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

Fabien Curto Millet
Balliol College, University of Oxford

21 January 2006

**Abstract**

A number of competing methods have been considered in the literature with the aim of quantifying qualitative survey data on inflation expectations. However, the issue of the optimal such quantification method has received relatively little attention, or was otherwise examined on limited sets of quantification techniques. This paper aims to address this issue in as comprehensive a manner as possible. We notably examine a number of methods derived from the Carlson-Parkin (1975) tradition, including several variants of the Batchelor and Orr (1988) approach, with two new proposed modifications; the 'regression method' of Pesaran (1984), and its reconsideration by Smith and McAleer (1995); and stochastic time-varying parameter approaches following Seitz (1988). Although a criterion of predictive ability is considered, the success of a quantification methodology is assessed on the basis of its ability to match quantitative expectations data for the United Kingdom and Sweden and on its behaviour in an important economic application, namely the modelling of wages in a number of European countries: France, Spain, the United Kingdom, Belgium, Germany, Sweden, Italy and The Netherlands. The best overall performance is achieved by the two modifications of the Carlson-Parkin tradition introduced here.

**JEL Classifications:** E31, E24, C32, C42, D84.

**Keywords:** Inflation expectations, survey data, quantification, wage equations.

**Contact Details:** Fabien Curto Millet

Balliol College
Oxford University
Oxford OX1 3BJ
United Kingdom

Tel: +44 (0)788 751 84 76
E-Mail: fabien.curtomillet@economics.ox.ac.uk

## ACKNOWLEDGEMENTS

This paper draws much of its material from my DPhil thesis (Curto Millet, 2006) under the supervision of Professor John Muellbauer, who has provided me with truly exceptional support and inspiration over the past three years, for which I am immensely grateful.

Some of this work is also based on my MPhil thesis (Curto Millet, 2004), which was co-supervised by Dr. Luca Nunziata, who was critical in helping me assemble and construct high-quality institutional data. That earlier work has benefited from comments and discussions with Professor David Vines, Dr. James Forder, two anonymous examiners and participants to my seminar in Nuffield College's "Macroeconomics and International Economics" series in March 2004.

I want to thank Professor Roy Batchelor, Professor Neil Shephard, Takamitsu Kurita and Samuel W. Malone for very insightful discussions on the new material here, as well as comments by participants to my presentation in the Maison Française d'Oxford on November 9, 2005.

The assembly of the database used in this article and a full understanding of its details was only possible thanks to the kind and competent 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 Instititutions 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 am also extremely grateful to Radhika Rathinasabapathy, Daniel Curto Millet and my family for their comments and extraordinary support of my work.

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.

# 1 INTRODUCTION

The importance of expectations is one of the central tenets of modern theoretical economics[1]. 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 frameworks[2]. For empirical economists, this has generated the fundamental challenge of seeking ways to measure these "animal spirits", to borrow the words of Keynes. This paper 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). 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 (Roberts, 1997; Curto Millet, 2006).

The failures of approaches based on "guessing" ex ante the form of expectations has made the development of direct measures of expectations 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). 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 sectors[3], and may involve additional complications (e.g. herding behaviour, time-varying risk premia, regulatory requirements[4]).

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 this data, the task of the present paper will be to find the optimal method of quantification to make this data useable.

[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 Englad/NOP survey. See Bank of England (2005), for instance.

[3] This is likely given the findings in Bryan and Venkatu (2001) or Carroll (2003)

[4] 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 27,000 consumers every month at present (DG ECFIN, 2003). Participants in the survey are notably 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 model 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 ][5]. Given that options 1 and 3 are essentially equivalent under both wordings, 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 course, to the extent that this assumption is flawed Spanish inflation expectations could suffer from measurement error[6].

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). After a major data recovery effort, we have managed to achieve the sample coverage detailed in Table 2[7]. This has left us with the most complete dataset to be used to date for consumer 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 |
| Germany[8] | 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 |
| Sweden[9] | N/A | N/A | January 1996 |
| United Kingdom | N/A | N/A | November 1981 |

[5] Options 1-3 of question 5 are also different and read [ very sharply / rather sharply/ slightly ].
[6] Gerberding (2001) proposes an alternative solution, the cost of which is the assumption that inflation perceptions are unbiased.
[7] Please refer to the acknowledgements to find a list of all the people who kindly contributed their time to this effort.
[8] 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, as the German wage equation in Section 7.4.8 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.
[9] The European Commission proposes 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 paper. Their output will be plotted for illustrative purposes in the case of the United Kingdom.

### 3.1 The Carlson-Parkin or Probability Approach

The probabilistic approach has a long track record of literature[10] (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 Survey's question 6, it is the latter method that we implement here.

The CP method assumes that respondents standing at time t-12 (months) have formed an expectation $\Pi_{it}^{\exp}$ 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 $f_i(\Pi_{it} \mid I_{i,t-12})$, conditional on the information set $I_{i,t-12}$ available at time t-12. The aggregation of the subjective probability distributions is denoted $g(\Pi_t \mid \Omega_{t-12})$, 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, $\Pi_t^{\exp}$.

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 probabilities[11].

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

1. There is a symmetric range of inflation values around zero (denoted $\varepsilon_t$) and around the perceived rate of inflation over the past twelve months $\Pi_{t-12}^{p}$ (denoted $\delta_t$) which the respondent cannot distinguish. This will lead the respondent to provide answers 4 and 2 respectively in their immediate vicinity[12].

2. Second, as implied by the above notation, we assume that the indifference intervals are equal for all agents, but allow for time-variation.

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

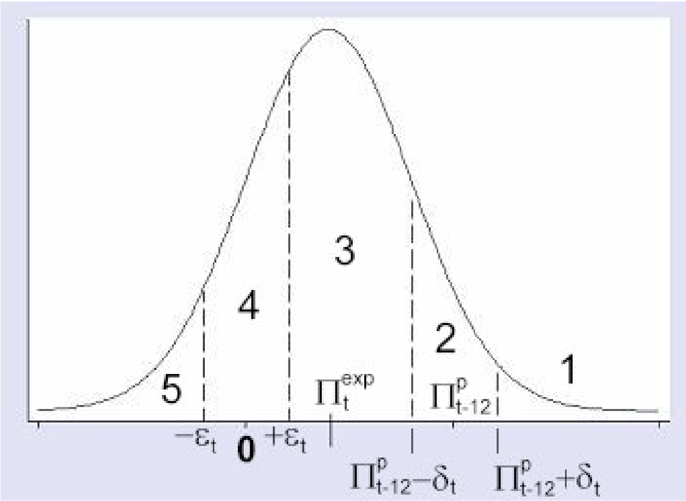
[11] We will focus on response options 1-5 of question 6 (cf. Table 1). The share of respondents choosing the 'don't know' option is allocated proportionally to the other categories (Visco, 1984).

[12] Such indifference intervals can be given psychological groundings, and are termed "difference limens" in the specialised literature (Osgood, 1953).

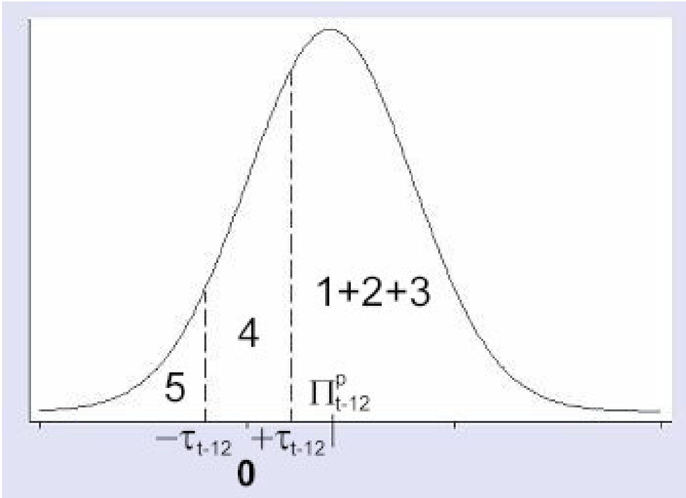
3. The shape of the aggregate distribution function is known and corresponds to the normal distribution[13].

**FIGURE 1: AGGREGATE DENSITY FUNCTIONS: EXPECTATIONS, PERCEPTIONS**

**(a) EXPECTATIONS**




**(b) PERCEPTIONS**




The shape of the assumed distribution has been a particularly contentious point in the literature[14]. Here, we rely on results by Nielsen (2003a) and Berk (1999) which suggest that alternatives make little difference to the derived expectations series.

