smith (1995). "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels." *Journal of Econometrics*, 68, pp. 79-113.
- Pesaran, M. H. and A. Timmermann (1990). "A Simple, Non-Parametric Test of Predictive Performance." *Cambridge Working Papers in Economics* 9021.
- Phelps, E. (1990). *Seven Schools of Macroeconomic Thought*. Oxford: Oxford University Press.
- Reckwerth, J. (1997). "Inflation and Output in Germany: The Role of Inflation Expectations." *Bundesbank Discussion Paper No.5*.
- Reis, R. (2005). "Inattentive Producers." *NBER Working Paper No. 11820*.
- Roberts, J. M. (1997). "Is Inflation Sticky?" *Journal of Monetary Economics*, 39, pp. 173-96.
- Roskelley, K. (2003). "Macroeconomic Volatility: Does Financial Market Development Matter?"
- Ruehmann, P. (2002). "European Monetary Union (EMU) and National Labour Markets." *ERASMUS-SOKRATES Intensive Program "European Integration. Ten Years After the Rio Summit"*. Sept. 2 nd –13 th 2002 in Coimbra, Portugal.
- Russo, G. and D. Veredas (2000). "Institutional Rigidities and Employment Rigidity in the Italian Large Industrial Firms." *CORE Discussion Paper, 2000/48*. Université Catholique de Louvain.
- Ruud, J. and K. Whelan (2001). "New Tests of the New Keynesian Phillips Curve." *Research and Statistics*, D.P.30:Federal Reserve Board of Governors.
- Sargent, T. J. (1973). "Rational Expectations, the Real Rate of Interest and the Natural Rate of Unemployment." *Brookings Papers on Economic Activity*, 2, pp. 429-72.
- Scheremet, W. (2000). "On the Convergence of Wage Policy After EMU." *ENEPRI Workshop "EMU and Asymmetries in Labour Markets"*. Paris.
- Schmitter, P.C. (1979). "Still the Century of Corporatism?," in *Trends Toward Corporatist Intermediation*. P.C. Schmitter and G. Lehmbruch eds. London: Sage, pp. 7-52.
- Schwarz, N., R. M. Groves, and H. Schuman (1998). "Survey Methods," in *Handbook of Social Psychology, 4th Edition*. N. Schwarz, R. M. Groves and H. Schuman eds. New York: McGraw-Hill, pp. 143-79.
- Schwarz, N., H. J. Hippler, B. Deutsch, and F. Strack (1985). "Response Categories: Effects on Behavioral Reports and Comparative Judgements." *Public opinion Quarterly*, 49, pp. 388-95.
- Seitz, H. (1988). "the Estimation of Inflation Forecasts from Business Survey Data." *Applied Economics*, 20, pp. 427-38.
- Sekine, T. (2001). "Modeling and Forecasting Inflation in Japan." *IMF Working Paper No. 01/82*.

Shapiro, C. and J. Stiglitz (1984). "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review*.

Siebert, H. (1997). "Labor Market Rigidities: At the Root of Unemployment in Europe." *Journal of Economic Perspectives*, 11:3, pp. 37-54.

Smith, J. and M. McAleer (1995). "Alternative Procedures For Converting Qualitative Response Data to Quantitative Expectations: An Application to Australian Manufacturing." *Journal of Applied Econometrics*, 10, pp. 165-85.

Soskice, D. (1990). "Wage Determination: The Changing Role of Institutions in Advanced Industrialised Countries." *Oxford Review of Economic Policy*, 6:4.

Sudman, S., N. M. Bradburn, and N. Schwarz (1996). *Thinking about Answers*. San Francisco, CA: Jossey-Bass.

Sveriges Riksbank (2003). "Deflation - An Outline of the Problems." *Inflation Report*, 3.

Swedish Institute (2004). "Social insurance in Sweden."

Swedish Institute (2005). "Labour Relations in Sweden."

Taylor, J. B. (1979). "Staggered Wage Setting in a Macro Model." *American Economic Review*, 69:2, pp. 108-13.

Taylor, J. B. (1980). "Aggregate Dynamics and Staggered Contracts." *Journal of Political Economy*, 88:LXXXVIII, pp. 1-23.

Theil, H. (1952). "On the Time Shape of Economic Microvariables and the Munich Business Test." *Revue de l'Institut International de Statistique*, 20, pp. 105-20.

Thomas, L. B. (1999). "Survey Measures of Expected US Inflation." *Journal of Economic Perspectives*, 13:4, pp. 125-44.

Traxler, F. (1994). "Collective Bargaining: Levels and Coverage," in *OECD Employment Outlook 1994*. OECD ed. Paris.

Traxler, F., S. Blaschke, and B. Kittel (2001). *National Labour Relations in Internationalised Markets. A Comparative Study of Institutions, Change and Performance*. Oxford: Oxford University Press.

Traxler, F. and B. Kittel (2000). "The Bargaining System and Performance: A Comparison of 18 OECD Countries." *Comparative Political Studies*, 33:9, pp. 1154-90.

van Lelyveld, I. P. P. (2000). "Inflation, Institutions and Preferences." *PhD Thesis, University of Nijmegen*.

Vandaele, Kurt and Tine Boucké (2003). "Conflict Resolution By Figures? A Neo-Institutional Approach to the Changing Mechanism of Wage Indexation in Belgium."

Vilmunen, J. and H. Laakkonen (2004). "How often do prices change in Finland ? Micro-level evidence from the CPI." *European Central Bank, Inflation Persistence Network*.

Viñals, J. and J. Jimeno (1996). "Monetary Union and European Unemployment." *CEPR Discussion Paper No. 1485*.

Visco, I. (1984). *Price Expectations in Rising Inflation*. Amsterdam: North Holland.

Weiss, A. (1980). "Job Queues and Layoffs in Labour Markets with Flexible Wages." *Journal of Political Economy*.

Winter, J. (2002). "Bracketing Effects in Categorized Survey Questions and the Measurement of Economic Quantities," in *Discussion Paper No. 02-35, Sonderforschungsbereich 504*: University of Mannheim.

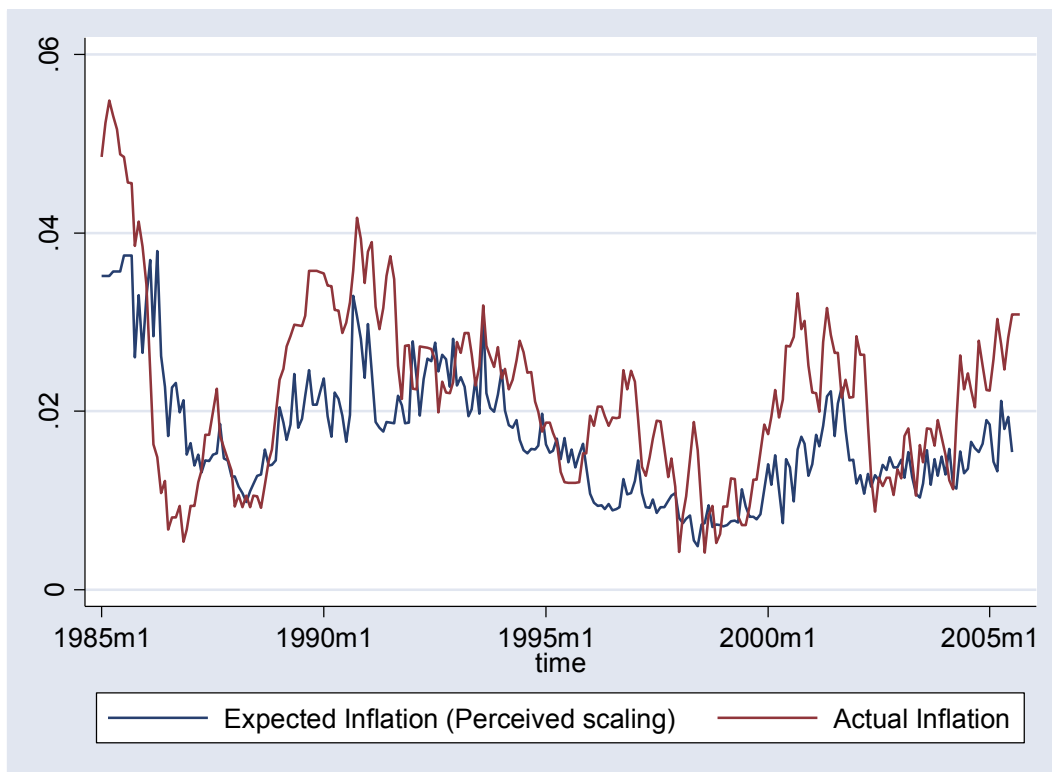
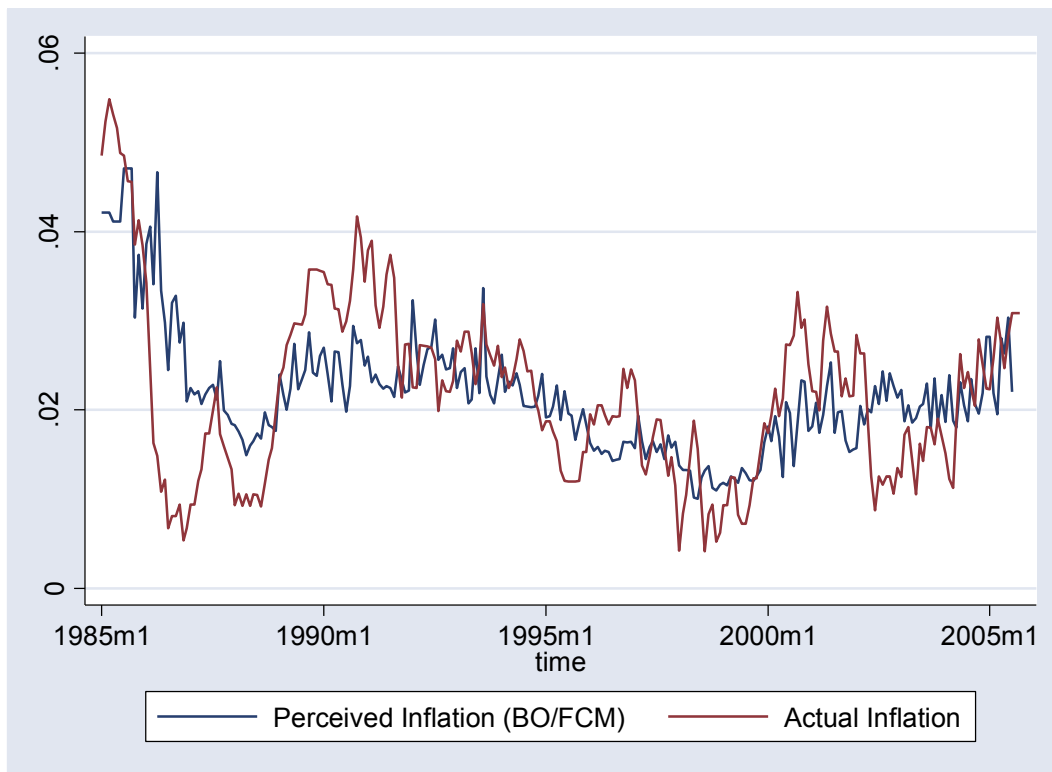
Young, D. (2003). "Employment Protection Legislation: Its Economic Impact and the Case for Reform." *European Economy, Economic Papers*:N°186.

APPENDIX I

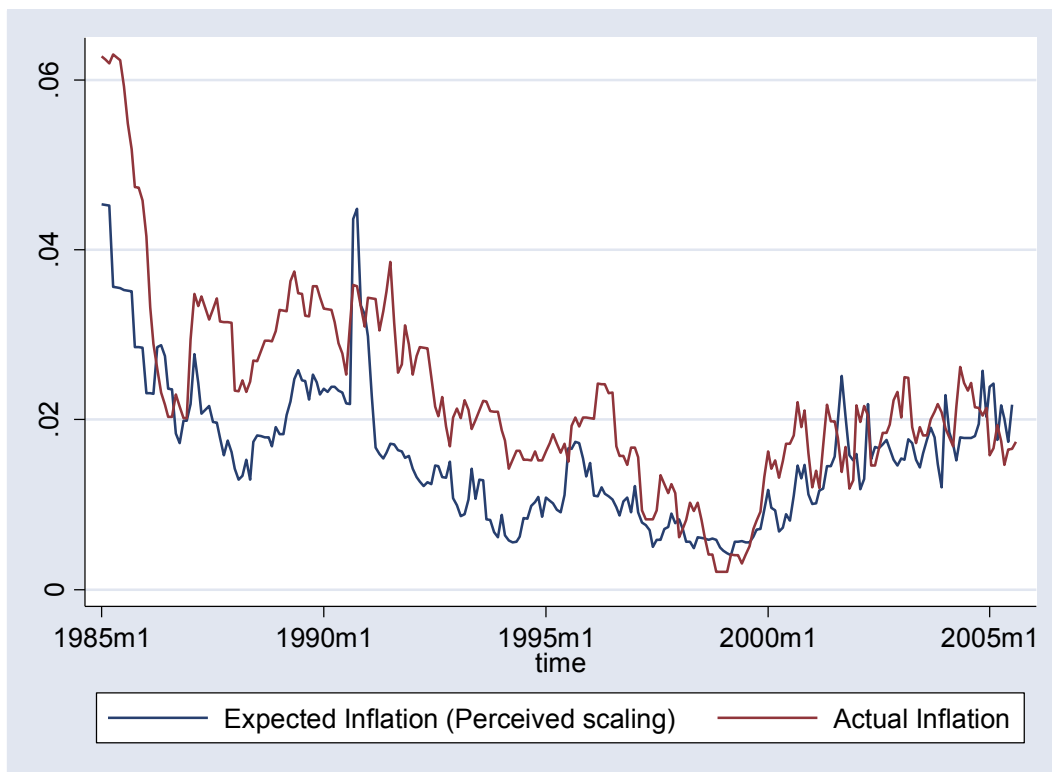
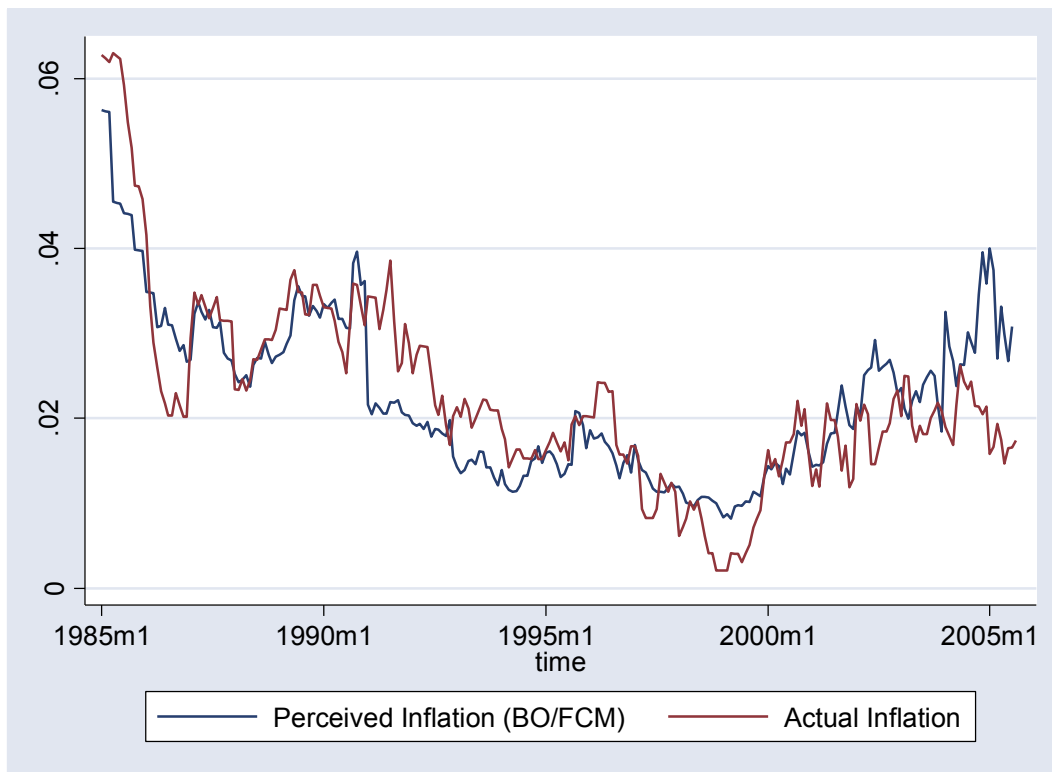
PLOTS OF INFLATION EXPECTATIONS AND PERCEPTIONS DATA

IMPORTANT NOTE: All the inflation expectations series are plotted at the time of their *formation* rather than at the time they *refer to* (+12 months).

