Economic Issues, Vol. 11, Part 2, 2006

Quantification of Expectations. Are They Useful for Forecasting Inflation?

Oscar Claveria, Ernest Pons, Jordi Suriñach[1]

ABSTRACT

Business tendency surveys are commonly used to provide estimations of a wide range of macroeconomic variables before the publication of official data. The qualitative nature of data on the direction of change has often led to quantifying survey results making use of official data, introducing a measurement error due to incorrect assumptions. Through Monte Carlo simulations it is possible to isolate the measurement error introduced by incorrect assumptions when quantifying survey results. By means of a simulation experiment we check the effect on the measurement error of respondents diverging from "rationality". We also analyse the predictive performance of different quantification methods for fourteen EU countries and the euro area. We find that allowing for asymmetric and stochastic response thresholds (indifference interval) produces a lower measurement error and more accurate forecasts.

1. INTRODUCTION

Business and consumer surveys provide detailed information about agents' expectations. The fact that survey results are based on the knowledge of the respondents operating in the market and are rapidly available makes them very valuable for forecasting purposes and decision-making. Survey results are presented as weighted percentages of respondents expecting a particular variable to go up, to go down or to remain unchanged. The qualitative nature of data on the direction of change has often led to quantifying survey results making use of official data.

The most common approach for quantifying survey expectations is assuming that respondents report a variable to go up or down if the mean of their subjective probability distribution lies above or below a threshold level (indifference interval). Carlson and Parkin (1975) suggested using a normal distribution together with symmetric and constant threshold parameters. Mitchell (2002) and Balcombe (1996) find evidence that normal distributions provide as accurate expectations as any other stable distribution.

Since survey data are approximations of unobservable expectations, they inevitably entail a measurement error. Several refinements have been proposed in order to reduce the measurement error introduced by incorrect assumptions (Mitchell *et al.* 2004; Löffler, 1999; Berk, 1999; Dasgupta and Lahiri, 1992; Seitz, 1988). Monte Carlo simulations can isolate the measurement error introduced when quantifying survey results, but there have been few attempts in the literature to compare quantification methods in a simulation context. Common (1985) and Nardo and Cabeza-Gutés (1999) analyse different quantification methods focusing on rational expectations testing rather than on their forecasting ability. Löffler (1999) estimates the measurement error introduced by the probabilistic method, and proposes a linear correction.

In this paper we design a simulation experiment in order to compare different quantification methods in terms of their forecasting performance, and to estimate the magnitude of the measurement error introduced by different assumptions. By means of Monte Carlo simulations we also check the effect on the measurement error introduced by the various procedures as respondents diverge from 'rationality'.

The paper also analyses the predictive performance of different quantification methods. Using survey data on price expectations for fourteen EU countries and the euro area, we find that allowing for an asymmetric and stochastic indifference interval results in lower measurement errors and produces more accurate forecasts than other quantification procedures.

The paper is organised as follows. The second section presents the various methods used for quantifying business survey data. Section three describes the simulation experiments. Section four describes the data and analyses the relative forecasting performance of the estimated series of expectations generated by the quantification methods described in Section two. Section five concludes.

2. QUANITIFICATION OF DATA ON THE DIRECTION OF CHANGE

Unlike other statistical series, survey results are weighted percentages of respondents expecting an economic variable to increase, decrease or remain constant. As a result, tendency surveys contain two pieces of independent information at time t, R_t and F_t, denoting the percentage of respondents at time t-1 expecting an economic variable to rise or fall at time t. The information therefore refers to the direction of change but not to its magnitude.

A variety of quantification methods have been proposed in the literature in order to convert qualitative data on the direction of change into a quantitative measure of agents' expectations. The output of these quantification procedures (estimated expectations) can be regarded as one period ahead forecasts of the quantitative variable under consideration. In this paper we apply the following quantification methods using agents' expectations about the future (prospective information):

2.1 Balance (BAL)

Anderson (1951) defined the balance statistic (R_t - F_t) as a measure of the average changes expected in the variable.

2.2 Anderson's Regression (AND)
Anderson (1952) was the first to formalise the relationship between actual changes in a variable and respondents' expectations. Let y_t be the actual average percentage change of an aggregate variable Y_t, and y_{it} the change for agent i, then up to a mean zero disturbance ξ_t:

$$y_t = \alpha R_t - \beta F_t + \xi_t \tag{1}$$

where α and β are two positive unknown parameters. OLS estimates of α and β are then used to obtain forecasts of y_t for one period ahead:

$$\hat{y}_{i+1} = \hat{\alpha} R_{t+1} - \hat{\beta} F_{t+1} \tag{2}$$

denoting R_{t+1} and F_{t+1} the percentage of firms at time t expecting industrial prices to rise and fall at time t+1, respectively.