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

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

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

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

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

[13] This can be justified by an application of the Lindeberg-Lyapunov Central Limit Theorem, for instance.
[14] See Curto Millet (2006) for an overview of this point

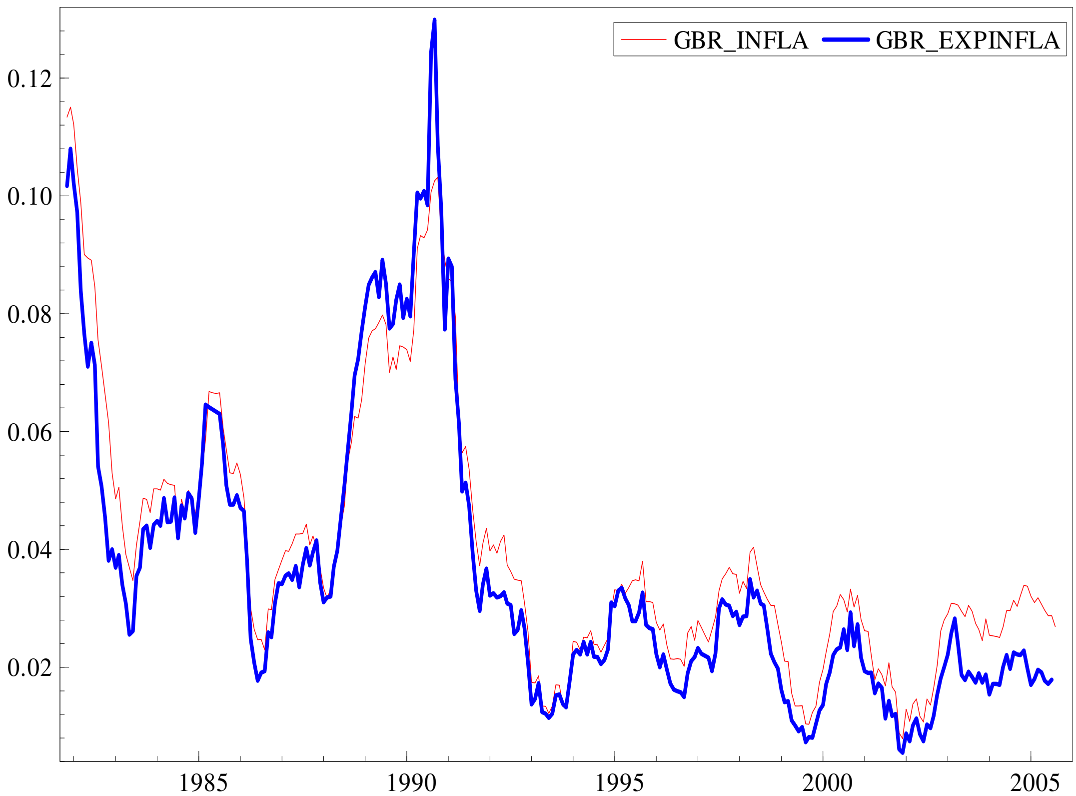
Having assumed the shape of the cumulative density function of $g(\Pi_t \mid \Omega_{t\text{-}12})$, it is possible to derive estimates of the Zs from the response proportions of the survey. The proof then proceeds by re-arranging equations (1) to (4) to obtain an expression for the expected inflation in terms of observables[15]:

$$\Pi_t^{\exp} = \Pi_{t-12}^{p} * [-\frac{Z_t^3 + Z_t^4}{Z_t^1 + Z_t^2 - Z_t^3 - Z_t^4}] \qquad (5)$$

The only missing element in this expression is the perceived inflation rate $\Pi_{t-12}^{p}$. 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**




[15] Refer to the Appendix in Berk (1999) for details of this derivation

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

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"; c.f. 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 manner[16]. With reference to Figure 1 (b) define:

$$W_{t-12}^{1} = \frac{+\tau_{t-12} - \Pi_{t-12}^{p}}{\omega_{t-12}} \tag{6}$$

$$W_{t-12}^{2} = \frac{-\tau_{t-12} - \Pi_{t-12}^{p}}{\omega_{t-12}} \tag{7}$$

The indifference thresholds are here denoted by "τ". The distributional assumptions and the response proportions allow us to derive the following expression for perceived inflation[17]:

$$\Pi_{t-12}^{p} = \left(\frac{W_{t-12}^{2} + W_{t-12}^{1}}{W_{t-12}^{2} - W_{t-12}^{1}}\right) * \tau_{t-12} \tag{8}$$

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:

$$\hat{\tau} = \frac{\sum_{t=1}^{T} \Pi_t}{\sum_{t=1}^{T} \left(\frac{W_t^{2} + W_t^{1}}{W_t^{2} - W_t^{1}}\right)} \tag{9}$$

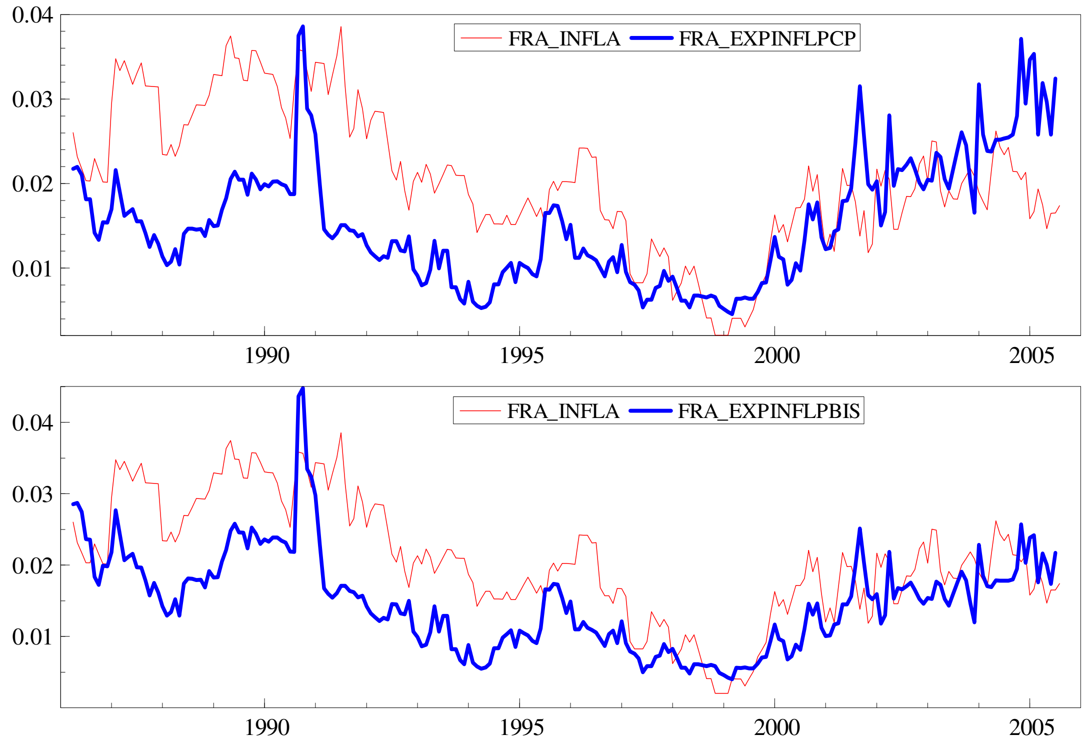
This assumption of constant indifference thresholds may be problematic, especially in the face of large structural changes such as EMU. Curto Millet (2004) showed that the introduction of the new currency was generally associated with a collapse in the proportion of respondents perceiving zero inflation 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. Curto Millet suggests that one interpretation of this is that the response thresholds have shifted inwards over the EMU period, so that it takes a smaller absolute

[16] The 'don't know' answers are, as before, proportionately assigned to the other response categories (Visco, 1984).
[17] See Carlson and Parkin (1975) for details

value of perceived inflation or deflation for the respondents to call it as such. 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).

The results for the previous two methods 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*.

**FIGURE 3: FRA, CP TRICHOTOMOUS QUANTIFICATION, 1981m11-2005m9**




### 3. Pentachomous quantification of question 5 (perceptions) to obtain $\Pi^{p}_{t-12}$

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”, $\Pi^{m}_{t}$. 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 $\Pi_t^m$ which are indistinguishable to respondents, and are enclosed by indifference thresholds. The relevant picture is then analogous to Figure 1 (a) for 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:

$$\Pi_{t-12}^p = \Pi_t^m * [-\frac{A_t^3 + A_t^4}{A_t^1 + A_t^2 - A_t^3 - A_t^4}] \tag{10}$$

The only missing item is then the moderate rate of inflation $\Pi_t^m$. We consider three proxies for this quantity[18]:

a. **The average value of the actual inflation rate over the whole sample[19].**

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 surveyed[20].**

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 inflations 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.

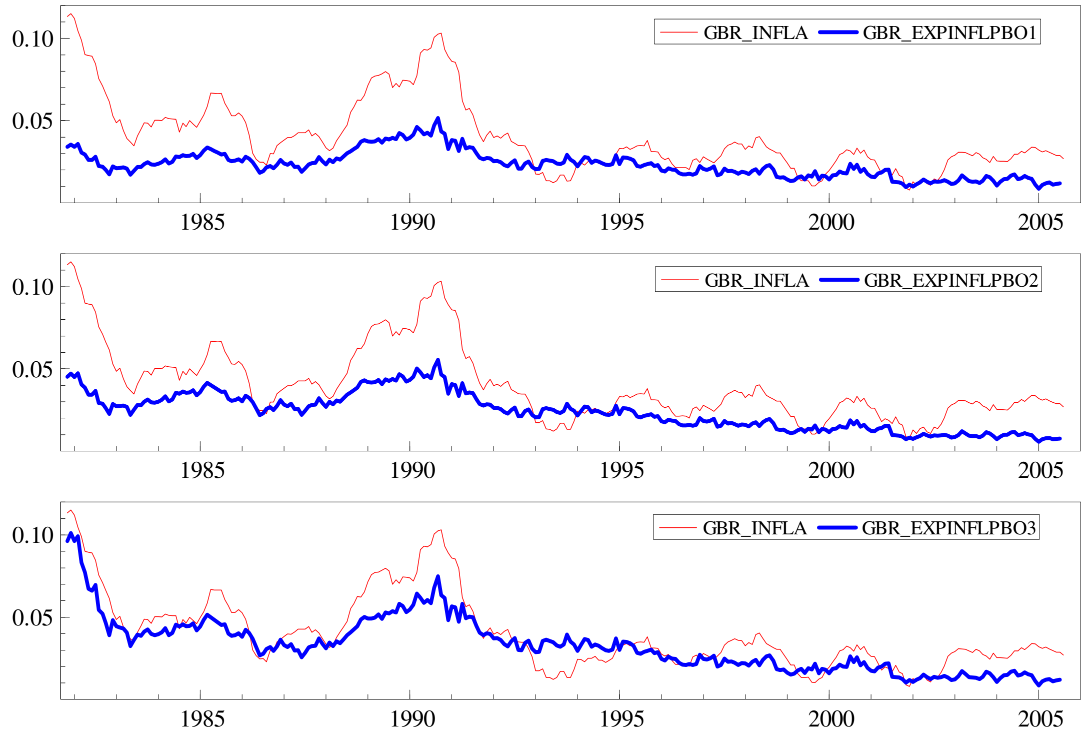
---

[18] 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.