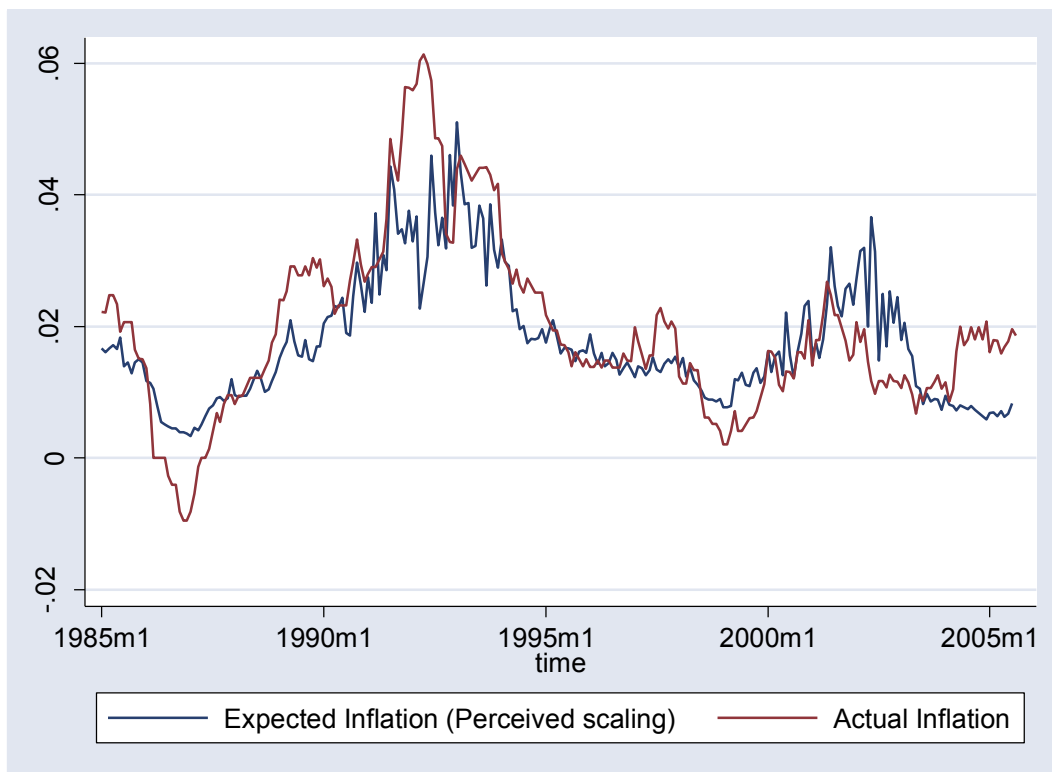
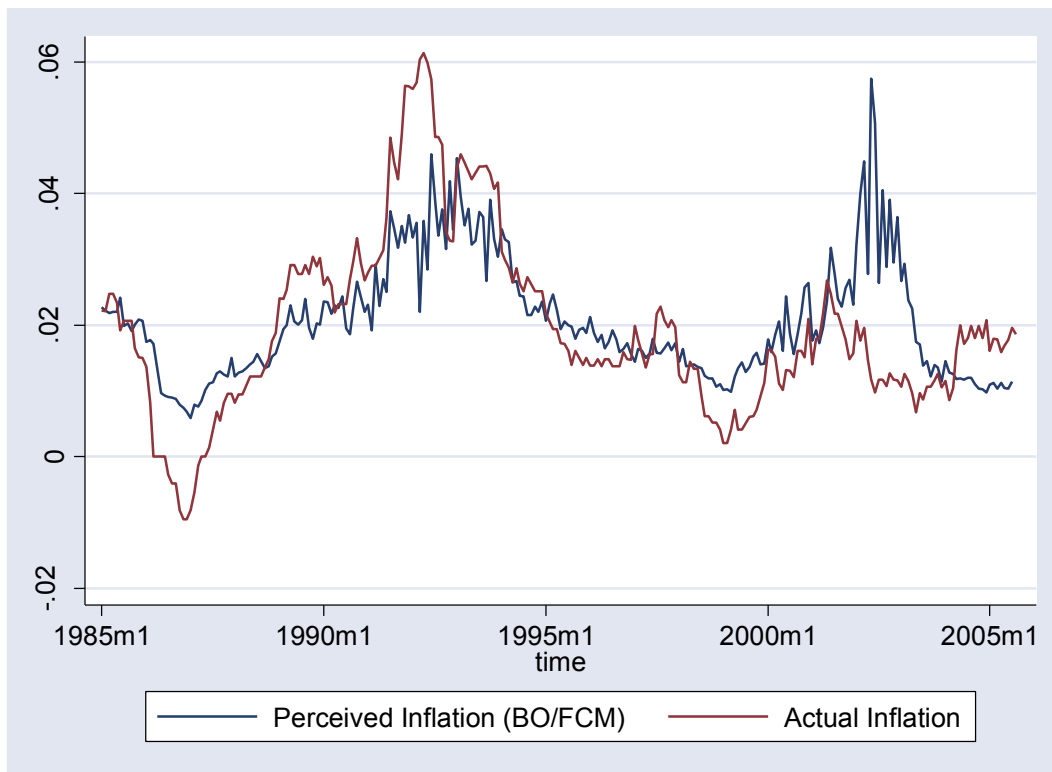
BELGIUM



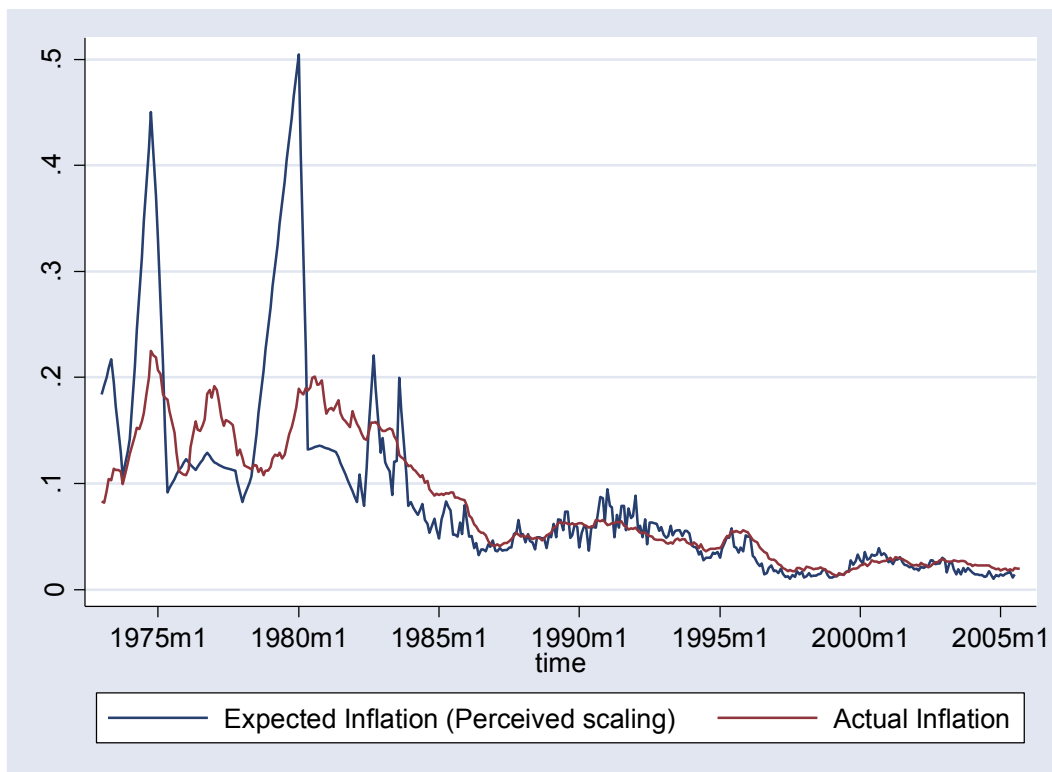
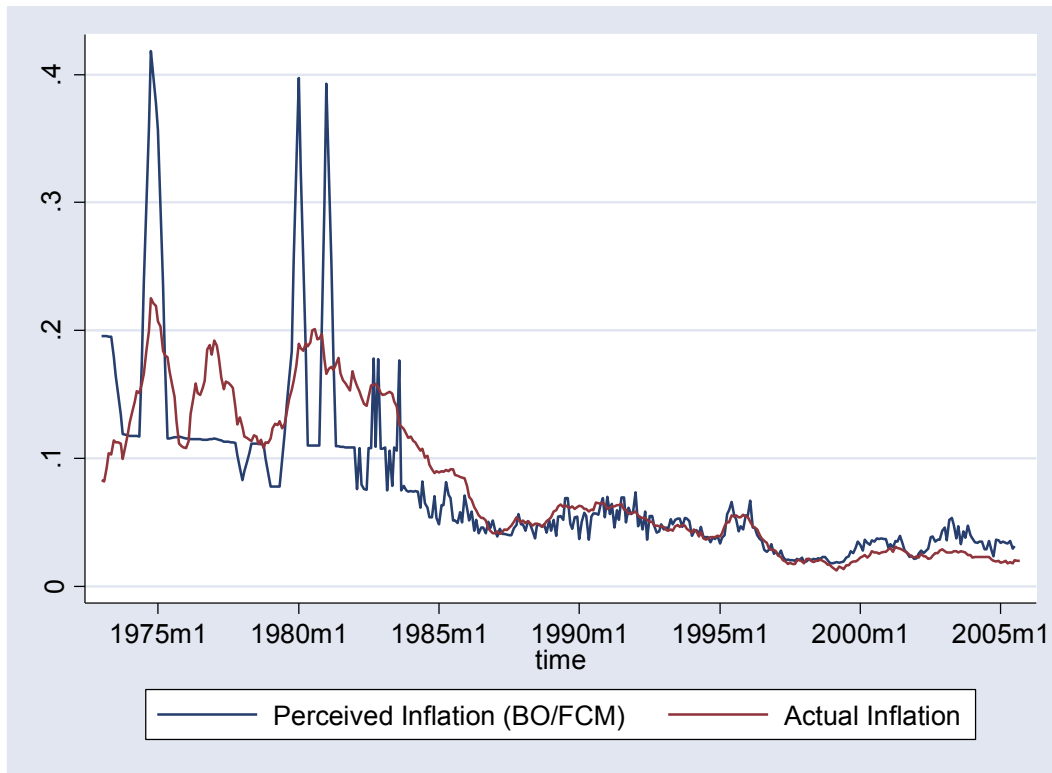
FRANCE



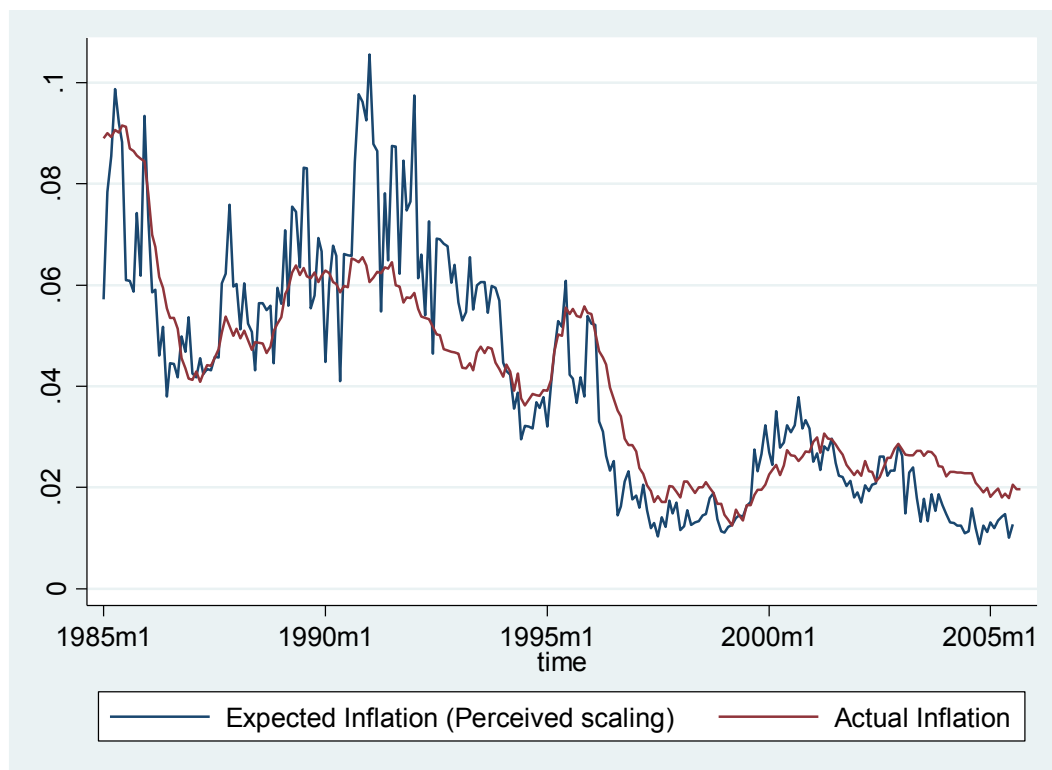
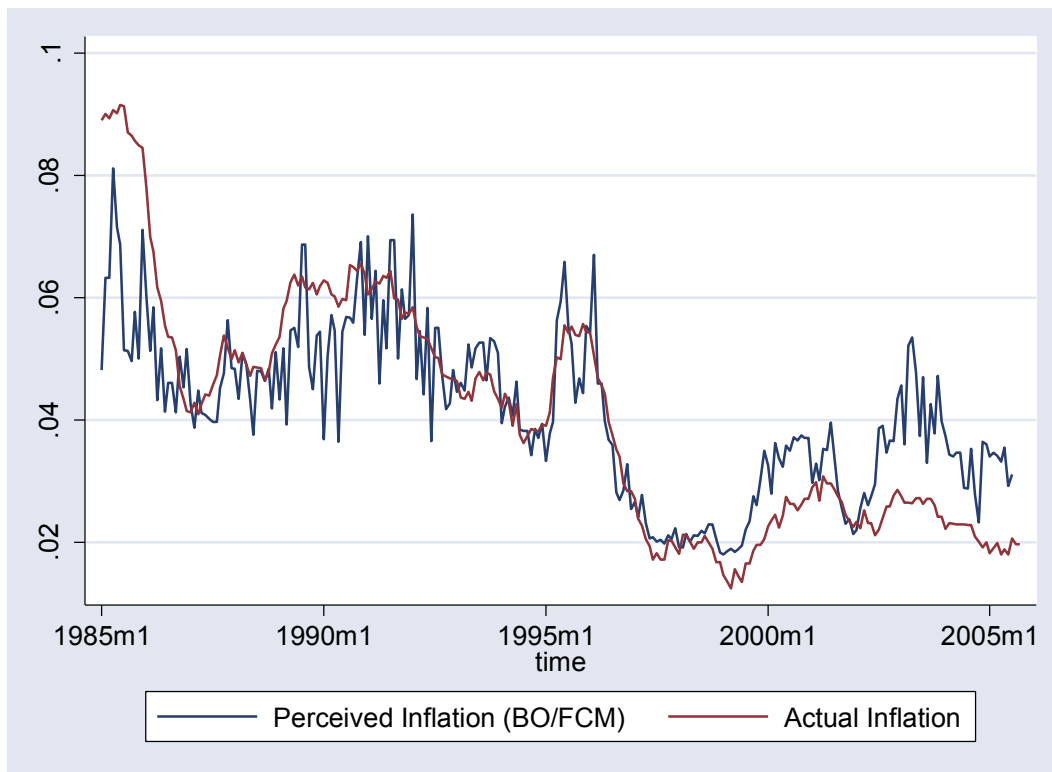
GERMANY