2.3 Pesaran's Regression (PES)
Pesaran (1984) extended the above model to allow for an asymmetrical relationship between y_t and y_{it} in periods of increasing inflation:

$$y_t = (\alpha_1 R_t - \beta F_t)/(1 - \lambda_1 R_t) + \xi_t \tag{3}$$

where ξ_t is not necessarily homoschedastic and uncorrelated. The non linear estimates of the parameters in (3) are then used to derive estimates of y_t for period t+1:

$$\hat{y}_{t+1} = \frac{\hat{\alpha}_1 R_{t+1} - \hat{\beta} F_{t+1}}{1 - \hat{\lambda}_1 R_{t+1}}$$

2.4 Carlson-Parkin's probability approach (CP)
The probability approach was developed by Carlson-Parkin (1975) along the lines suggested by Theil (1952). The method is based on the assumption that each respondent i answers according to a subjective probability distribution which is conditional on the information set available up to that moment, Ω_{it}. As a result, they report 'no change' in $Y_{i,t+1}$ if their expectation $\hat{y}_{i,t+1}$ lies inside an indifference interval $(-a_{i,t+1}, b_{i,i+1})$, an increase if $\hat{y}_{i,t+1} \geq b_{i,t+1}$, and a fall if $\hat{y}_{i,t+1} \leq -a_{i,t+1}$. The response thresholds $a_{i,t+1}$ and $b_{i,t+1}$ are assumed to be symmetric and fixed both across firms and over time, $a_{i,t+1} = b_{i,t+1} = \lambda$, $\forall\, i,t$. If

the subjective probability distributions are assumed to be independent and to have the same form across respondents, an aggregate density function $f(y_{t+1},\Omega_t)$ can be derived, yielding:

$$\hat{y}_{t+1} = \lambda\left(\frac{f_{t+1}+r_{t+1}}{f_{t+1}-r_{t+1}}\right) \tag{5}$$

Carlson and Parkin (1975) assumed that f_i (.) were normally distributed and estimated λ by assuming that over the sample-period $\hat{y}_{t+1}$ is an unbiased estimate of y_t:

$$\lambda_{CP} = \left(\sum_{t=1}^{T} y_t\right) \Big/ \left(\sum_{t=1}^{T}\left(\frac{f_{t+1}+r_{t+1}}{f_{t+1}-r_{t+1}}\right)\right) \tag{6}$$

2.5. Carlson-Parkin's probability approach with an asymmetric indifference interval (ACP)

By relaxing the assumption that response thresholds are symmetric, equation (5) becomes:

$$\hat{y}_{t+1} = \frac{bf_{t+1}+ar_{t+1}}{f_{t+1}-r_{t+1}} \tag{7}$$

Parameters a and b are unknown; they can be estimated by regressing y_t on survey expectations:

$$y_t = bx_{bt} - ax_{at} + u_t \tag{8}$$

where $u_t \sim N(0,\sigma_u^2)$ and $x_{bt} = f_t/(f_t - r_t)$ and $x_{bt} = f_t/(f_t - r_t)$.

2.6. Berk's probability approach (BK)

Berk (1999) proposed estimating l in (5) as:

$$\lambda_{BK} = y_{t-1} + \left[\left(\sum_{t=1}^{T} y_t\right) \Big/ \left(\sum_{t=1}^{T}\left(\frac{f_{t+1}+r_{t+1}}{f_{t+1}-r_{t+1}}\right)\right)\right] \tag{9}$$

2.7. Seitz's probability approach (TVP)

By relaxing the assumption that thresholds a_{it} and b_{it} are fixed across time, equation (3) becomes:

$$\hat{y}_{t+1} = x_{t+1}\beta_{t+1}$$

where $x_{t+1} = \lfloor x_{bt} \quad x_{at} \rfloor$ and $\beta_{t+1} = \lfloor -a_{t+1} \quad b_{t+1} \rfloor$. In order to obtain estimates of β_t, Seitz (1988) used Cooley and Prescott's (1976) time-varying parameter model, assuming that the parameter vector was subject to permanent and temporary shocks:

$$\begin{aligned} \beta_t &= \beta_t^p + \varphi_t \\ \beta_t^p &= \beta_{t-1}^p + \zeta_t \end{aligned} \tag{11}$$

where β_t^p is the permanent component of the variation in the parameters, and φ_t is its temporary component. $\varphi_t \sim N(0,(1-\gamma)\sigma^2\Sigma)$ and $\zeta_t \sim N(0,\gamma\sigma^2\Sigma)$ are assumed to be mutually uncorrelated and independent.