[19] This approach is followed by Reckwerth (1997)

[20] This is suggested in Nielsen (2003a)

**FIGURE 4: UK, CP PENTACHOMOUS QUANTIFICATION, 1981m11-2005m9**




## 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 author[21].

Relate $\pi_t$ over the past 12 months to a weighted combination of respondents having experienced increasing (superscript +) or falling (superscript –) prices:

$$\pi_t = \sum w_i^+ \pi_{t,i}^+ + \sum w_i^- \pi_{t,i}^- \tag{11}$$

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:

$$\begin{cases} \pi_{t,i}^+ = \alpha + \lambda \pi_t + v_{i,t}^+ & \alpha \geq 0,\ 0 \leq \lambda \leq 1 \\ \pi_{t,i}^- = -\beta + v_{i,t}^- & \beta \geq 0 \end{cases} \tag{12}$$

[21] 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.

Weighing all respondents equally and substituting the relations in (12) into (11) suggests the following nonlinear regression:

$$\pi_t = \frac{\alpha R_t - \beta F_t}{1 - \lambda R_t} + u_t \tag{13}$$

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 (13) 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:

$$\hat{\pi}^{\exp}_{t,t+12} = \frac{\hat{\alpha} R^{\exp}_t - \hat{\beta} F^{\exp}_t}{1 - \hat{\lambda} R^{\exp}_t} \tag{14}$$

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[22] that a regression like (13) "*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*".

We estimated (13) for our UK data by nonlinear least squares using the TSP module of GiveWin2 and obtained the results in column(1) of Table 3[23]. 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.

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

[22] See Pesaran (1987), page 211.
[23] The initial values for the procedure were taken from Smith and McAleer (1995)

nevertheless estimated for interest a model with an AR(2) correction, as in Pesaran (1984) and report the results in Table 3, column(3):

$$\pi_t = \frac{\alpha R_t - \beta F_t + \rho_1[(1-\lambda R_{t-1})\pi_{t-1} - \alpha R_{t-1} + \beta F_{t-1}] + \rho_2[(1-\lambda R_{t-2})\pi_{t-2} - \alpha R_{t-2} + \beta F_{t-2}]}{1-\lambda R_t} + \varepsilon_t \quad (15)$$

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 is uninformative relative to lags of inflation, which undermines the use of this modification of the approach.

This begs 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% level[24]. 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) – 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) error[25]. 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-stationarity[26].

Smith and McAleer (1995) propose to modify Pesaran's approach by specifying the following symmetric response model instead of (12):

$$\begin{cases} \pi^{+}_{t,i} = \alpha + \lambda \pi_t + v^{+}_{i,t} & \alpha \geq 0,\ 0 \leq \lambda \leq 1 \\ \pi^{-}_{t,i} = -\beta + \lambda_2 \pi_t + v^{-}_{i,t} & \beta \geq 0 \end{cases} \quad (16)$$

Which leads to the regression:

$$\pi_t = \frac{\alpha R_t - \beta F_t}{1 - \lambda R_t + \lambda_2 F} + u_t \quad (17)$$

[24] This is a modified Dickey-Fuller t-test for a unit root in which the series has been transformed by a generalized least-squares regression. Detailed test results are available from the author.
[25] 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.
[26] 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.

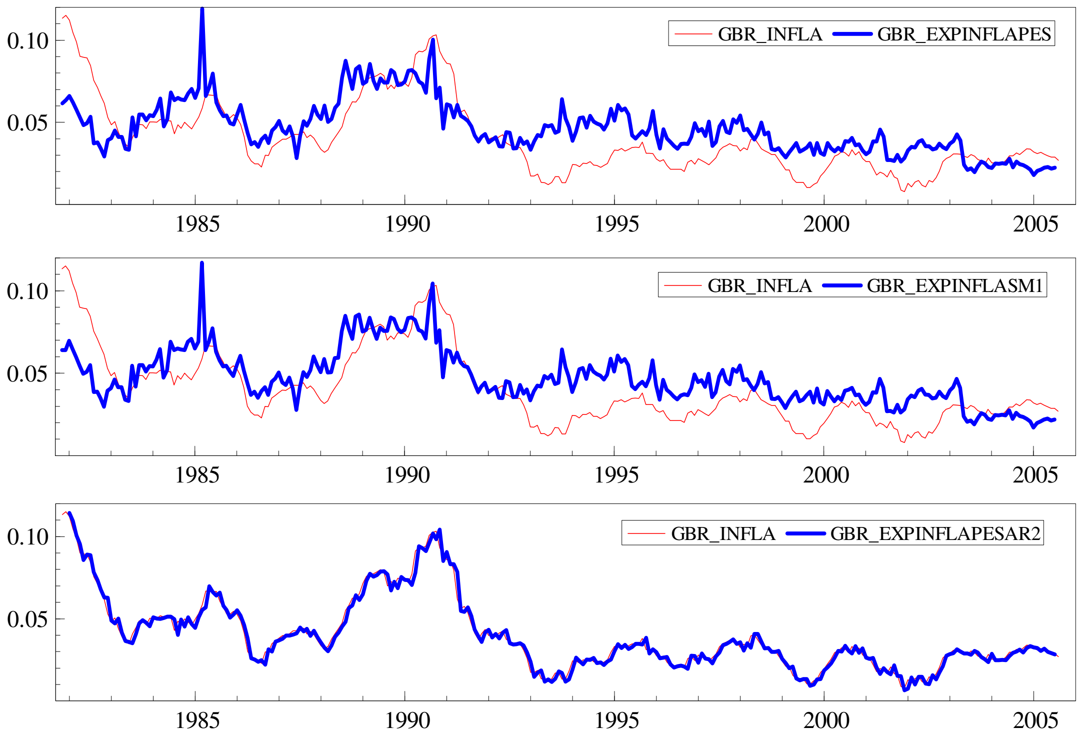
Column(2) of Table 3 makes clear that this is a very minor modification empirically, the additional coefficient being insignificant, and the others staying close to their earlier values (β excepted). The resulting expectations are plotted in Figure 5 (middle).

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

| **Coefficient** | **Standard Pesaran Model** [Eq. (13)] | **Smith & McAleer Extension** [Eq.(17)] | **Pesaran AR(2) Correction** [Eq.(15)] |
|---|---|---|---|
| **$\alpha$** | .010104*** | .00982227*** | -.00354601 |
| **$\beta$** | -.029650*** | -.012423 | -.000135834 |
| **$\lambda$** | .950394*** | .950129*** | .434539** |
| **$\lambda_2$** | - | .807317 | - |
| **$\rho_1$** | - | - | 1.30519*** |
| **$\rho_2$** | - | - | -.314790*** |
| **$R^2$** | .786618 | .788001 | .977825 |
| **Eq. Std. Error** | .010403 | .010385 | .328193E-02 |
| **LM Het. Test** | 28.5278 [.000] | 27.5464 [.000] | 12.4186 [.000] |
| **Durbin-Watson** | .278174 [.000,.000] | .283020 [.000,.000] | 2.04975 [.570,.746] |

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.

**FIGURE 5: ALTERNATIVE EXPECTATIONS SERIES, PESARAN TRADITION**



Smith and McAleer further propose to write parameters α and β as weighted averages of past $\pi_t$ values, turning the equivalent of equation (13) into a dynamic nonlinear regression model.

$$\begin{cases} \alpha_t = \alpha + \lambda_1 \pi_t + \gamma_{11}\pi_{t-1} + \ldots + \gamma_{1m}\pi_{t-m} \\ \beta_t = \beta + \lambda_2 \pi_t + \gamma_{21}\pi_{t-1} + \ldots + \gamma_{2n}\pi_{t-n} \end{cases} \tag{18}$$

$$\pi_t = \frac{\alpha R_t - \beta F_t + R_t \sum_{j=1}^{m} \gamma_{1j}\pi_{t-j} - F_t \sum_{j=1}^{n} \gamma_{2j}\pi_{t-j}}{1 - \lambda_1 R_t + \lambda_2 F_t} + v_t \tag{19}$$

The results of this approach are presented in Table 4[27]. The implausible coefficient values, very high $R^2$ and a ratio of $\gamma_{21}$ and $\lambda_2$ which is close to one indicate that we have (predictably) collided with the same problem that occurred above with the AR correction.

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

| | α | β | λ | $\lambda_2$ | $\gamma_{11}$ | $\gamma_{12}$ | $\gamma_{13}$ | $\gamma_{14}$ | $\gamma_{15}$ | $\gamma_{21}$ | $\gamma_{22}$ | $\gamma_{23}$ | $\gamma_{24}$ | $\gamma_{25}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **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] D-W = 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.

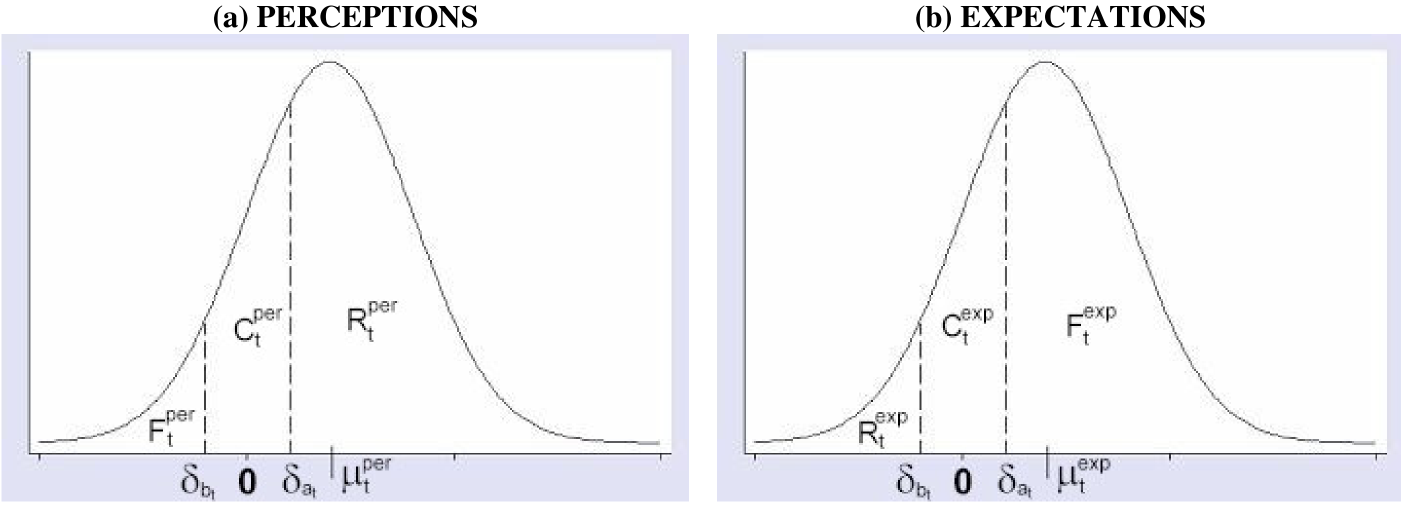
Given the results presented in this section, we choose to retain only the original Pesaran method for the comparisons that are carried out later in this study.