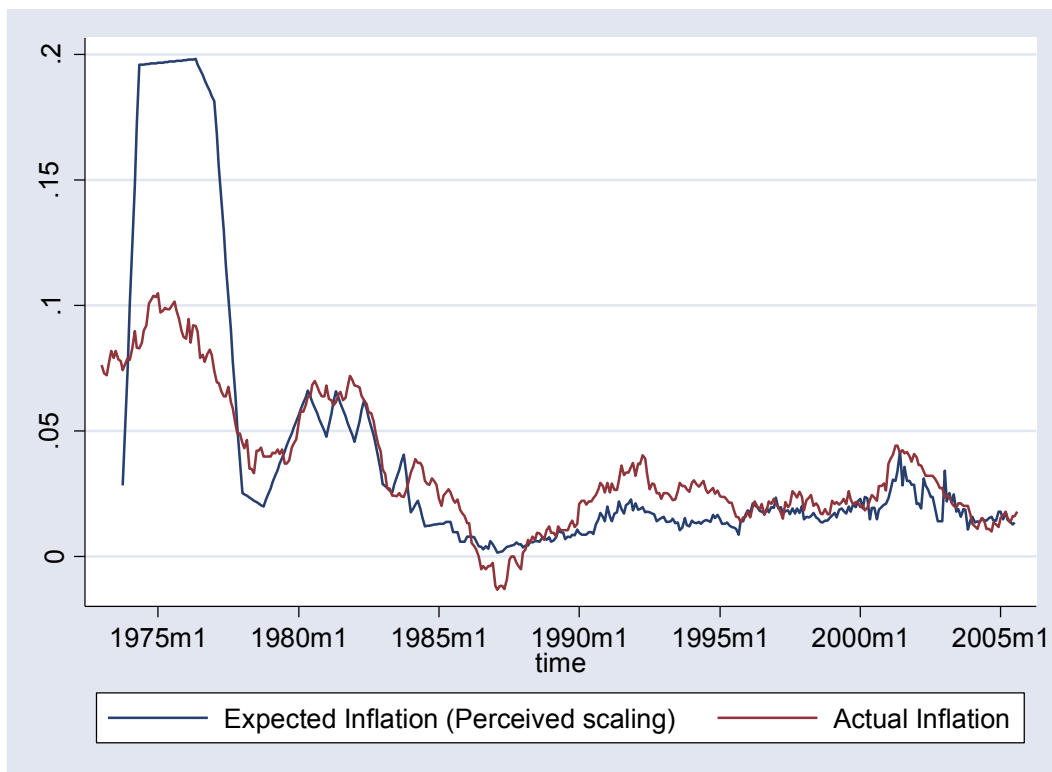
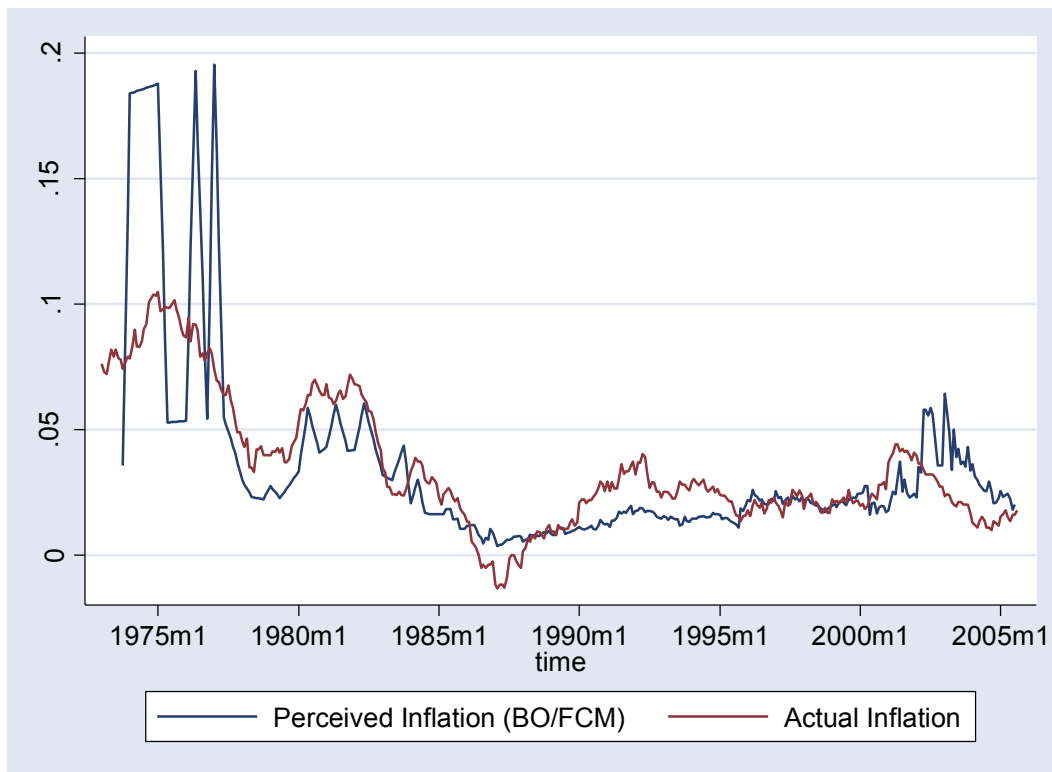
ITALY



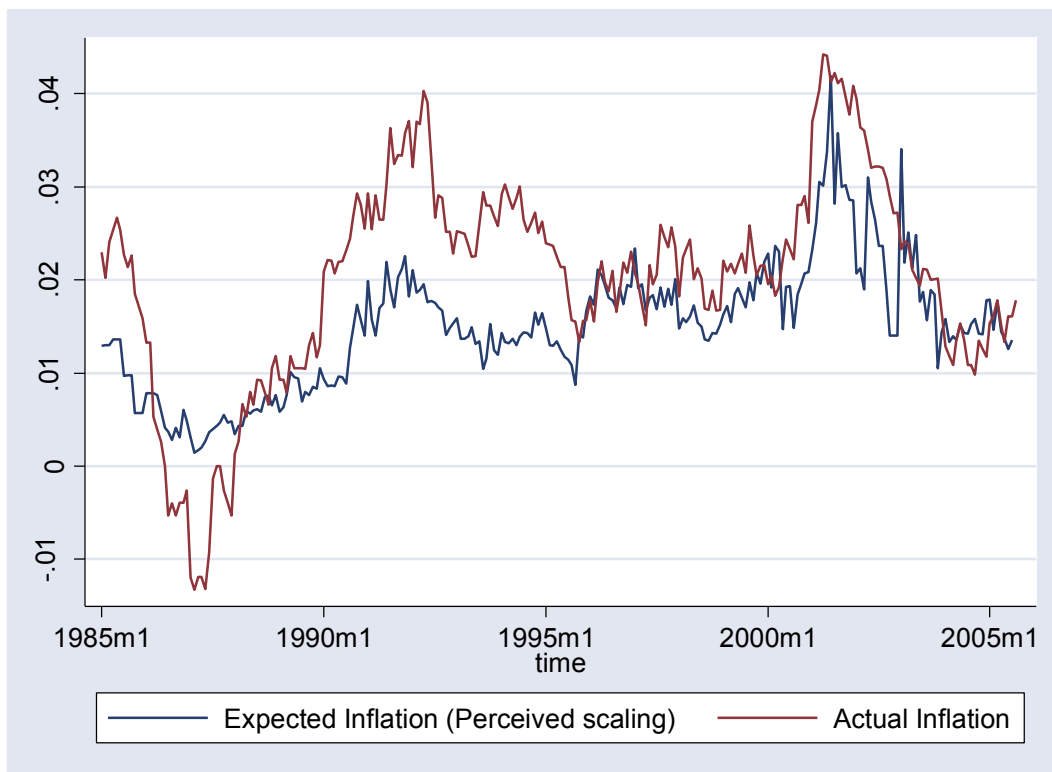
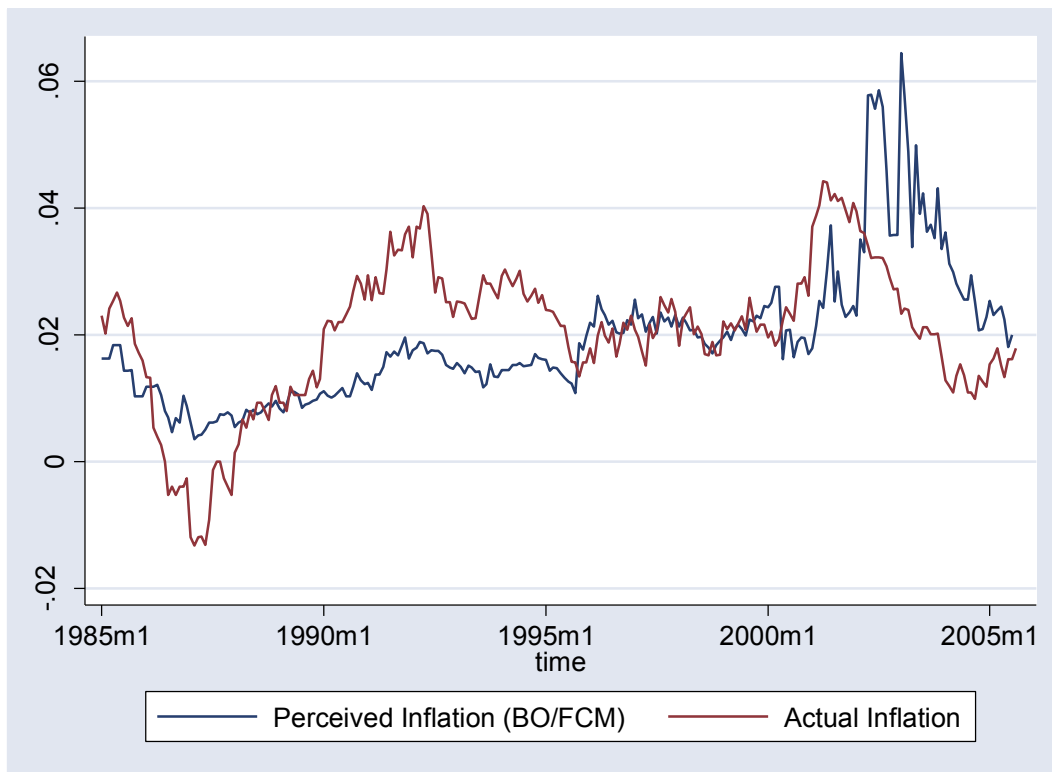
ITALY: PERIOD POST-1985



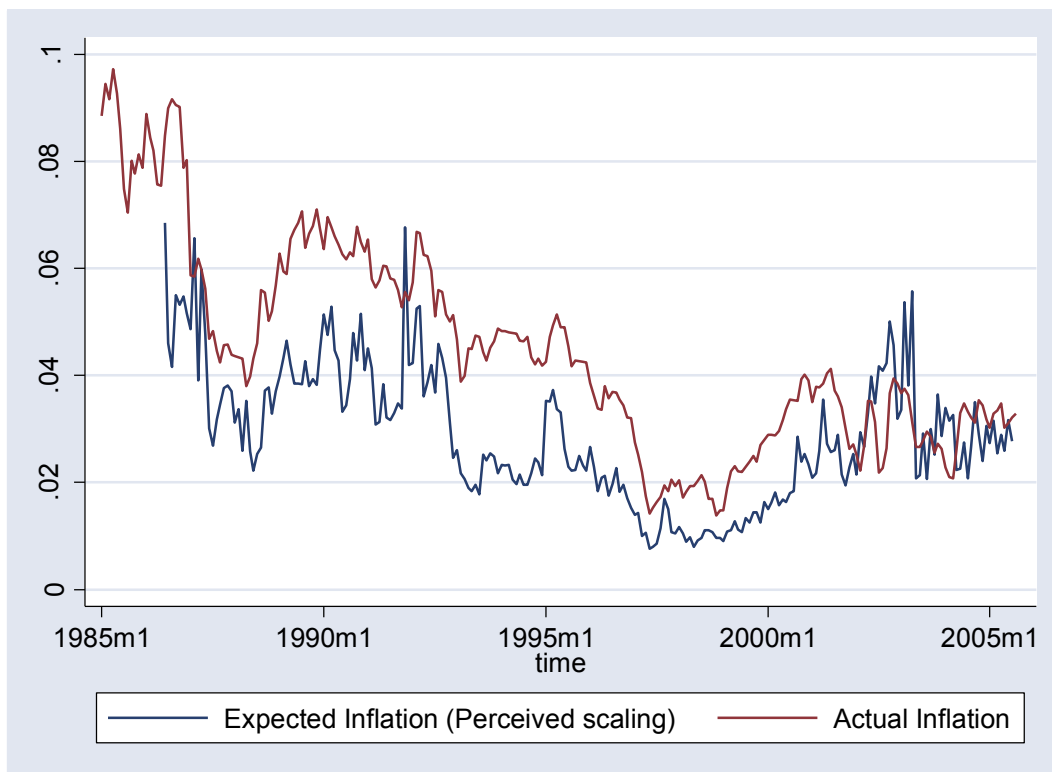
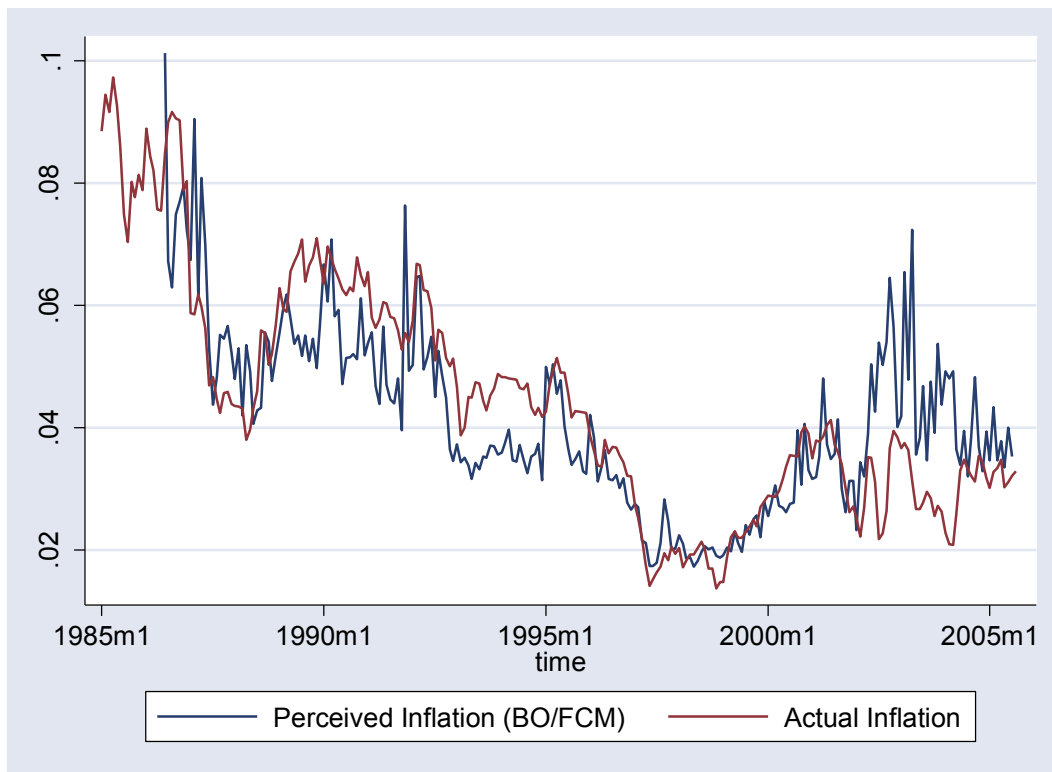
THE NETHERLANDS