### 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:

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

**FIGURE 6: SEITZ METHOD; PERCEPTIONS & EXPECTATIONS DISTRIBUTIONS**




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

$$\mu_{per,t} = \frac{\delta_{a,t} W_{2,t} - \delta_{b,t} W_{1,t}}{W_{2,t} - W_{1,t}} \tag{20}$$

$W_{2,t}$ 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 (20) as:

$$\begin{aligned} \mu_{per,t} &= \delta_{a,t} x_{1,t} + \delta_{b,t} x_{2,t} \\ &= x_t' \beta_t \end{aligned} \tag{21}$$

Where $\beta_t$ denotes the vector of threshold coefficients at time t: $(\delta_{a,t}, \delta_{b,t})$[28].

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

[28] There is a constraint on this vector – the claim to the contrary in Smith and McAleer is puzzling.

$$\begin{cases} \beta_t = \beta_t^p + u_t \\ \beta_t^p = \beta_{t-1}^p + v_t \end{cases} \tag{22}$$

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:

$$\begin{cases} \mathrm{cov}(u_t) = (1-\gamma)\sigma^2 \Sigma_u \\ \mathrm{cov}(v_t) = \gamma\sigma^2 \Sigma_v \end{cases} \tag{23}$$

The parameter $\gamma \in [0,1]$ shows the relative variance of permanent and transitory changes in the β, with γ=0 corresponding to a purely random coefficients model. Matrices $\Sigma_u$ and $\Sigma_v$ 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 (21) 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 (22) into (21) to obtain a state-space representation with the following measurement and transition equations:

$$\begin{cases} \mu_{per,\,t} = x_t'\beta_t^p + x_t'u_t \\ \beta_t^p = \beta_{t-1}^p + v_t \end{cases} \tag{24}$$

Finally, replace $\mu_{p,\,t}$ in (24) by the actual inflation rate at t, $y_t$ – hence assuming an unbiased perception of past inflation.

The natural step to take at this point is 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:

$$\begin{cases} y_t = x_t'\beta_t^p + x_t'u_t + w_t \\ \beta_t^p = \beta_{t-1}^p + v_t \end{cases} \tag{25}$$

The problem with this step of course is that the system is not identified anymore. This is obviously unsatisfactory yet is ignored in the literature, which seems to simply substitute $\mu_{p,t}$ 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:

$$\begin{cases} y_t = x_t'\beta_t^p + \varepsilon_t \\ \beta_t^p = \beta_{t-1}^p + v_t \end{cases} \tag{26}$$

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 $\beta_t$ then follows two steps:

1. Follow Cooley-Prescott (1973, 1976) to obtain the parameter values $\beta_0^p$, $\gamma$ and $\sigma^2$.

2. Implement the Kalman filter on (26) given the results from the first step.

A derivation of the concentrated likelihood function and the relevant estimators for step 1 is provided in Curto Millet (2006). This procedure was implemented in Ox code, the validity of which was successfully checked by simulation experiments[29].

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

$$\begin{cases} \hat{\beta}_{t|t-1}^p = \hat{\beta}_{t-1}^p \\ P_{t|t-1} \equiv Var(\hat{\beta}_{t|t-1}^p) = P_{t-1} + Q_t \\ \hat{\beta}_t^p = \hat{\beta}_{t|t-1}^p + P_{t|t-1}x_t f_t^{-1}(y_t - x_t'\hat{\beta}_{t|t-1}^p) \\ P_t = P_{t|t-1} - P_{t|t-1}x_t f_t^{-1} x_t' P_{t|t-1} \end{cases} \tag{27}$$

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

[29] These are also described in Curto Millet (2006).

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

$$TRUTH: \quad \beta_t \begin{cases} \delta_{a,0} = 0.025 \\ \delta_{b,0} = -0.040 \end{cases} \qquad \gamma = 0.24, \qquad \sigma^2 = 0.0013, \qquad \Sigma_u = \Sigma_v = I \tag{28}$$

We plot in Figure 7 the $\delta$ 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.

Consider now the empirical implementation of this model. As noted in Nardo (2003), a situation in which the lower threshold $\delta_b$ 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 $\beta_t^p$ and their evolution through time, as derived by the filter:

$$\delta_{a,t} > 0 \qquad and \qquad \delta_{b,t} < 0 \quad \forall \ t \in \{1,...,T\} \tag{29}$$

In practice, this involves carrying out a grid-search over possible values of $\gamma$ and finding the one associated with the highest concentrated likelihood, subject to the constraint. The initial conditions $\beta_0^p$ and $\sigma^2$ associated with this value of $\gamma$ are then inputted with the latter to initialise the Kalman filer algorithm, and it is checked whether the constraint holds dynamically. Should this not be the case, the next-preferred value of $\gamma$ is chosen to initialise the Kalman filter, until a satisfactory specification is found.

**FIGURE 7: KALMAN FILTER SIMULATION RESULTS, 1981m11-2005m6, UK**

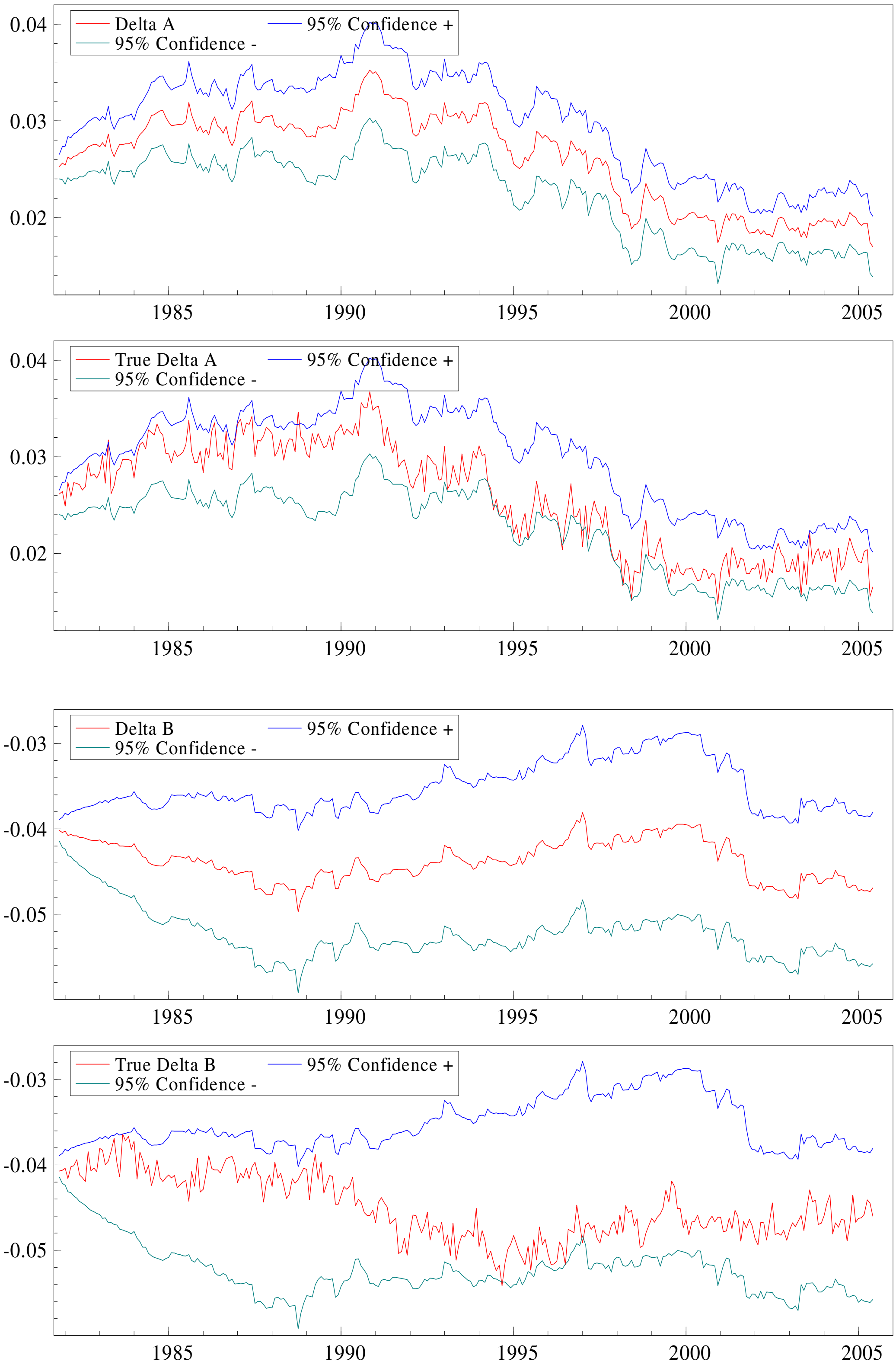


**TABLE 5: SUMMARY OF SPECIFICATION SELECTION FOR TVP MODEL**

| Country | γ maximising unconstrained LogL | γ yielding acceptable $\beta_0^p$ | 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 is 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 (29) and indeed seem to think it justifiable to do so[30], 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 make the focus of this study[31], so comparisons are of limited value – an observation that is compounded by the differences between survey methodologies.

The selected indifference thresholds for the UK are plotted in Figure 8, and the derived expectations are presented against actual inflation in Figure 9. As for every figure in this document, expectations are plotted at the time they are *surveyed* rather than at the time they are *referring to* (+12 months).

[30] Figure 2 on p.172 shows that their results violate the conditions listed in (29). They assert that "in this model, no restrictions are imposed on $a_t$ or $b_t$."
[31] See Seitz (1988) p.428.

**FIGURE 8: TIME-VARYING INDIFFERENCE THRESHOLDS, UK**

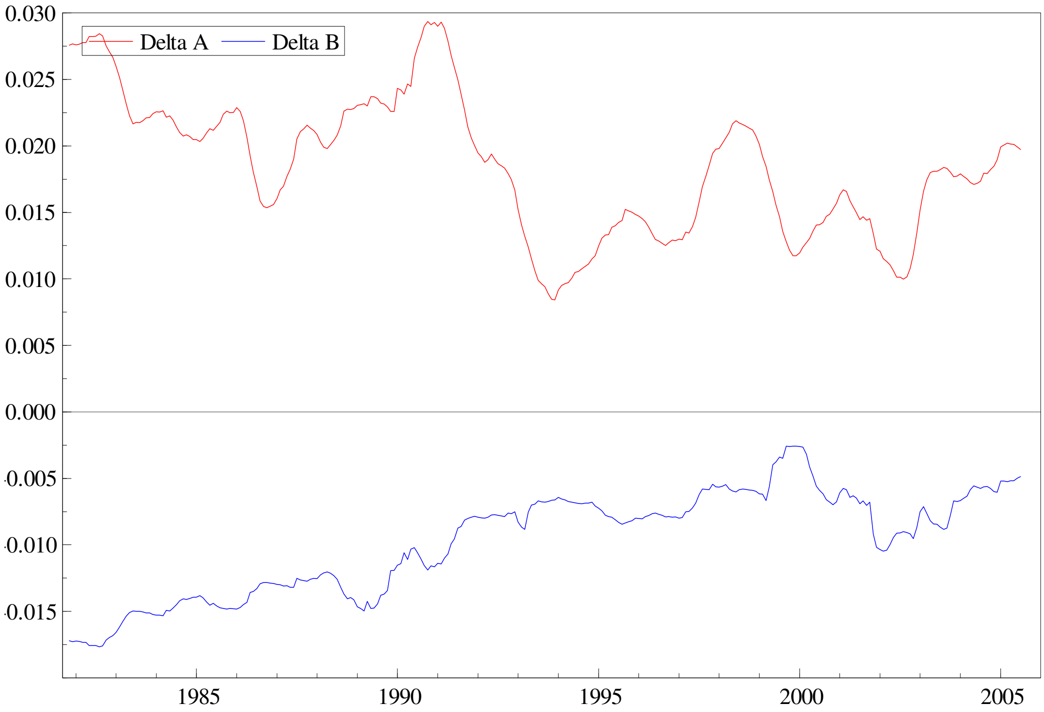
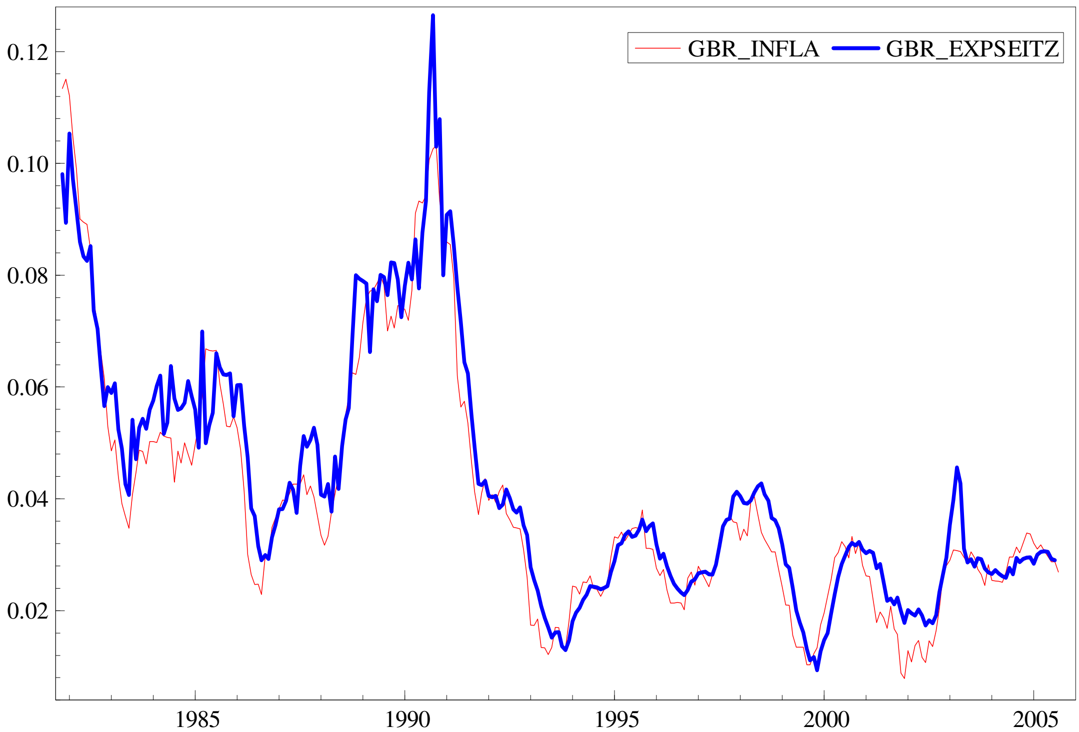



**FIGURE 9: UK INFLATION EXPECTATIONS, SEITZ METHODOLOGY**



### 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 |
|---|---|---|
| Carlson-Parkin Tradition | EXPINFLA | Past inflation perceived correctly (p.8) |
| | EXPINFLPCP | Trichotomous quantification of perceptions (p.9) |
| | EXPINFLPBIS | Trichotomous perceptions, with threshold adjustment. (p.10) |
| | EXPINFLPBO1 | Pentachomous perceptions. Moderate rate is full sample inflation mean. (p.11) |
| | EXPINFLPBO2 | Pentachomous perceptions. Moderate rate from linear interpolation. (p.11) |
| | EXPINFLPBO3 | Pentachomous perceptions. Moderate rate is running average of inflation. (p.11) |
| | EXPINFLAPES | Pesaran's regression approach. (p.12) |
| | EXPINFLASEITZ | Time-varying stochastic parameter model for thresholds. (p.16) |

The relative performance of these measures is chiefly an empirical issue. However, it is possible to make a number of comments regarding 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 breaks down when one of the response categories attracts a zero share of the 'vote', as this directly contradicts the normality assumption made for the aggregate distribution.

**TABLE 7: FREQUENCY OF ZEROES IN EC MONTHLY SURVEY DATA**

| | Perceptions data | | | | Expectations data | | |
|---|---|---|---|---|---|---|---|
| Country | 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 |

[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.

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 when it assumes constancy for response thresholds belonging to the perceived inflation distribution, 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 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, 2003).

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 militates 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 on the legitimacy of estimating the relationship between official inflation data and the survey data of question 5 before applying that 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 in the indifference thresholds seriously. However, the random walk assumption for $\beta_t$ 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 perception 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.

## 4 COMPARING METHODS ON THE BASIS OF PREDICTIVE ABILITY

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 the non-parametric predictive performance test statistic of 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 a world of rational expectations – 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.

### 4.1 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.

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. Inflation expectations data is 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 under performs 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 ^ 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:

$$MSE = \frac{1}{T}\sum_{t=1}^{T} e_t^2 = (\bar{\hat{y}} - \bar{y})^2 + (\hat{\sigma} - r_{y\hat{y}}\sigma)^2 + (1 - r_{y\hat{y}}^2)\sigma^2 \tag{30}$$

Where $r_{y\hat{y}}$ denotes the correlation coefficient between predicted and actual values, while $\hat{\sigma}$ and $\sigma$ correspond to the standard deviations of predictions and observations, repectively. 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

| | France | | Spain | | Belgium | | Germany | |
|---|---|---|---|---|---|---|---|---|
| **Measure** | % | **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 |
| **Nº Observations** | 220 | | 218 | | 227 | | 286 | |

| | Italy | | United Kingdom | | Sweden | | Netherlands | |
|---|---|---|---|---|---|---|---|---|
| **Measure** | % | **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 |
| **Nº 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 over performance relative to a fair coin flip. In fact, the concern is actually the opposite – in many cases, the measures significantly *under* perform chance according to binomial tests[33]. Table 13 summarises the information above. On balance, EXPINFLPBIS, EXPINFLPCP and EXPINFLAPES are the best performers.

[33] 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 in 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 *under perform* (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 comparison of our quantification methodologies with the two sources of comparable quantitative expectations data available in Europe, with the aim of making a sounder assessment in Section 6. These are first presented in the next section.