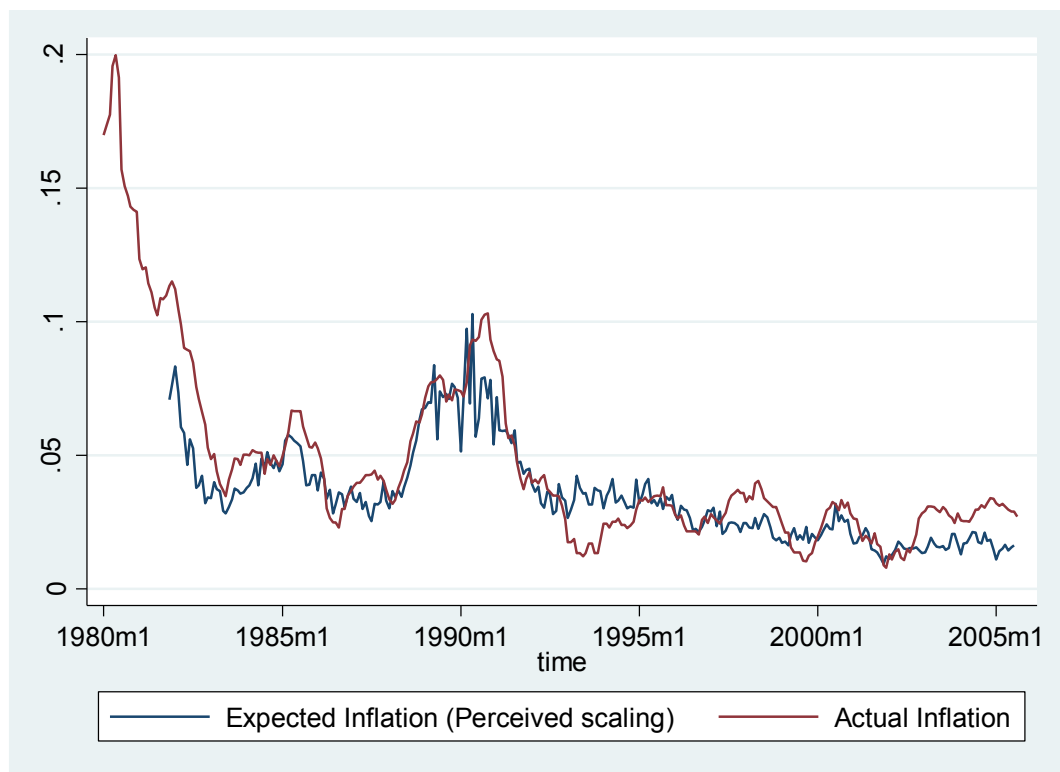
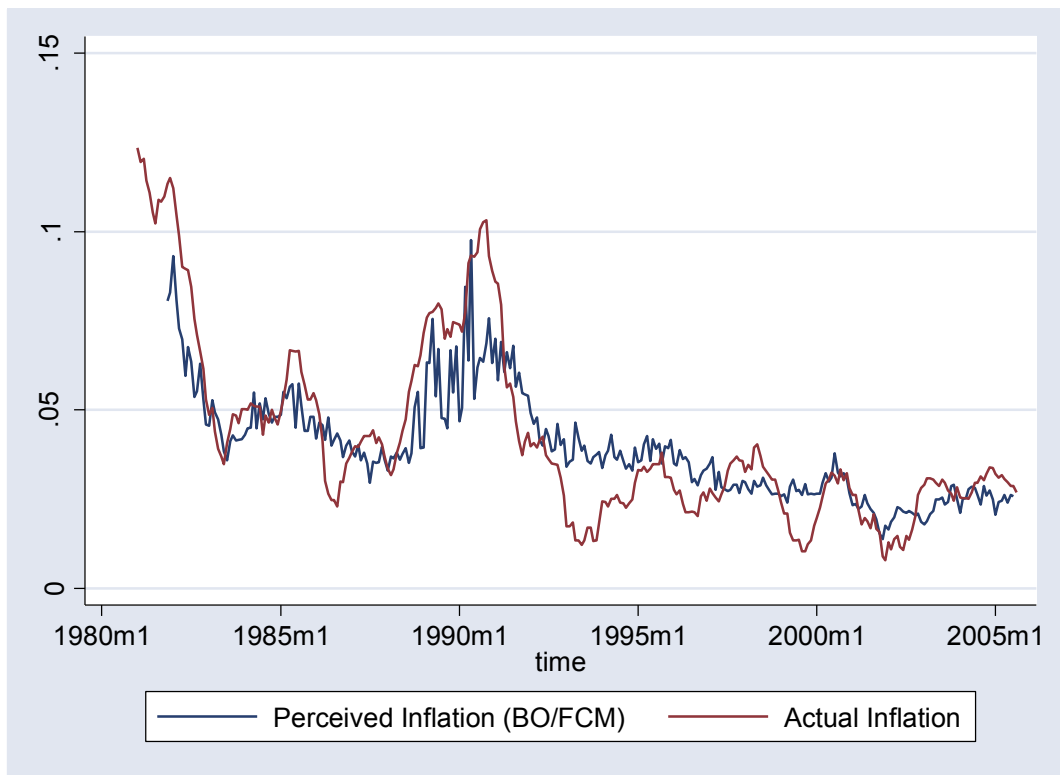
THE NETHERLANDS: PERIOD POST-1985



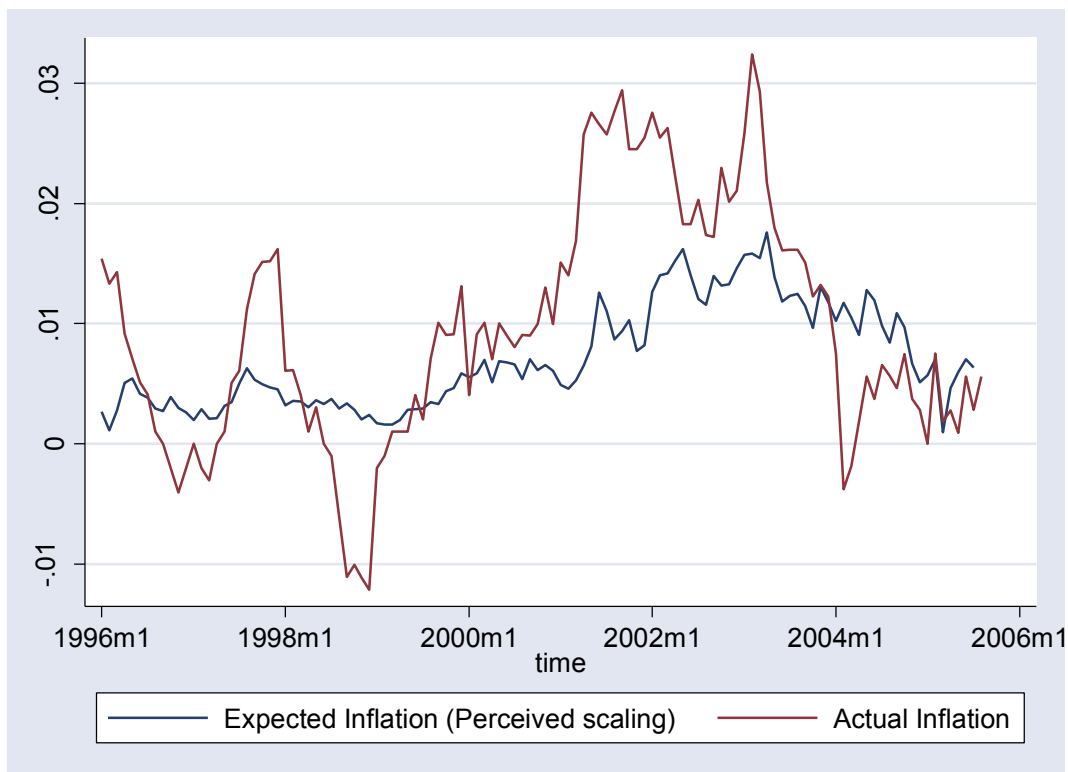
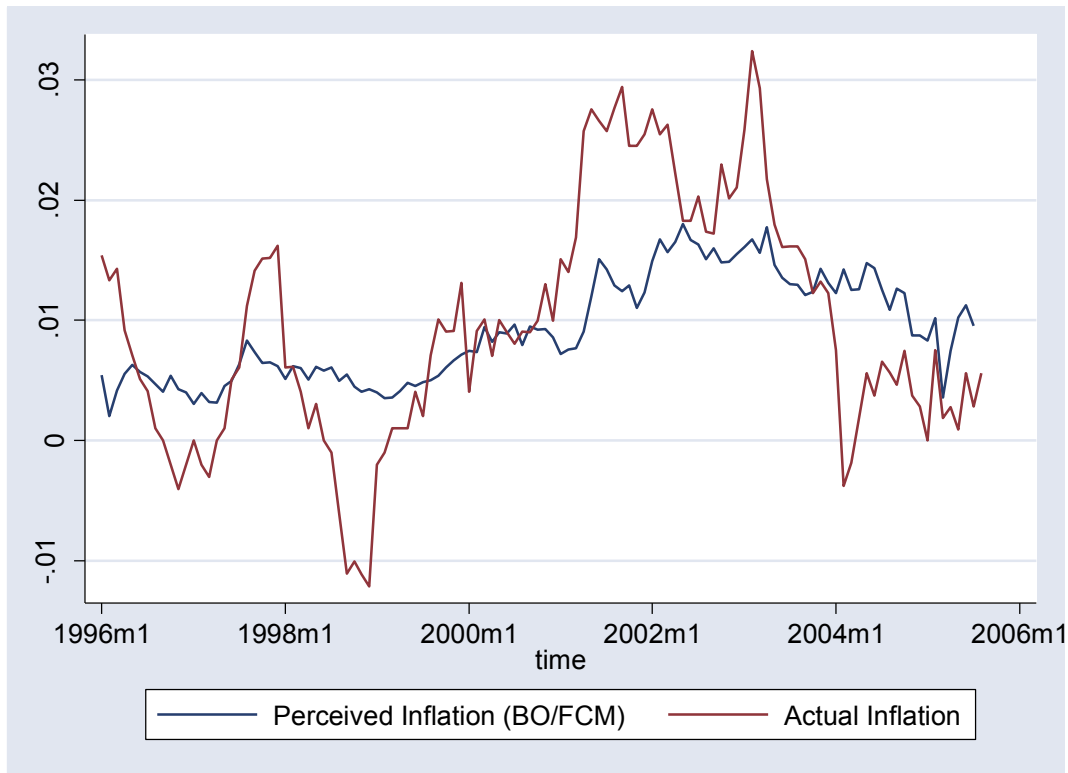
SPAIN



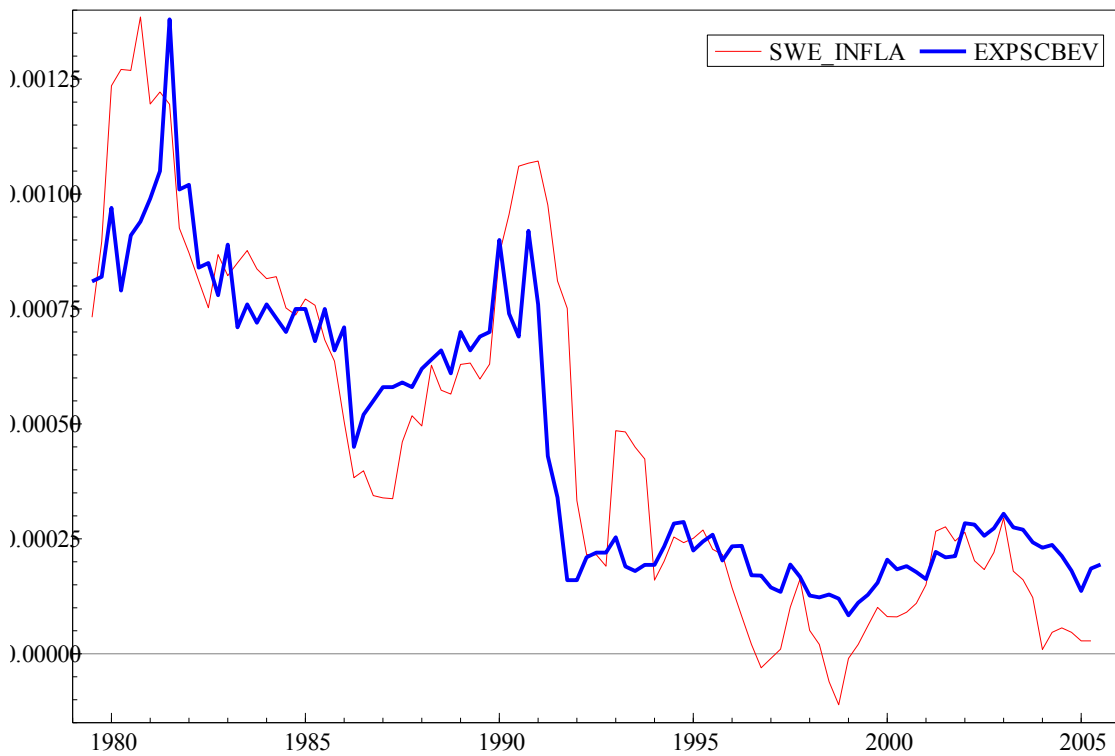
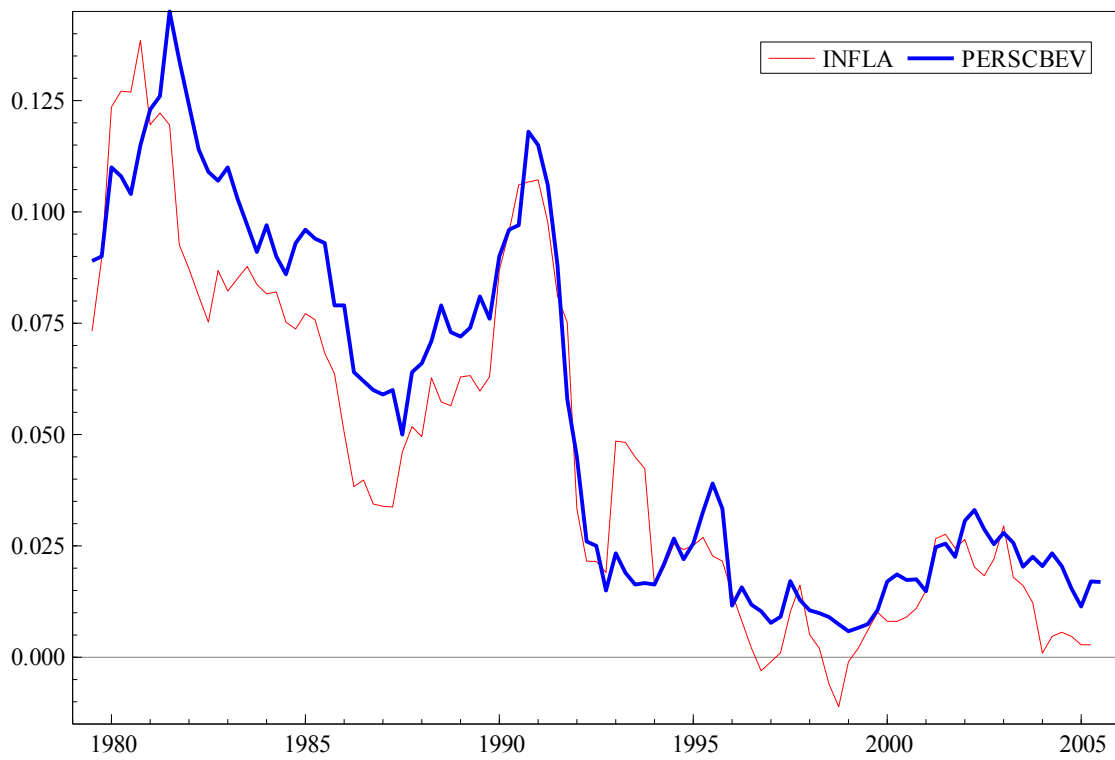
UNITED KINGDOM



SWEDEN (QUALITATIVE DATA)



SWEDEN (QUANTITATIVE DATA)



APPENDIX II

ADDITIONAL DATA AND DIAGRAMS FOR CHAPTER III

I. ADDITIONAL TABLE AND FIGURES FOR SECTION 2.3

TABLE 38: WEIGHTS OF HICP COMPONENTS FOR SWEDEN, 1990-2005

time	cp01	cp02	cp03	cp04	cp05	cp06	cp07	cp08	cp09	cp10	cp11	cp12
1990	20.52%	6.45%	7.48%	15.01%	7.30%	0.54%	19.79%	2.25%	11.58%	0.00%	5.67%	3.43%
1991	19.22%	6.21%	8.16%	17.74%	7.48%	0.57%	17.42%	2.33%	11.97%	0.00%	5.43%	3.47%
1992	18.87%	6.37%	7.62%	18.86%	7.31%	0.74%	18.09%	2.44%	11.56%	0.00%	4.83%	3.32%
1993	18.57%	6.79%	7.90%	19.25%	6.68%	0.76%	17.69%	2.47%	11.37%	0.00%	4.83%	3.70%
1994	18.93%	6.66%	7.50%	20.56%	6.53%	0.81%	15.98%	2.79%	11.27%	0.00%	5.25%	3.74%
1995	18.95%	6.54%	7.05%	20.23%	6.18%	1.16%	16.60%	3.01%	11.27%	0.17%	4.87%	3.92%
1996	17.98%	6.22%	6.90%	20.89%	6.40%	1.21%	17.73%	2.91%	10.82%	0.17%	4.77%	4.00%
1997	17.51%	5.83%	7.20%	21.36%	6.00%	1.22%	17.70%	3.14%	10.82%	0.17%	5.36%	3.69%
1998	17.64%	6.02%	7.09%	20.84%	6.06%	0.66%	17.92%	3.45%	11.07%	0.16%	5.44%	3.65%
1999	17.65%	5.07%	7.42%	20.60%	6.41%	0.38%	17.65%	3.78%	11.60%	0.16%	5.64%	3.63%
2000	15.69%	5.30%	7.27%	19.45%	5.80%	2.99%	15.75%	3.58%	11.44%	0.35%	6.35%	6.03%
2001	15.40%	5.30%	7.36%	18.12%	6.21%	3.13%	15.68%	3.39%	11.74%	0.35%	6.59%	6.75%
2002	15.62%	5.11%	7.38%	18.80%	6.24%	3.00%	13.95%	3.48%	11.95%	0.41%	6.52%	7.55%
2003	15.02%	5.10%	7.26%	18.13%	6.13%	3.11%	14.95%	3.91%	12.70%	0.30%	6.25%	7.17%
2004	14.74%	4.74%	7.02%	18.44%	6.01%	3.20%	15.20%	3.77%	12.44%	0.30%	6.18%	7.97%
2005	14.90%	4.74%	6.74%	18.17%	5.95%	3.35%	15.68%	3.89%	12.56%	0.34%	6.17%	7.52%

SOURCE: Eurostat

FIGURE 90: INFLATION PERCEPTIONS EQCM, GRAPHICAL ANALYSIS

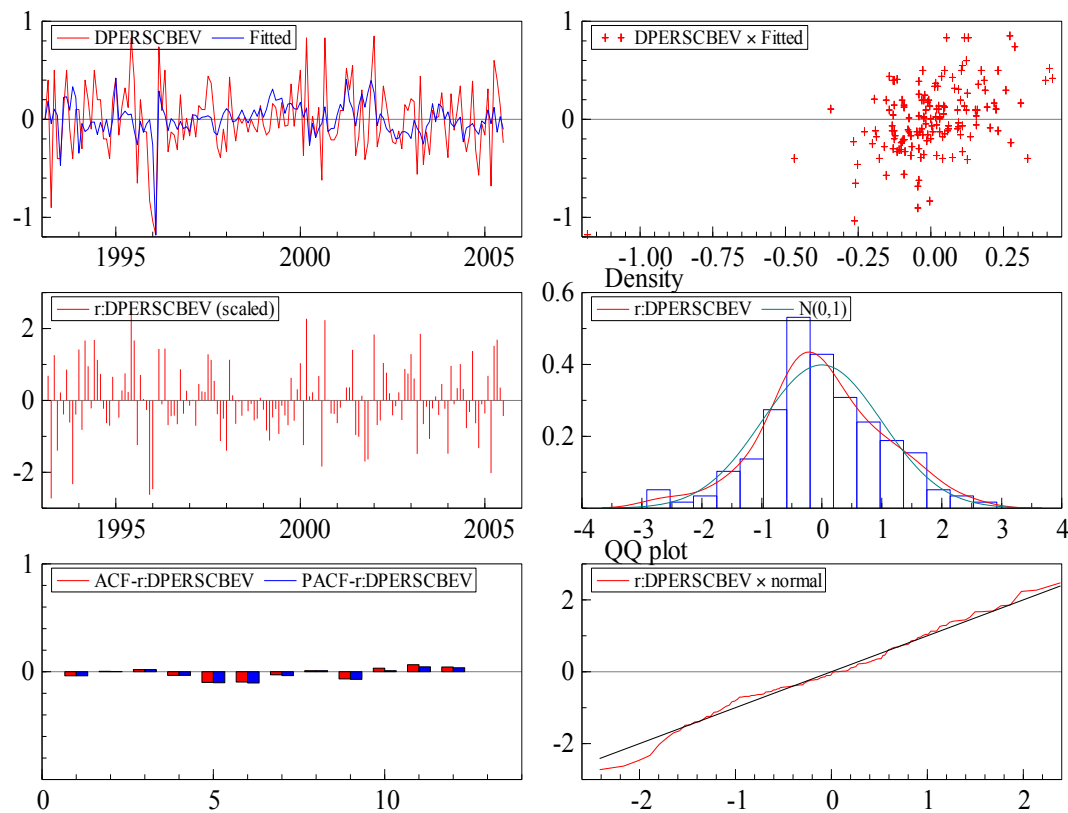
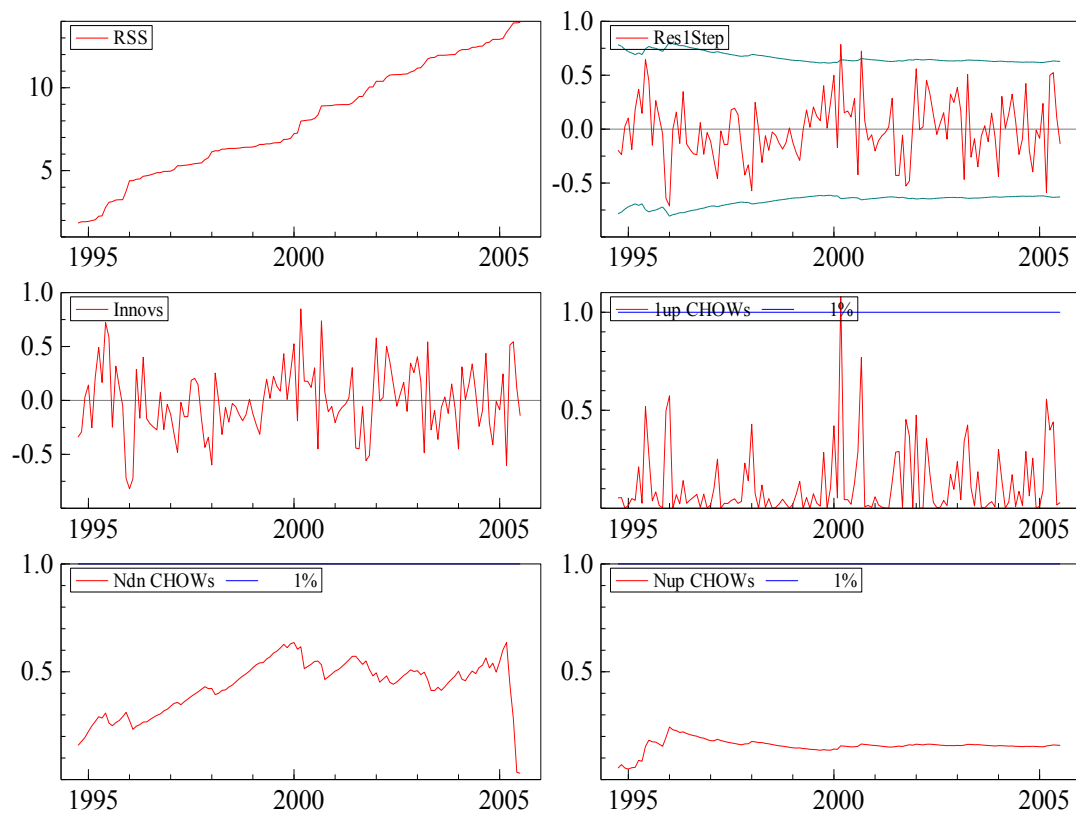
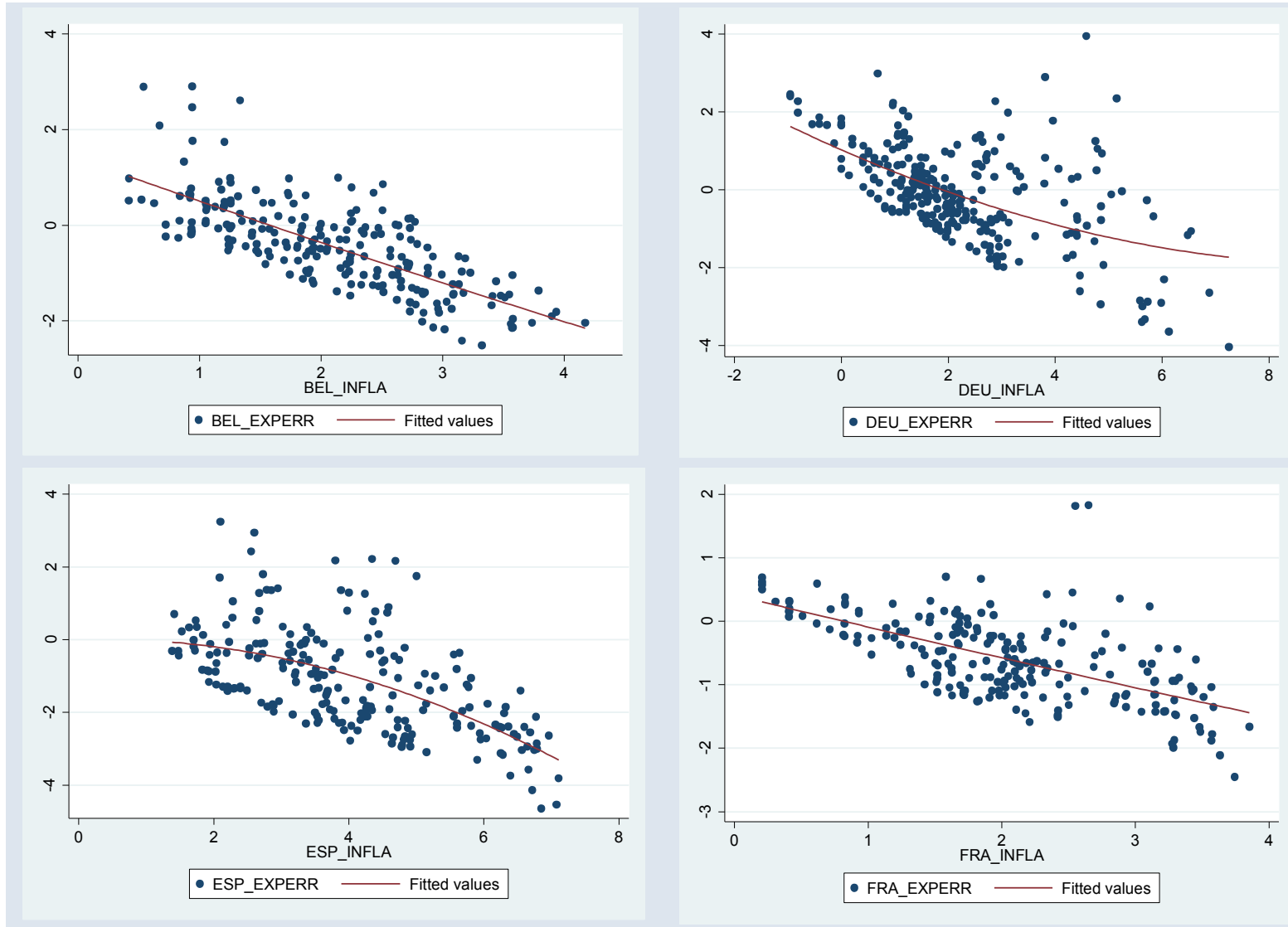