# 5 QUANTITATIVE SURVEY DATA

## 5.1 United Kingdom: the Gallup Survey

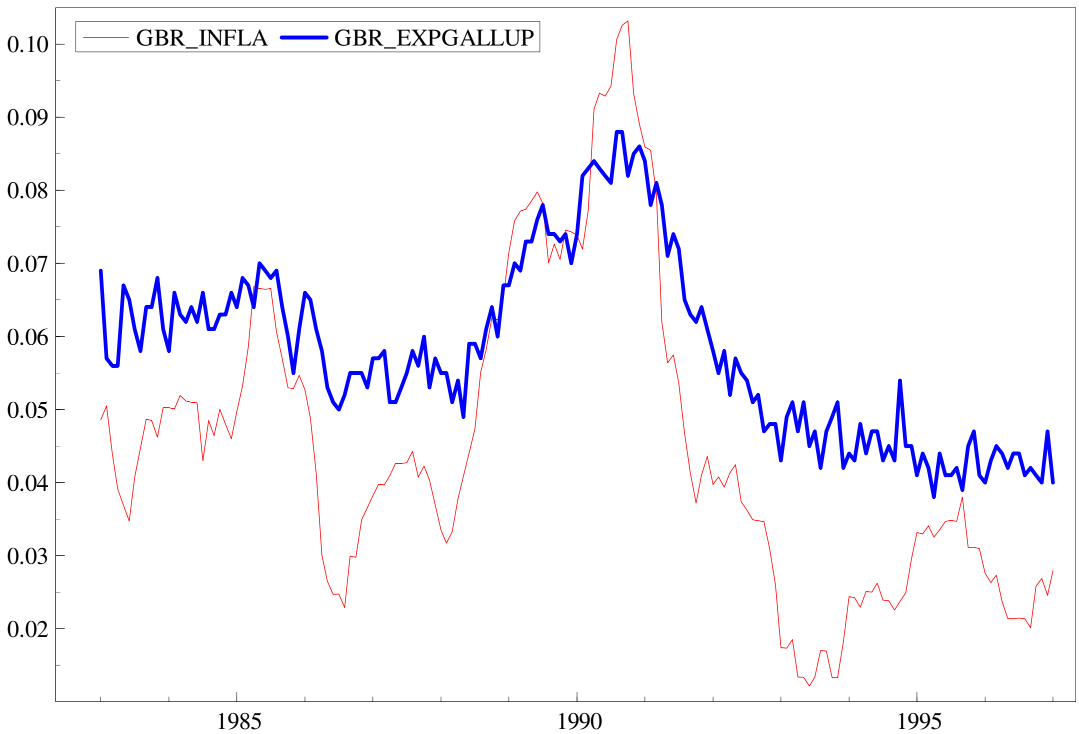
*Social Surveys (Gallup Poll) Ltd* carried out a monthly survey on a sample of employees in the United Kingdom continuously between January 1983 and January 1997. The results of this were published in the *Gallup Political Index* reports, subsequently renamed *Gallup Political and Economic Index* reports.

The particularity 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 10:

**FIGURE 10: 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 period of the 1990s to lower these. The graph also makes it clear that expectations exhibit extremely important 'adaptive' characteristics.

## 5.2 Sweden: the HIP Survey

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 expectations[34] (in addition to the quantitative ones) from January 1996.

The quantitative questions included in the HIP survey have the following phrasing[35]:

**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 consisting of individuals aged 16 through 84[36].

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 remains unexplained. In the case of inflation expectations, the authors suggest 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 survey[37].

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

[34] See TABLE 1, page 4.
[35] See Palmqvist and Strömberg (2004) for further details on the practical implementation of the survey
[36] See GfK (2003), Consumer Survey 2003.
[37] 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 11: HIP PERCEPTIONS vs ACTUAL INFLATION, SWEDEN, 1979q1-2005q3**

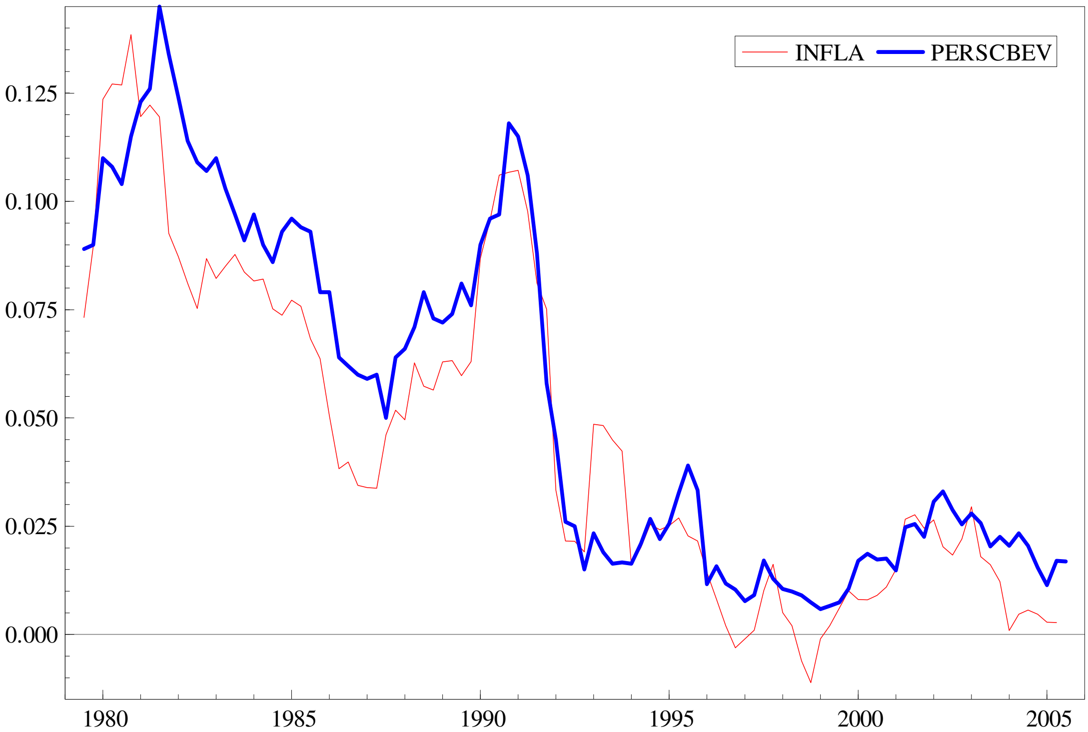
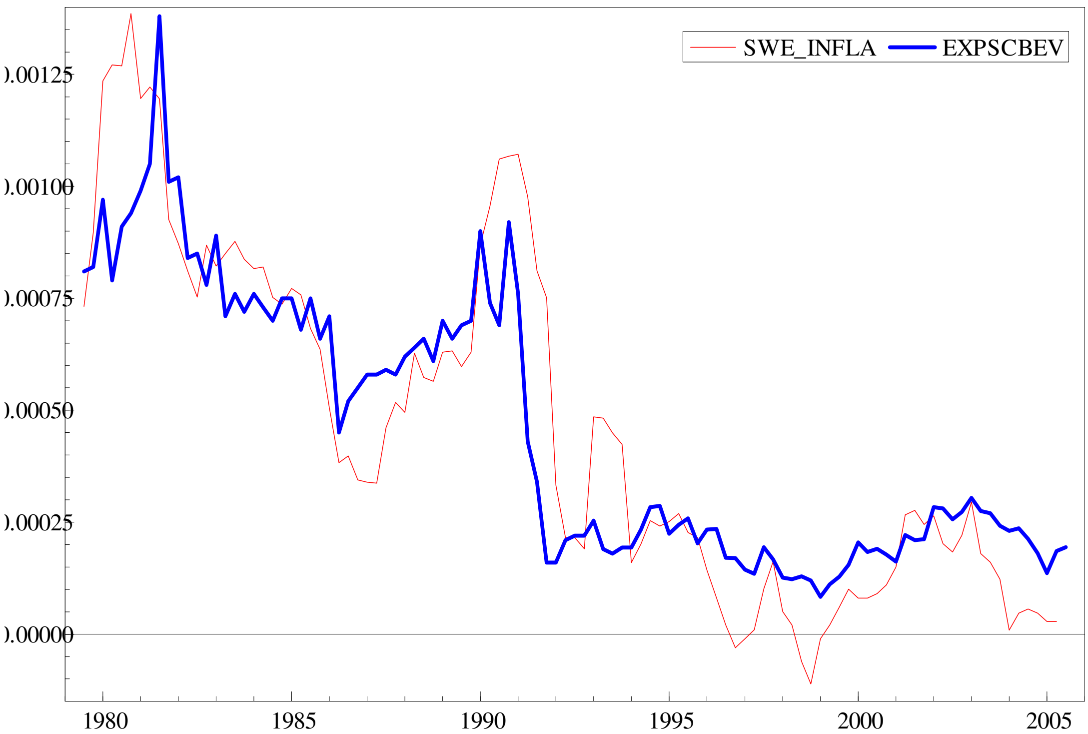



**FIGURE 12: HIP EXPECTATIONS vs ACTUAL INFLATION, SWEDEN, 1979q1-2005q3**



# 6 BENCHMARKING EC 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 will 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 is 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 sample[38] 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 begs 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: 1996m1-2005m8) are very different in terms of inflation behaviour – low

[38] The Bank of England/NOP Inflation Attitudes Survey also surveys the expectations of the general public, but is only available quarterly since February 2001.

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 |

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

| | UK | | | | | Sweden | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Measure | 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**

| | UK | | Sweden | |
|---|---|---|---|---|
| **Measure** | % | **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 |
| **Nº Observations** | 168 | | 114 | |

It is interesting to note that for the United Kingdom, none of the measures is 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 presents 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 variable[39]. 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 paper.

[39] 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 variables:

**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. Second, the constant in the equations is significant, confirming the earlier results that the measures are biased relative to the quantitative data. This is however not problematic for the use of these measures as controls in regression equations. 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. Interestingly, in neither case can homogeneity of first degree be rejected for the equations. Finally, we note that only measures in the Carlson-Parkin and Seitz traditions are retained in *both* equations. Furthermore, it can be seen that the coefficient on the Carlson-Parkin measure displays a significantly larger weight in both equations relative to the others.

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 or 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 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, as compared to our 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. Sections 7.2 and 7.3 clarify the construction of the data and the econometric approach taken. The results are presented in section 7.4.

### 7.1 Model specification and sign priors

We consider the wage equation with an equilibrium correction structure proposed in Curto Millet (2004). 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.

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 22 lists our sign priors.