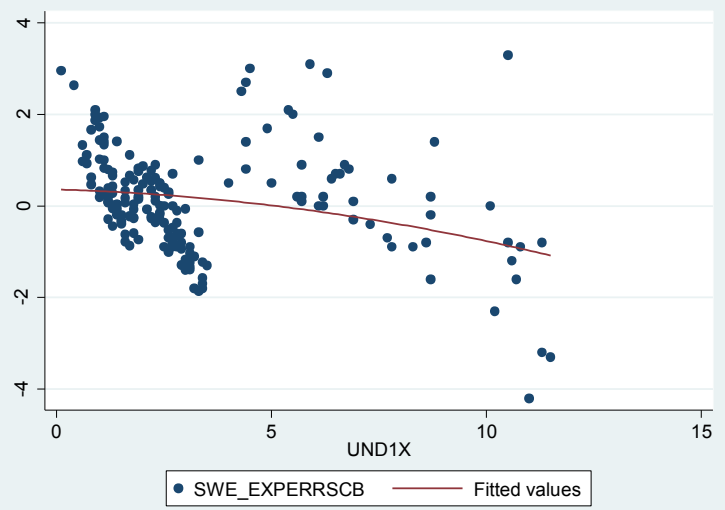
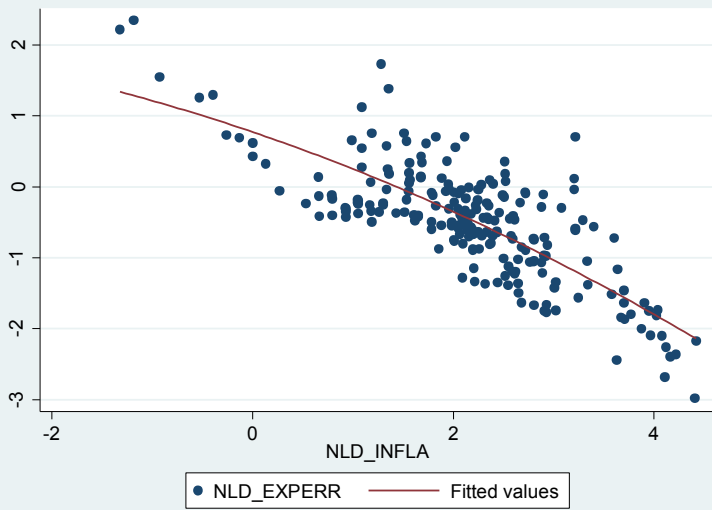
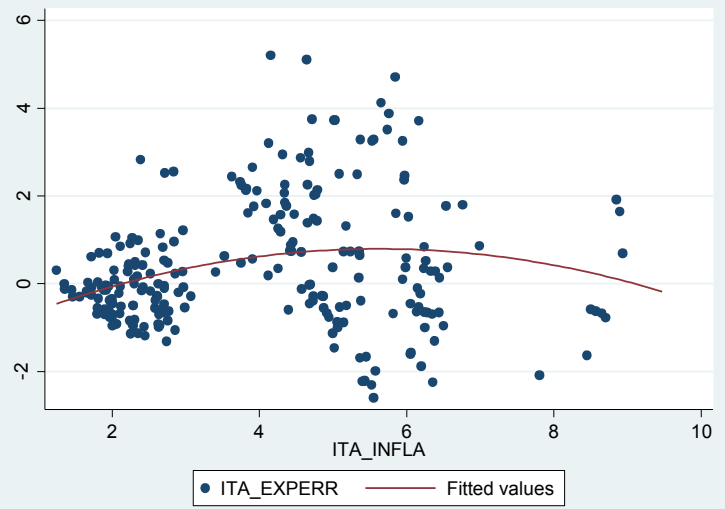
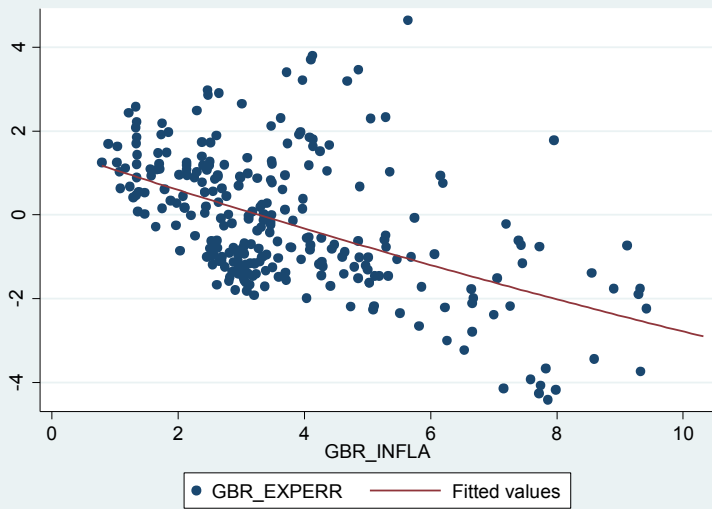
FIGURE 91: INFLATION PERCEPTIONS MODEL, STABILITY TESTS



II. ADDITIONAL FIGURES FOR SECTION 3.3

FIGURE 95: AGGREGATE EXPECTATIONS ERRORS AND INFLATION





III. ADDITIONAL FIGURES FOR SECTION 4.4.2

FIGURE 97: FIT FOR GERBERDING'S RULE-OF-THUMB MODEL, 1991Q1-1999Q4

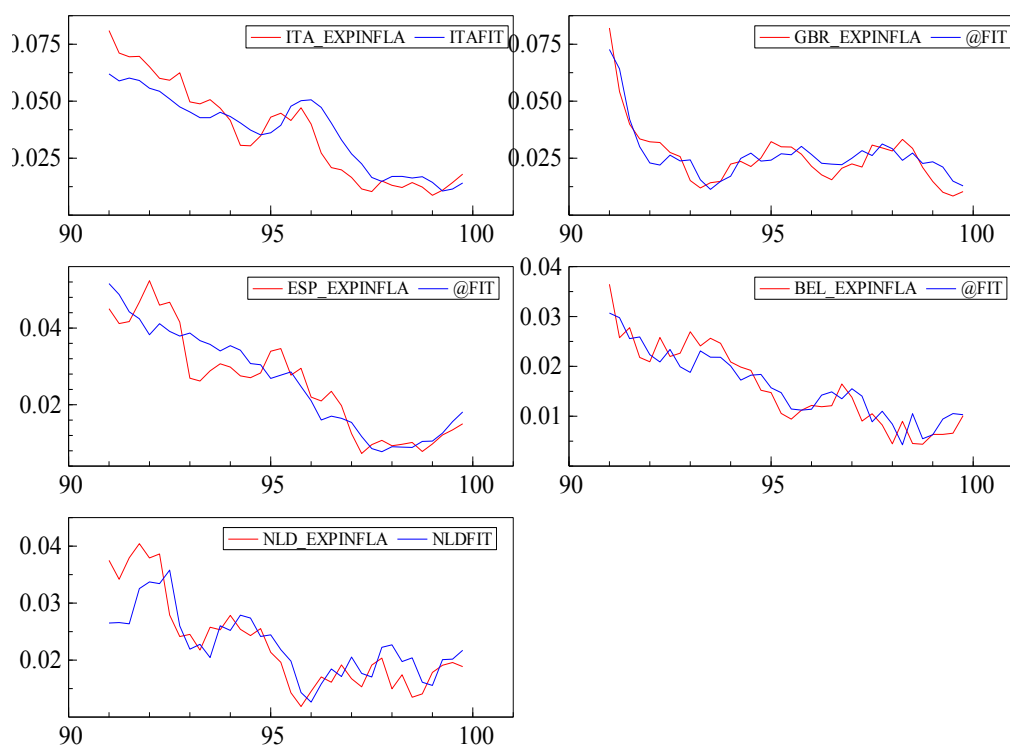
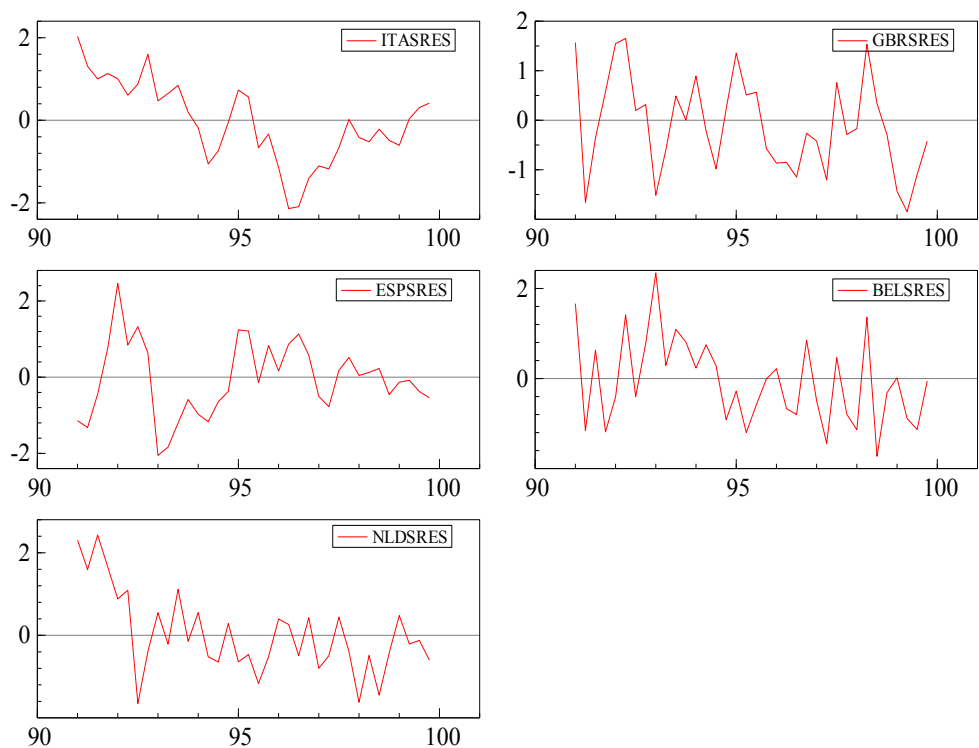


FIGURE 98: STANDARDISED RESIDUALS FOR THE MODEL, 1991Q1-1999Q4



IV. ADDITIONAL FIGURES FOR SECTION 4.4.3

FIGURE 99: RULE-OF-THUMB MODEL, FITTED & ACTUAL VALUES, FS

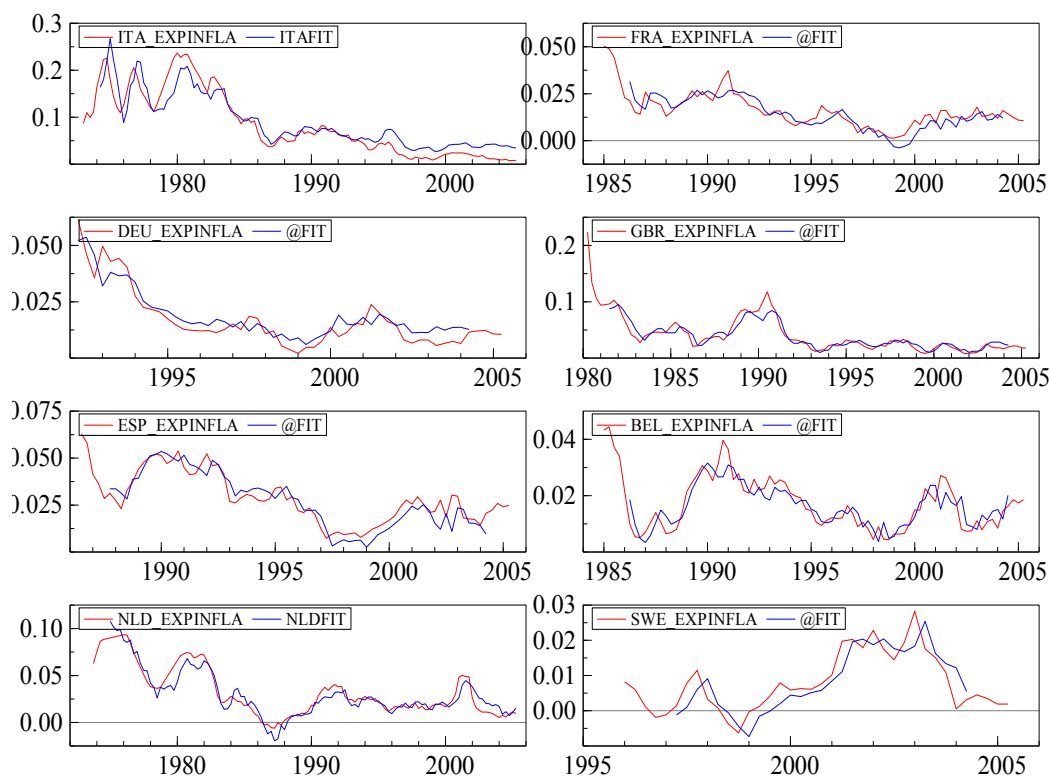
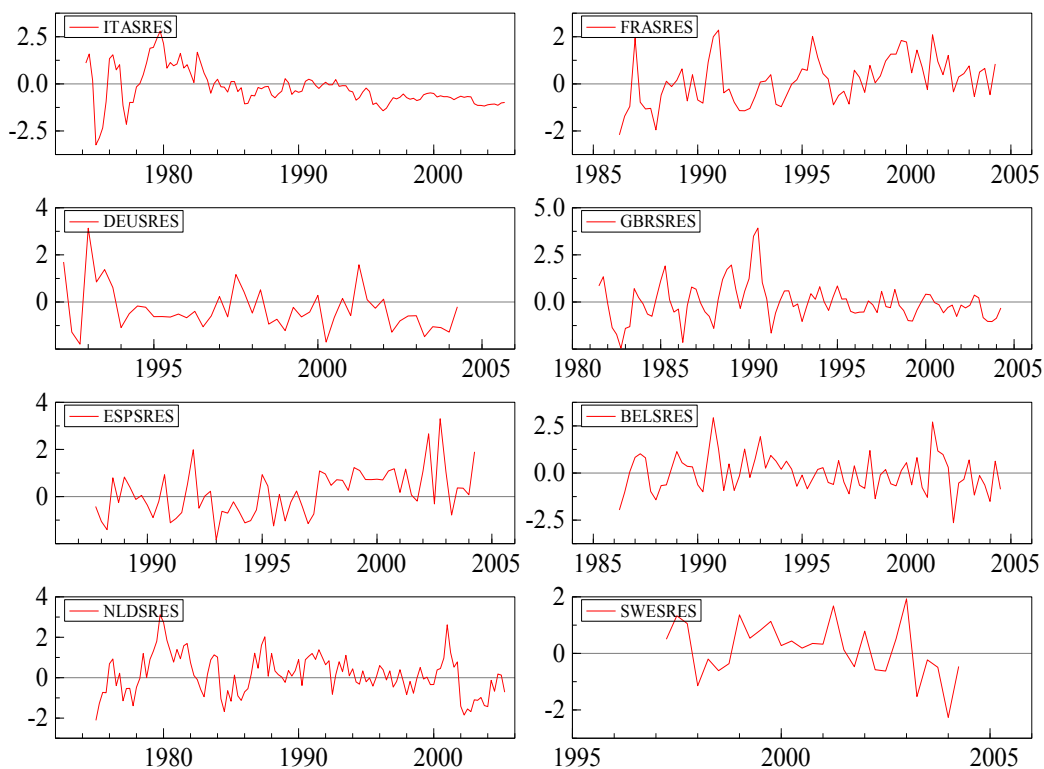


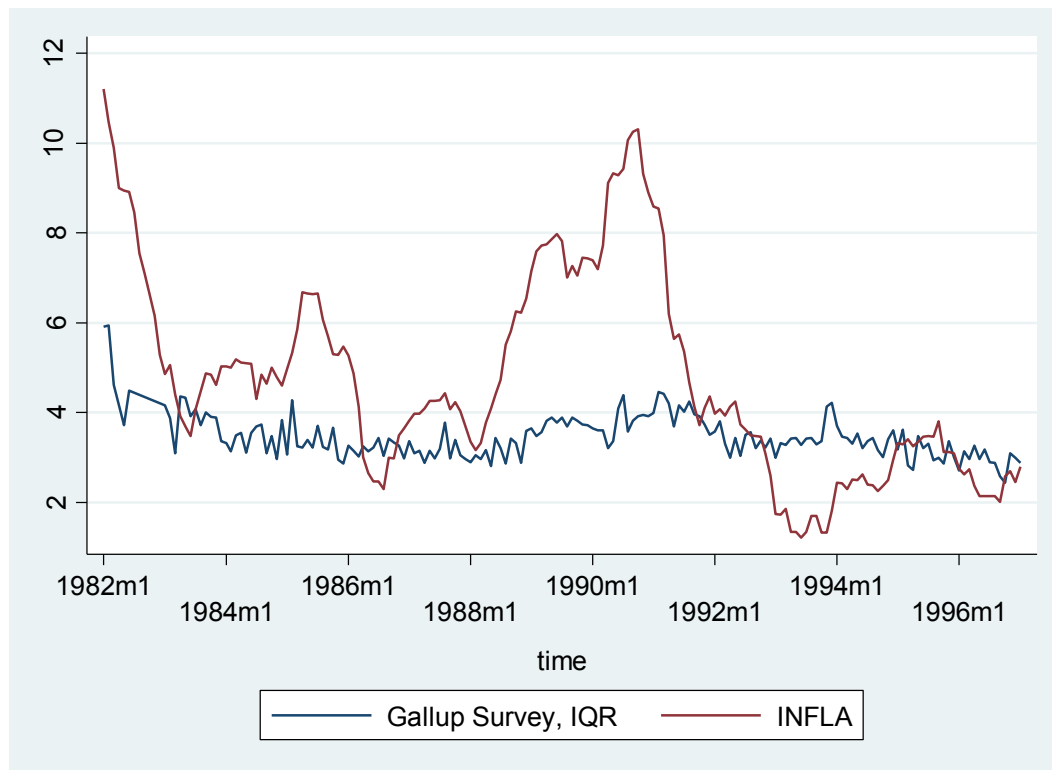
FIGURE 100: RULE-OF-THUMB MODEL, STANDARDISED RESIDUALS, FS



V. GRAPHICAL EVIDENCE FOR SECTION 6.4

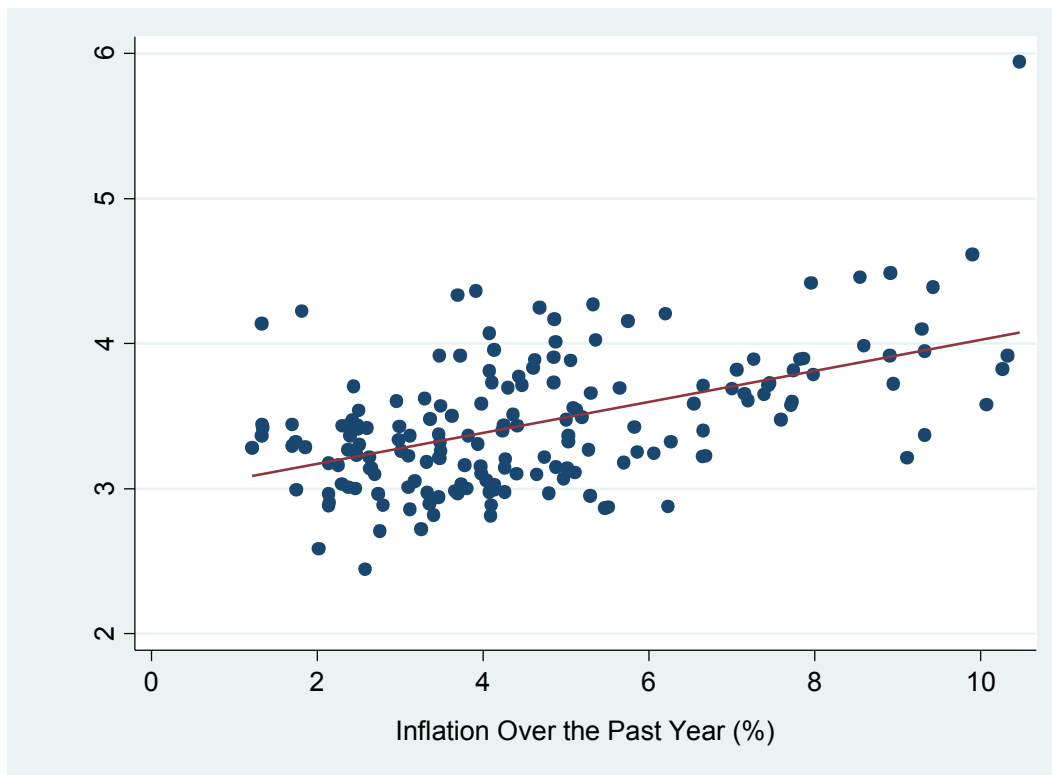
Figure 117 plots the Gallup interquantile range against CPI inflation to explore the relationship between both variables.

FIGURE 117: GALLUP INTERQUANTILE RANGE AND CPI INFLATION



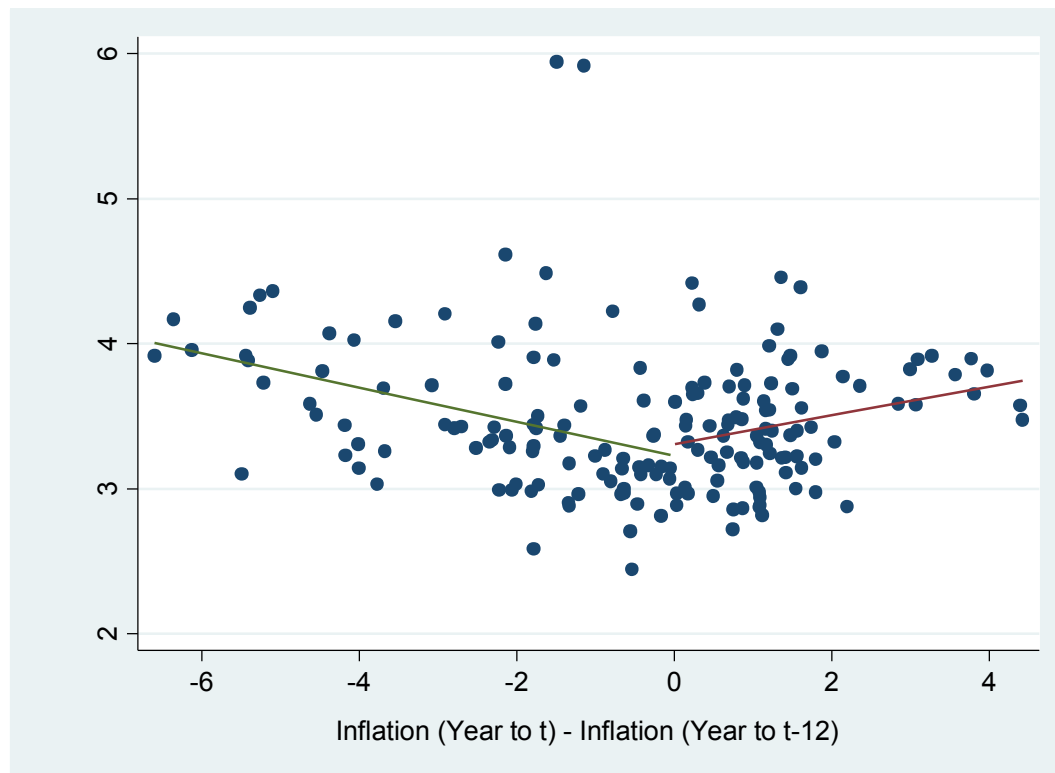
There would seem to be a weak but positive relationship between actual inflation and the extent of disagreement about inflation expectations among consumers. This is explored further in Figure 118. The slope of this relationship appears to be much weaker than that found for US data by MRW.

FIGURE 118: GALLUP INTERQUANTILE RANGE AND CPI INFLATION, SCATTER



The next figure examines the relationship between the interquantile range and the change in inflation. The aim is to check whether disagreement increases when inflation is more volatile. The issue is notably made interesting by the fact that finding such a relationship would be consistent with a staggered adjustment of expectations.

FIGURE 119: GALLUP INTERQUANTILE RANGE AND CHANGES IN INFLATION

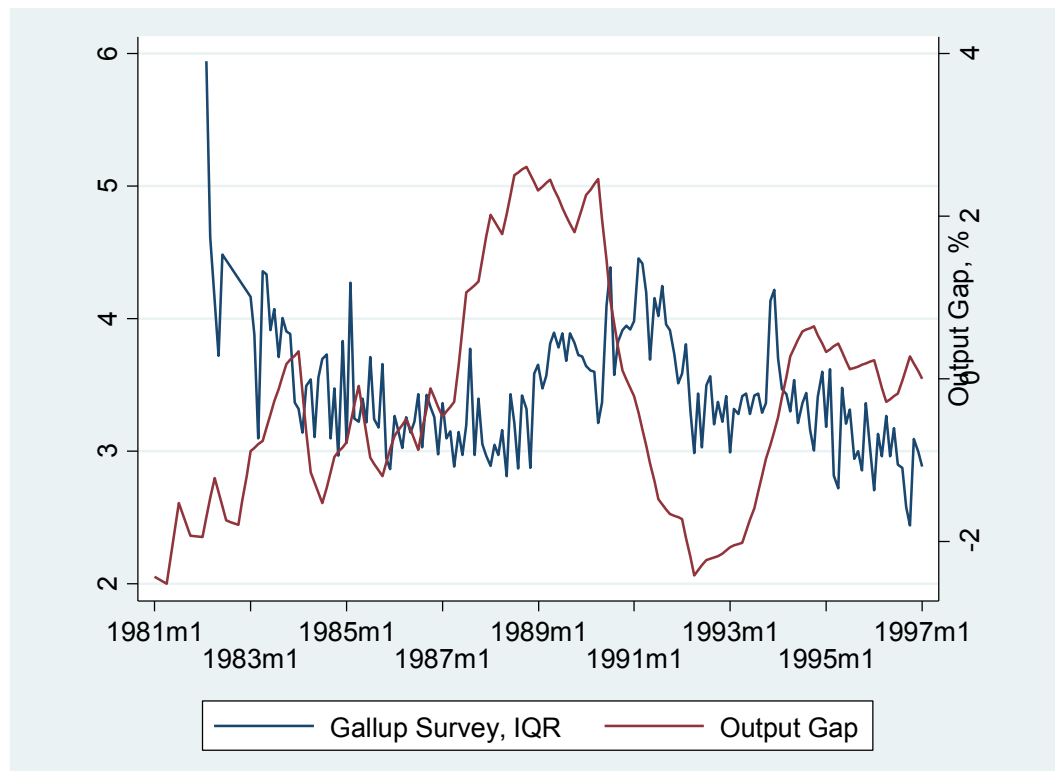


Although there does seem to be a tendency for larger changes in inflation to be associated with more disagreement, the relationship is once again much weaker than that apparent on US data in MRW (2003).

We next consider the relationship between disagreement and the real economy. The cyclical state of the latter is proxied by an output gap variable, constructed as the difference between a quarterly volume measure of GDP (from the OECD's Main Economic Indicators database) and its trend (obtained from a Hodrick-Prescott filtering exercise, with the conventional lambda parameter of 1600).

Figure 120 plots the Gallup interquantile range against this output gap measure.

FIGURE 120: GALLUP INTERQUANTILE RANGE AND THE OUTPUT GAP



It would appear that disagreement tends to rise during booms and fall during recessions. This is the exact opposite of the finding claimed by MRW (although their data are actually ambiguous in this respect).

Finally, we consider how the dispersion of inflation rates across different HICP sub-components can influence disagreement. To study this question, we construct the interquartile range for the distribution of inflation across twelve HICP subcategories (listed in Section 2.3). This is an exercise analogous to that in Bryan and Cecchetti (1994), who produce a weighted median CPI. Essentially, we assume that the HICP weight of a sub-component represents the percentage of the distribution of all prices experiencing that sub-component's amount of inflation. The resulting 75th percentile of

inflation, 25th percentile and interquantile range are plotted in Figure 121. Figure 122 shows that there appears to be a weakly positive relationship between disagreement and the dispersion of HICP component-level inflation rates; once again, the slope is much lower than that of its US counterpart in MRW (2003).

FIGURE 121: DISTRIBUTION OF PRICE CHANGES ACROSS HICP COMPONENTS

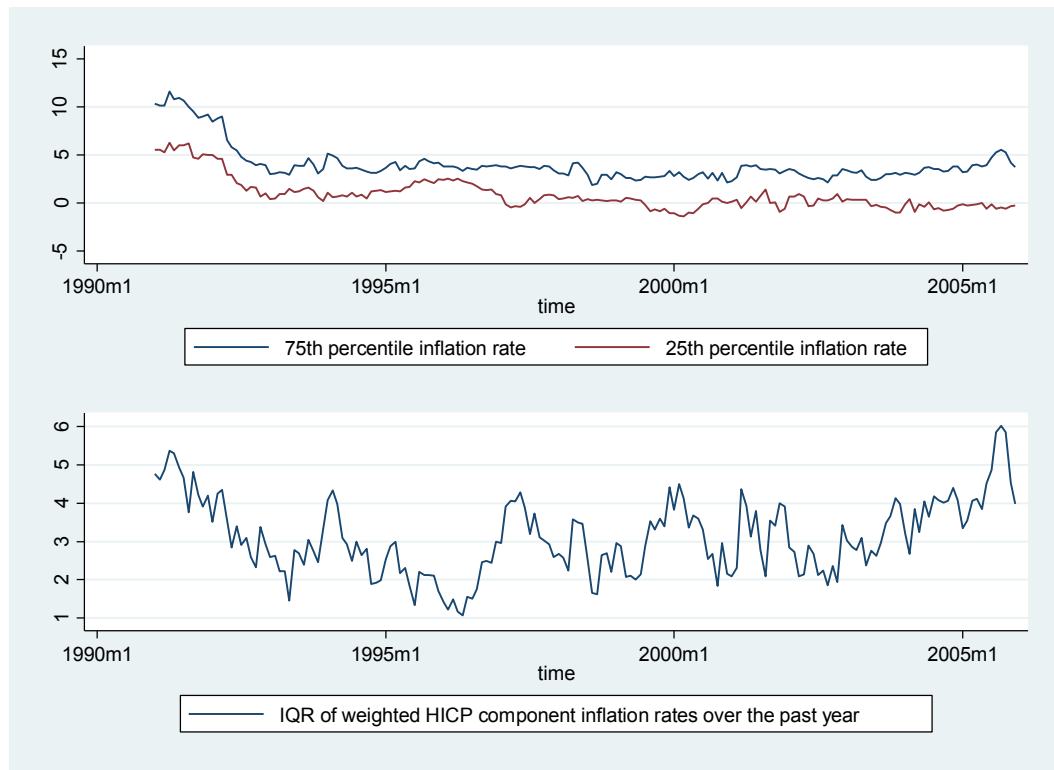
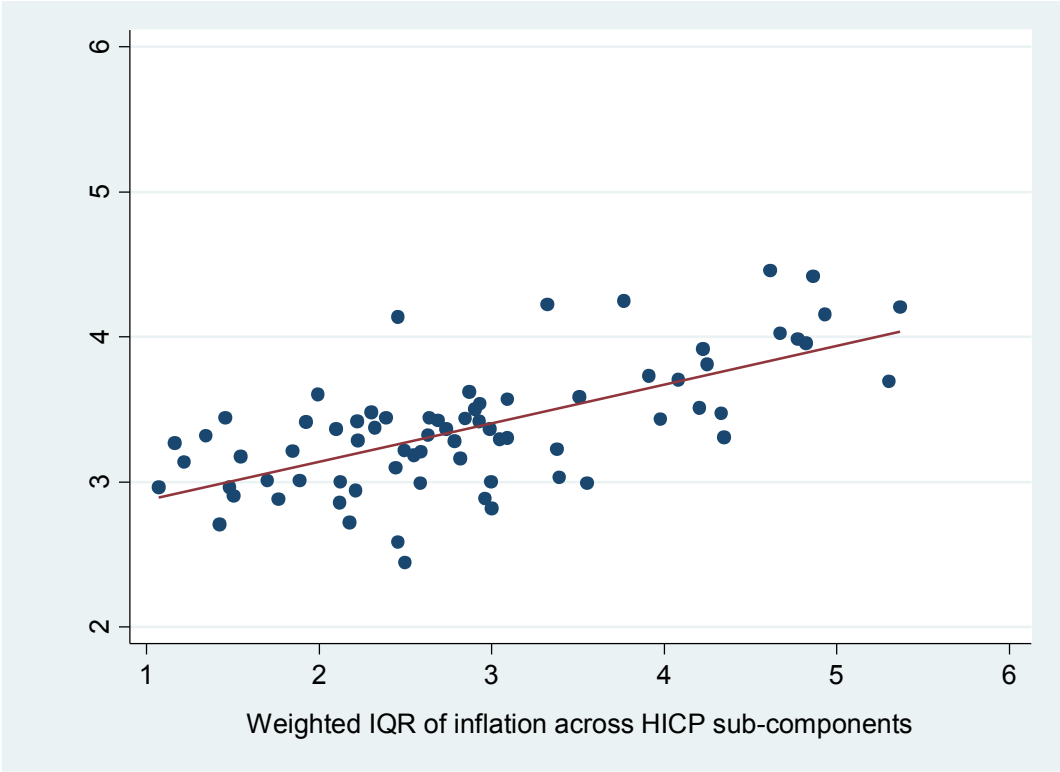