**TABLE 21: WAGE EQUATION STRUCTURE**

$$\begin{aligned}
\Delta \ln COMPTH_t = {} & \alpha + a_1 INFLEXP_{t,t+4} + a_2 INFLEXP_{t-1,t+3} + a_3 INFLEXP_{t-2,t+2} \\
& + a_4 INFLEXP_{t-3,t+1} + a_5 INFLEXP_{t-4,t} \\
& + b_1 \Delta lCPI_t + b_2 \Delta lCPI_{t-1} + b_3 \Delta lCPI_{t-2} + b_4 \Delta lCPI_{t-3} \\
& + c_1 \Delta \ln COMPTH_{t-1} + c_2 \Delta \ln COMPTH_{t-2} + c_3 \Delta \ln COMPTH_{t-3} \\
& + c_4 \Delta \ln COMPTH_{t-4} + d_1 \Delta \ln U_t \qquad (31) \\
& + e_1 [\ln(\frac{COMPTH}{P_c})_{t-1} - f_1 \ln PRODIT_{t-1} \\
& - h_1 \ln(1 - ITAX)_{t-1} - i_1 \ln(1 + CTAX)_{t-1} \\
& - j_1 \ln U_{t-1} - k_1 \ln BRR_{t-1} - l_1 ACOV_{t-1} - m_1 \ln EPL_{t-1} \\
& - n_1 \ln UDENS_{t-1} - o_1 \ln BD_{t-1} - p_1 lRELPMPC_{t-1} - q_1 lCOORD]
\end{aligned}$$

| | | | |
|---|---|---|---|
| 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 ratio | | |

**TABLE 22: 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. |
| $c_i \leq 0$ $or$ $\geq 0$ for all i (variable) | Eq.(31) 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[40]. 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[41]. 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. |

[40] This shifts the Bargained Real Wage (BRW) curve up in the Carlin and Soskice (1990) framework
[41] This shifts the Price Determined Real Wage Curve (PRW) down in the Carlin and Soskice (1990) framework

| INSTITUTIONAL VARIABLES | |
|---|---|
| $k_i \geq 0 \;\; o_i \geq 0$ | Higher benefit replacement ratios and duration make unemployment less unpalatable and lead to tougher union bargaining. |
| $l_i \geq 0$ | Higher bargaining coverage of unions means 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. Possibly nonlinear relation (Calmfors and Driffill, 1988). |

### 7.2 Some explanatory notes on our data

Curto Millet (2006) provides a full explanation concerning the data used and its construction. We provide here only a few key details to enhance understanding of the results.

**General variables**

The dependent variable is the hourly compensation of *workers* in the economy. This is computed on the basis of the aggregate compensation of employees from the OECD's Economic Outlook 77, which we adjust for self-employed workers following Batini et al (2000). Total hours worked in the economy can be extracted from the OECD Productivity Database. The dependent variable can then be derived by dividing the aggregate compensation measure by the total hours worked in the economy.

It is reasonable to expect that short-term variations in productivity will not overly impact wage bargains and that consequently productivity trends are more relevant measures. Thus, our productivity measure is obtained by applying a Hodrick Prescott filter with lambda parameter 1600[42] to a productivity index provided by the OECD.

**Tax 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 are relevant here. Our data in this respect is an extension of the CEP-OECD dataset. The underlying data are taken from the OECD Revenue Statistics and OECD Economic Outlook.

The **direct tax rate** is obtained as:

$$ITAX = \frac{WC + IT}{HCR} \tag{32}$$

[42] This is standard for a quarterly frequency.

Its components are employees' social security contributions (WC), income taxes (IT) and household current receipts (HCR).

The **indirect tax rate** is obtained as:

$$CTAX = \frac{TX - SB}{CC} \tag{33}$$

That is, indirect taxes (TX) minus subsidies (SB) divided by private final consumption expenditure (CC).

### Institutional variables

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

$$BD = \alpha \frac{BRR_2}{BRR_1} + (1-\alpha)\frac{BRR_4}{BRR_1} \tag{34}$$

BRR1 is the unemployment 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 ($\alpha = 0.6$).

The **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 **bargaining coverage variable (ACOV)** is *adjusted* so as to refer the number of employees covered by a collective agreement as a percentage of employees *equipped with the right to bargain* (Traxler, 1994).

To measure **bargaining coordination (COORD)**, we extended an index compiled by Ochel (2000)[43]. The index comprises the following dimensions, measured on the scales provided in parentheses:

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

- 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. The summary measure computed from the above indices has a scale (1-3).

The **Employment Protection Legislation (EPL)** indicator is the result of chaining a multidimensional index from the OECD Employment Outlook 2004 with data on severance pay from Lazear (1990) for the period previous to 1985.

### 7.3 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 Generating 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 Generating 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.

We take the model presented in Eq.(31) 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 temporal position of the ECM 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.(31), if $e_1 = b_2 = b_3$, this would suggest an ECM lagged three times. We are also especially attentive at the possibility of restrictions that would suggest fourth difference effects in the dynamics. For instance, $b_1 = b_2 = b_3 = b_4$ would suggest a $\Delta_4 lCPI_t$ 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 the 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 further delta effects. In the actual modelling, further such lags have been considered for the unemployment variable. Other delta effects (e.g. institutional variables) were not considered, as theoretically they appear to be more naturally features of the equilibrium than affecting the wage bargain in the short-run. 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.

### 7.4 Results

We have carried out 64 general-to-specific model selection procedures in total[44]. Note that all the equations obtained as a result are congruent excepting cases mentioned explicitly. We will start by commenting on the summary results, before providing a discussion of the specific results in each country.

Table 23 indicates the retained lags of the expectations variable (if any) and their significance.

**TABLE 23: 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**] | - | - |

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.

44 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.

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 notable 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 further comments on the individual country results. The Swedish specification is presented in full detail in the next sub-section. For the benefit of concision, the remaining set of results are presented descriptively, but can be found in their entirety in Curto Millet (2006).

### 7.4.1 Sweden

**TABLE 24: RESULTS FROM THE SWEDISH WAGE EQUATION**

**Modelling SWE_D12lCOMPTH; Estimation Sample: 1983q1-2005q1**

| | Coefficient | Std.Error | t-value | t-prob | Part.R^2 |
|---|---|---|---|---|---|
| Constant | -13.5521 | 2.230 | -6.08 | 0.000 | 0.3454 |
| SWE_D4lCOMPTH_1 | 0.321190 | 0.06271 | 5.12 | 0.000 | 0.2726 |
| SWE_EXPSCBEV | 0.314704 | 0.1296 | 2.43 | 0.018 | 0.0777 |
| SWE_DlCPI | 0.405203 | 0.1492 | 2.72 | 0.008 | 0.0953 |
| SWE_DlCPI_1 | 0.364944 | 0.1362 | 2.68 | 0.009 | 0.0930 |
| SWE_DlCPI_2 | 0.490487 | 0.1359 | 3.61 | 0.001 | 0.1568 |
| SWE_DlCPI_3 | 0.587426 | 0.1380 | 4.26 | 0.000 | 0.2056 |
| SWE_D4lCPI_4 | 0.832525 | 0.07115 | 11.7 | 0.000 | 0.6617 |
| SWE_D4lCPI_8 | 0.840667 | 0.09915 | 8.48 | 0.000 | 0.5067 |
| SWE_D4lU | -0.0182446 | 0.009260 | -1.97 | 0.053 | 0.0525 |
| SWE_D4lU_4 | -0.0274556 | 0.01122 | -2.45 | 0.017 | 0.0787 |
| SWE_D4lU_8 | -0.0717109 | 0.007928 | -9.05 | 0.000 | 0.5389 |
| I1999 | -0.0212432 | 0.004866 | -4.37 | 0.000 | 0.2140 |
| I1995:1 | -0.0235829 | 0.008237 | -2.86 | 0.006 | 0.1048 |
| I2002:1 | 0.0215551 | 0.008209 | 2.63 | 0.011 | 0.0897 |
| SWE_lRCOMPTH_12 | -0.718932 | 0.05035 | -14.3 | 0.000 | 0.7444 |
| SWE_lPRODIT_12 | 0.540040 | 0.07743 | 6.97 | 0.000 | 0.4100 |
| SWE_lU_12 | -0.0715063 | 0.009462 | -7.56 | 0.000 | 0.4493 |
| SWE_lACOV_12 | 2.53557 | 0.5299 | 4.78 | 0.000 | 0.2465 |

| | | | |
|---|---|---|---|
| sigma | 0.00762516 | RSS | 0.00407001202 |
| R^2 | 0.988069 | F(18,70) = | 322.1 [0.000]** |
| log-likelihood | 318.392 | DW | 1.81 |
| no. of observations | 89 | no. of parameters | 19 |
| mean(SWE_D12lCOMPTH) | 0.166093 | var(SWE_D12lCOMPTH) | 0.00383307 |

**Specification Tests**

| | | | |
|---|---|---|---|
| AR 1-5 test: | F(5,65) = | 1.4472 | [0.2194] |
| ARCH 1-4 test: | F(4,62) = | 0.57369 | [0.6827] |
| Normality test: | Chi^2(2) = | 2.1774 | [0.3367] |
| hetero test: | F(33,36) = | 0.72145 | [0.8268] |
| RESET test: | F(1,69) = | 0.33559 | [0.5643] |

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, 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.

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 (2004) the current formalised system of adopted a 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 coefficient is found to be 72%. 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 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 1999 that was set in the 1998 bargaining round (EIRO, 1999). 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% 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. 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.

As stated earlier, the remaining set of results are presented descriptively, but can be found in their entirety in Curto Millet (2006).

### 7.4.2 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 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 is 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).

### 7.4.3 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 suggested in Curto Millet (2004).

The remarkable feature of the selected French models is the entire 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 these were 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 and significant, the coefficient magnitudes involved are not particularly strong, especially when compared to other countries such as Sweden which *does* exhibit a significant role for expectations. This suggests that a different interpretation may be relevant here.