FIGURE 122: DISPERSION OF GALLUP EXPECTATIONS & HICP INFLATION



VI. GRAPHICAL EVIDENCE FOR SECTION 6.6.4

FIGURE 140: UK INFLATION EXPECTATIONS, GALLUP DATA

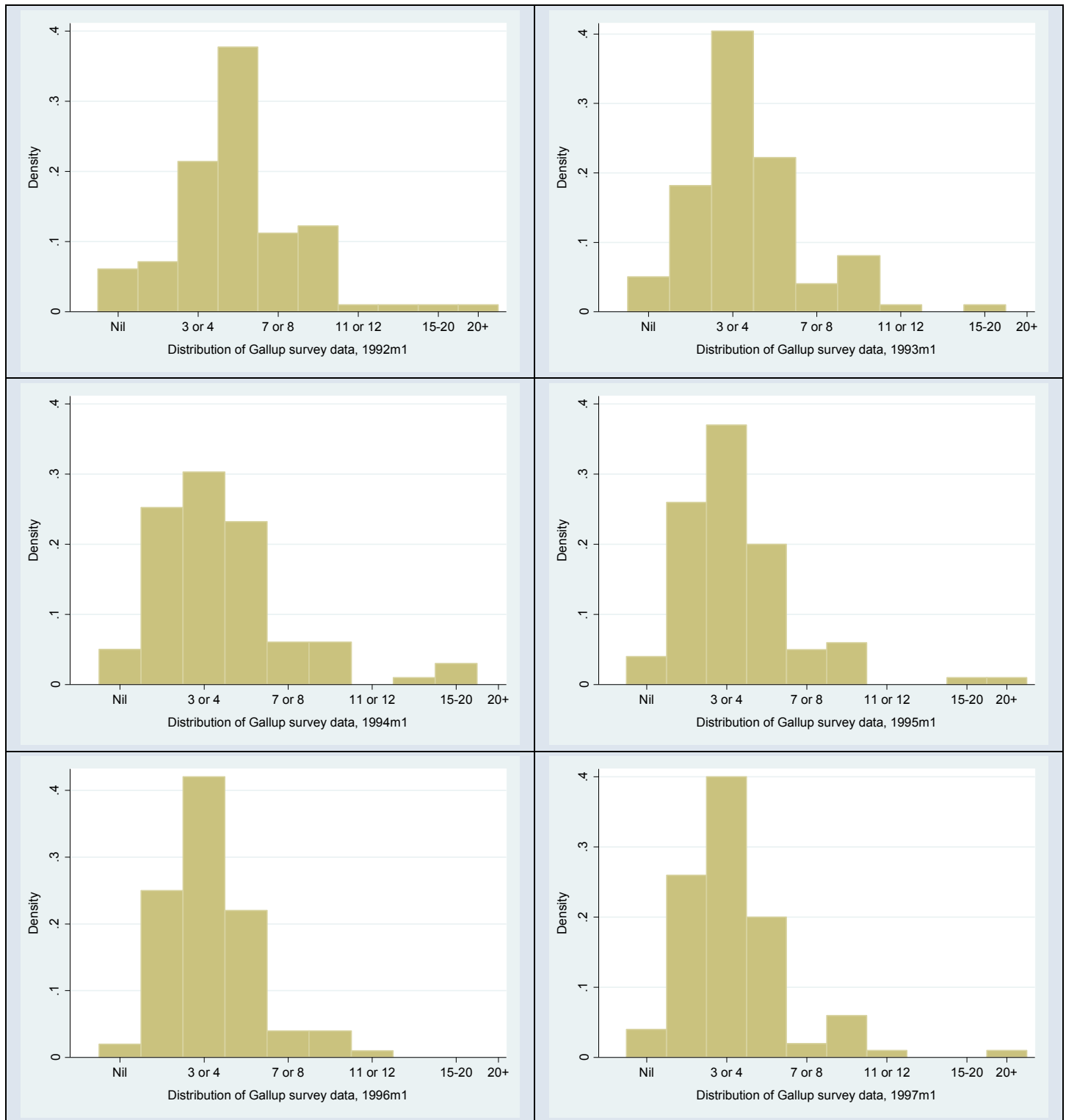
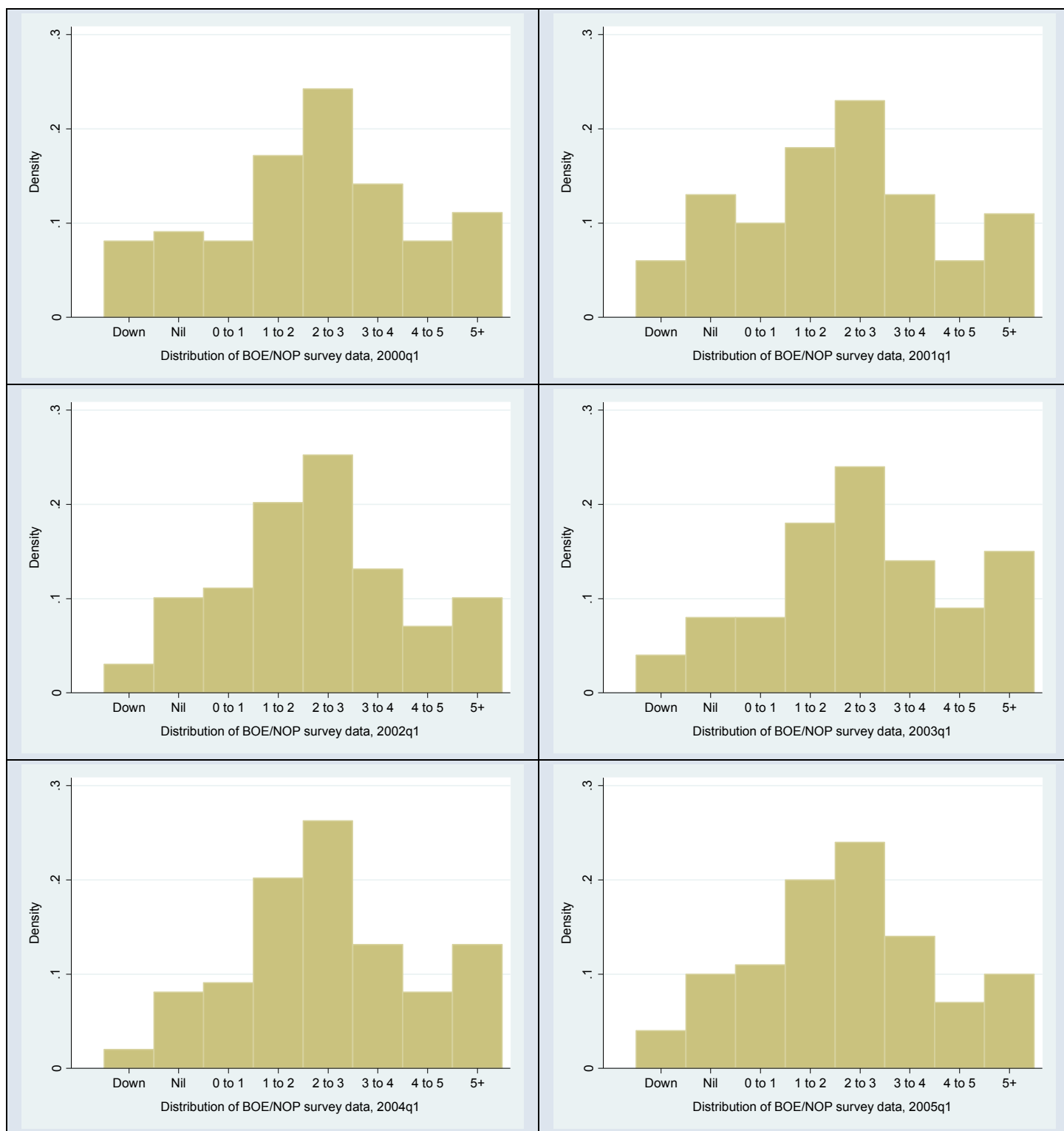


FIGURE 141: UK INFLATION EXPECTATIONS, BOE/NOP DATA



ACADEMIC ACKNOWLEDGEMENTS

I am greatly indebted to my supervisor, Professor John Muellbauer, who provided me with truly exceptional support and inspiration throughout this research. The present study owes much to his challenging contributions, as does my thinking about economics; I am immensely grateful for his help.

Some of this work was initiated as part of my MPhil thesis (Curto Millet, 2004), which was co-supervised by Dr. Luca Nunziata. Luca was critical in helping me assemble and construct high-quality institutional data on labour markets, and I am extremely grateful for his support.

The research presented here has benefited from many insightful discussions. I want to thank in particular Professor Roy Batchelor, Professor Neil Shephard, Dr. Mark Williams, Professor Dennis Snower, Takamitsu Kurita, Samuel W Malone and my two DPhil examiners, Professor Peter Sinclair and Dr. Christopher Bowdler; the two anonymous examiners of my MPhil provided me with very helpful written suggestions.

I also wish to highlight the extraordinary support of Dr. James Forder and Professor David Vines, my tutors at Balliol College. This thesis was yet another opportunity to realise how indebted I am to their teaching.

The material of this thesis has been presented in a number of forums: Nuffield College, as part of the “Macroeconomics and International Economics” series in March 2004; the Maison Française d’Oxford on November 9, 2005; the Gorman Workshop at the Department of Economics in Oxford on February 24, 2006; and the National Bank of Poland on February 10, 2006, as part of the conference on “The Role of Inflation Expectations in Modelling and Monetary Policy Making” in Warsaw. Comments by the participants to those events are gratefully acknowledged; I want to make special reference in this respect to those by my discussant at the NBP, Dr. Jerzy Mycielski.

The assembly of the database used in this article and a full understanding of its details was only possible thanks to the kind and diligent efforts of many.

Data on inflation perceptions and expectations greatly benefited from the input of: Cédric Viguié and Antonio Fuso (DG ECFIN A3, European Commission); Dr. Marco Malgarini and Bianca Maria Martelli (ISAE), Dr. Francesco Papadia (European Central Bank) and Dr. Aviram Levy (Bank of Italy) for the Italian data; Dr. AH Kroese and Dr. Gert Buiten (CBS) for the Dutch data; Klas-Göran Warginger (Konjunkturinstitutet) for the Swedish data; Rolf Bürkl and Klaus Hilbinger (GfK Marktforschung GmbH) for the German data; Hélène Poncet and Philippe Scherrer (INSEE) for the French data; Isabelle De Greef and Jean-Paul Vonck (Banque Nationale de Belgique) for the Belgian data; Janet Lewis, Catty May and Dr. Karen Dunnell (Office for National Statistics) for the British data.

Regarding the collection of institutional and other data, I want to express my appreciation to: Dr. Luca Nunziata for access to the Labour Markets Institutions Database (LMIDB); Glenda Quintini (OECD) for her considerable help in the provision and the interpretation of several OECD datasets; Dominique Paturot (OECD) for the data on benefit replacement rates; Dana Blumin (OECD) for help regarding employment protection legislation data and collective bargaining coverage; Alexandra Wijmenga-Daniel (OECD) for information concerning social security contributions. Dr. Wolfgang Ochel was instrumental in helping me extend his indicators on bargaining coordination, and kindly commented on my first round of updates; Professor Bernhard Ebbinghaus contributed helpful comments on the updating of net union density figures. I also want to thank Lene Mejer of Eurostat Unit D4 for her clarifications regarding HICP data.

Finally, I want to thank all those who kindly sent me papers or references that would otherwise have been unavailable and contributed to my work. Among them are Dr. Janine Aron, Dr. Iman van Lelyveld, Professor Jürgen Reckwerth, Laurence Coupé-Mbida and Ángel Estrada. The kind help of the staff in the Bodleian Library, the Social Sciences Library and the British Library was also warmly appreciated.

I want to acknowledge the financial support for my research received from the University of Oxford through the Scatcherd European Scholarship (2002-2005) and the Department of Economics through the Departmental Teaching Studentship (2005-2006).

All remaining errors are my own, as are the views expressed in this document. In particular, they should not be taken to reflect the views of the European Commission, which kindly provided some of the data used here, nor of NERA Economic Consulting.

PERSONAL ACKNOWLEDGEMENTS

My coming so far owes immensely to the affection and support of my wife, Radhika; she gave up a lot of our common time for this project but was ever encouraging of it, helping me keep perspective and get to the finish line without ending up in too many rabbit holes. I also want to thank my entire family for their support – my parents Nicole and Federico, my brother Daniel, Appa and Amma, Prabha, Lina and Mohan, Chitra and Roman.

I have also been fortunate to have many friends in these few years of research.

Mark Williams and my colleagues at NERA Economic Consulting have been an inspiration, and injected a good dose of real-world economics thinking into my work.

I thank the Maison Française d'Oxford for hosting me and supporting me throughout the year 2005-2006. It has allowed me to finish this Thesis in the midst of a sensational group – Sonia Tébbakh, Guillaume Lachenal, César Ramirez-Montes, Christophe Miqueu, Julie Vatain, Sophie Vasset, Guillaume Clément, Benjamin Sylvand, François et Aude Zanetti, Isaure Devauchelle, Flore Coulouma, Caterina Guenzi, Berny Sèbe et Esmeralda, Géraldine Vaughan, Catherine Saint-Pierre, Nina Sylvanus, Amirouche Moktefi, Régis Ponsard, Cédric.

My friends in the Department of Economics have made these years of hard work extraordinarily enjoyable. I am thinking of Simone Ghislandi, Takamitsu Kurita, Sam Malone, Mohammad Niaz Asadullah, Yuwi Manachotphonng, César Calvo, Gonzalo Salinas, Victor Pavon-Villamayor, Gaston Yalonetzky, Marco Hernandez, Andy Zeitlin, Justin Sandefur, Karlygash Kuralbaeva, Maria Ana Lugo, Florencia Lopez Boo, Andrea Pataconi, Sylvano, Tiago Mendes, Karen Croxson, Serineh Najarian, Zenon Quispe.

Quiero también tener un recuerdo especial para los que nos están aquí para leer este trabajo, que me han querido tanto, y que echo tanto de menos: pienso en mis abuelos, en mi tío, el Profesor Gaspar Morocho Gayo, y en mi gran amigo André Lambert.