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 of the economic situation.

A much more straightforward and perhaps more plausible explanation is simply that the inflation expectations data is of significantly worse quality than that of the other European countries given the measurement issue exposed in Section 2 due to 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.

#### 7.4.4 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 is of lower quality than that of the other countries, as it was only available at a half-yearly frequency[45]. 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 for some of the equations.

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 benefit replacement ratios but drops unemployment. The equations for EXPINFLPBIS and EXPINFLPBO2 are similar and have a more conventional long-run but simpler expectations dynamics. The EXPINFLPBIS model appears marginally more sensible, as the four quarter-lagged expectations of EXPINFLPBO2 seems rather long.

[45] Refer to Curto Millet (2006) for details of its construction

### 7.4.5 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 the unemployment, the benefit replacement ratio, union density and the relative price of imports to the CPI. In all cases, long-run homogeneity for productivity can be imposed. Dynamic homogeneity can also be imposed for the expectations coefficient or the sum thereof when there are several.

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.

### 7.4.6 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:

$$MA[abs(D4lCPI - D4lCPI\_4)]$$

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 the consumption tax, thereby making producer prices the relevant price concept in equilibrium. The models also include significant unemployment, coordination and volatility effects, as well as a benefit replacement rate or benefit duration term.

EXPINFLPBO3 has a more interesting lag structure of expectations, although this comes at the expense of actual inflation dynamics, which can be retained in the EXPINFLPBIS equation.

### 7.4.7 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 EXPINFLPCP, EXPINFLPBIS and EXPINFLPBO1. Long-run homogeneity can be imposed for productivity as well as for the consumption tax variable, thereby making producer prices the relevant equilibrium prices. 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 relatively 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%.

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%).

### 7.4.8 Germany

We model the half-yearly change in hourly compensation with an ECM term lagged 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 of both taxation variables, as well as the presence of the benefit replacement rate and bargaining coverage. Unemployment has an insignificant long-run effect but rather 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 reasonable speeds of adjustment given a usual duration of collective agreements of 1-2 years in Germany (European Commission, 2004) and the imposition of long-run homogeneity for productivity is possible. 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 paper 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 approaches 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 under perform 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, although 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 what regards the long-run equilibrium. Both measures are modifications of the Carlson-Parkin tradition suggested here and used in Curto Millet (2004) for the former. 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 to the difference in wording of the French questionnaire bears some responsibility, suggesting that special care should be taken in the quantification and use of this 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 to 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 article, 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 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.

Department of Economics, Oxford University, 21 January 2006.

## REFERENCES

Aron, J., J. Muellbauer, et al. (2004). Modelling the Inflation Process in South Africa. Keynote Address, Eighth Annual Conference on Econometric Modelling for Africa, Stellenbosch University, South Africa.

Bakhshi, H. and A. Yates (1998). "Are UK Inflation Expectations Rational?" Bank of England, Working Paper No81.

Bank of Canada (1998). Information in Financial Asset Prices. Conference, Ottawa, Bank of Canada.

Bank of England (2005). Inflation report, November 2005.

Batchelor, R. A. (1986). "Quantitative v. Qualitative Measures of Inflation Expectations." Oxford Bulletin of Economics and Statistics 48(2): 99-120.

Batchelor, R. A. and P. Dua (1987). "The Accuracy and Rationality of UK Inflation Expectations: Some Quantitative Evidence." Applied Economics 19: 819-828.

Batchelor, R. A. and A. B. Orr (1988). "Inflation Expectations Revisited." Economica 55: 317-331.

Batini, N., B. Jackson, et al. (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."

Bell, B. and N. Dryden (1996). The CEP-OECD Data Set (1950-1992). London, Centre for Economic Performance (CEP), London School of Economics.

Berk, J. M. (1999). "Measuring Inflation Expectations: A Survey Data Approach." Applied Economics 31: 1467-1480.

Berk, J. M. (2000). "Consumers' Inflation Expectations and Monetary Policy in Europe." Research Memorandum, Vrije Universiteit, Amsterdam, 2000-20, June.

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.

Bryan, M. F. and S. Palmqvist (2005). "Testing Near-Rationality Using Detailed Survey Data." European Economy(228).

Bryan, M. F. and G. Venkatu (2001). The Demographics of Household Inflation Surveys. Economic Commentary, Federal Reserve Bank of Cleveland.

Carlin, W. and D. Soskice (1990). Macroeconomics and the Wage Bargain.

Carlson, J. A. and M. Parkin (1975). "Inflation Expectations." Economica 42(May): 123-138.

Carroll, C. D. (2003). "Macroeconomic Expectations of Households and Professional Forecasters." Quarterly Journal of Economics(February).

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, et al. (1998). "Testing for a Unit Root in Variables with a Double Change in the Mean." Economics Letters 59: 175-182.

Clements, M. P. and D. F. Hendry (1999). Forecasting Non-stationary Economic Time Series.

Cooley, T. F. and E. C. Prescott (1973). "Varying Parameter Regression: A Theory and Some Applications." Annals of Economic and Social Measurement 2: 463-473.

Cooley, T. F. and E. C. Prescott (1976). "Estimation in the Presence of Stochastic Parameter Variation." Econometrica 44: 167-184.

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). "Labour Markets, Inflation Expectations and EMU." University of Oxford DPhil Thesis.

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): 391-400.

DG ECFIN (European Commission) (2004). "The Joint Harmonised EU Programme of Business and Consumer Surveys, User Guide."

EIRO (1997-2004). "EIRO National Reports." (various countries).

Ericsson, N. R. and J. S. Irons (1995). The Lucas Critique in Practice: Theory Without Measurement. Macroeconometrics: Developments, Tensions and Prospects. K. D. Hoover. Dordrecht, Kluwer Academic Press.

European Commission (2004). Industrial Relations in Europe 2004.

Evans, G. W. and S. Honkapohja (2001). Learning and Expectations in Macroeconomics, Princeton University Press.

Forder, J. (2000). "Could Reputation-Bias be a Bigger Problem than Inflation-Bias?" Oxford University, Department of Economics Working Paper No22.

Fregert, K. and L. Jonung (1998). "Monetary Regimes and Endogenous Wage Contracts: Sweden 1908-1995."

Gerberding, C. (2001). "The Information Content of Survey Data on Expected Price Developments for Monetary Policy." Deutsche Bundesbank ERC Discussion Paper 9/01.

GfK (2003). Consumer Survey 2003, GfK Sverige AB.

Harvey, A. C. (1989). Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge, Cambridge University Press.

Hendry, D. F. (1995). Dynamic Econometrics. Oxford, Oxford University Press.

Hendry, D. F. (2001). "Modelling UK Inflation 1875-1991." Journal of Applied Econometrics 16: 255-275.

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): 1-42.

INSEE (2004). "Enquête Mensuelle de Conjoncture Auprès des Ménages (Note Méthodologique)."

Konjunkturinstitutet Hushållens Inköpsplaner (Consumer Survey, User Manual).

Lazear, E. (1990). "Job Security Provision and Employment." Quarterly Journal of Economics.

Lyziak, T. (2003). "Consumer Inflation Expectations in Poland." ECB Working Paper No. 287.

Mankiw, N. G., R. Reis, et al. (2003). "Disagreement about Inflation Expectations." NBER Working Paper No.9796.

Nickell, S. (2003). "Employment and Taxes." Tax Policy and Employment Conference, CESifo.

Nickell, S. and L. Nunziata (2001). "Labour Market Institutions Database."

Nielsen, H. (2003a). "Inflation Expectations in the EU - Results from Survey Data." Discussion Paper No. 13/2003(Sonderforschungsbereich 373, Humboldt Universität zu Berlin).

Nolte, I. and W. Pohlmeier (2004). "Using Forecasts of Forecasters to Forecast."

Ochel, W. (2000). "Collective Bargaining (Centralization/Co-ordination)." Ifo Institute.

OECD (2004). OECD Employment Outlook 2004.

Osgood, C. E. (1953). Method and Theory in Experimental Psychology. New York, Oxford University Press.

Palmqvist, S. and L. Strömberg (2004). "Households' Inflation Opinions - a Tale of Two Surveys." Sveriges Riksbank Economic Review 4: 23-42.

Papadia, F. and V. Basano (1981). "EEC-DG II Inflationary Expectations. Survey Based Inflationary Expectations for the EEC Countries." DG ECFIN Economic Papers.

Pesaran, M. H. (1984). Expectations Formation and Macroeconometric Modelling. Contemporary Macroeconomic Modelling. P. Malgrange and P. A. Muet. Oxford, Blackwell.

Pesaran, M. H. (1985). "Formation of Inflation Expectations in British Manufacturing Industries." Economic Journal 95(December 1985): 948-975.

Pesaran, M. H. (1987). The limits to rational expectations. Oxford, UK, Basic Blackwell.

Pesaran, M. H. and A. Timmermann (1990). "A Simple, Non-Parametric Test of Predictive Performance." Cambridge Working Papers in Economics 9021.

Reckwerth, J. (1997). "Inflation and Output in Germany: The Role of Inflation Expectations." Bundesbank Discussion Paper No.5.

Roberts, J. M. (1997). "Is Inflation Sticky?" Journal of Monetary Economics 39: 173-196.

Seitz, H. (1988). "the Estimation of Inflation Forecasts from Business Survey Data." Applied Economics 20: 427-438.

Sekine, T. (2001). "Modeling and Forecasting Inflation in Japan." IMF Working Paper No. 01/82.

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: 165-185.

Social Surveys (Gallup Poll) Ltd (various). Gallup Political and Economic Index

Social Surveys (Gallup Poll) Ltd (various). "Gallup Political Index ".

Swedish Institute (2005). Labour Relations in Sweden.

Theil, H. (1952). "On the Time Shape of Economic Microvariables and the Munich Business Test." Revue de l'Institut International de Statistique 20: 105-120.

Visco, I. (1984). Price Expectations in Rising Inflation. Amsterdam, North Holland.