2.8. State-Space approach with random walk response thresholds (SS1)

Instead of using Cooley and Prescott's time-varying parameter model to obtain estimates of y_{t+1} in (10), Claveria et al (2004) suggested a more general model that allows for asymmetric and dynamic response thresholds and that would include Seitz's method as a particular case. In this generalisation, response thresholds are generated by a random walk process:

$$y_t = b_t x_{bt} - a_t x_{at} + u_t \qquad u_t \sim N(0,\sigma_u^2) \tag{12}$$

$$\begin{cases} b_t = b_{t-1} + v_t \\ a_t = a_{t-1} + w_t \end{cases} \tag{13}$$

where v_t and w_t are two independent and normally distributed disturbances with mean zero and variance σ_v^2 and σ_w^2. The initial conditions are assumed to be zero.[2] The relationship between y_t and the response thresholds is linear and is expressed in the measurement equation (12). The unknown state is assumed to vary in time according to the linear transition equation (13). The Kalman filter is used to estimate the variances and to derive estimates of y_{t+1}.

2.9. State-Space approach with autoregressive response thresholds (SS2)

This generalisation, in which response thresholds are generated by a first-order Markov process, is based on the following state-space representation:

$$y_t = b_t x_{bt} - a_t x_{at} + u_t \qquad u_t \sim N(0,\sigma_u^2) \tag{14}$$

$$\begin{cases} b_t = \phi b_{t-1} + v_t \\ a_t = \rho a_{t-1} + w_t \end{cases} \tag{15}$$

where ϕ and ρ are the autoregressive parameters. As in (13), v_t and w_t are two independent and normally distributed disturbances with mean zero and variances σ_v^2 and σ_w^2. Supposing null initial conditions, the Kalman filter is used to estimate the variances and the autoregressive parameters.

3. THE SIMULATION EXPERIMENT

The differences between the actual values of a variable and quantified expectations may arise from three different sources (Lee, 1994): measurement or conversion error due to the use of quantification methods, expectational error due to the agents' limited ability to predict the movements of the actual variable, and sampling errors. Since survey data are approximations of unobservable expectations, they inevitably entail a measurement error.

Through Monte Carlo simulations it is possible to distinguish between these three sources of error and to isolate the measurement error introduced by incorrect assumptions when quantifying survey results, but there have been few attempts in the literature to compare quantification methods in a simulation context. The aim of this section is to design a simulation experiment that allows us to analyse the size of the measurement error of each estimated series of expectations and the forecasting performance of each of the quantification methods described in Section two.

3.1. Description of the experiment

In order to compare different quantification methods in terms of their forecasting performance, and to estimate the magnitude of the measurement error introduced by different assumptions, the experiment is designed in four consecutive steps:

(i) The simulation begins by generating a series of actual changes of a variable. We consider 50 agents and 200 time periods. Let y_{it} indicate the percentage change of variable Y_{it} for agent i from time t-1 to time t. Additionally we suppose that the true process behind the movement of y_{it} is given by:

$$y_{it} = d_{it} + \varepsilon_{it} \tag{16}$$

$i = 1,\dots,50$, $t = 1,\dots,200$ and $d_{it} = -0.05 + 0.9y_{i,t-1}$ where d_{it} is the deterministic component. The initial value, $y_{i0} = 0.9$, is assumed to be equal for all agents.[3] ε_{it} is an identical and independent normally distributed random variable with mean zero and variance $\sigma_\varepsilon^2 = 8$. The average rate of change, y_t, is given by $y_t = \frac{1}{50}\sum_i y_{it}$ The same weight is given to all agents.

(ii) Secondly, we generate a series of agents' expectations about y_t under the assumption that individuals are rational in Muth's sense:[4]

$$y_{it}^e = d_{it} + \zeta_{it} \qquad \zeta_{it} \sim N(0, \sigma_\zeta^2) \tag{17}$$

where y_{it}^e has the same deterministic part as y_{it} but a different stochastic term ζ_{it}. We derive $y_t^e = 1/50\sum_i y_{it}^e$ Additionally, we assume that $\sigma_{i\zeta}^2 = \sigma_\zeta^2 = 1$. All the values given to σ_ε^2 and σ_ζ^2, and to the indifference interval are set to simulate actual business survey series.

(iii) The third step consists of constructing the answers to the business surveys. The answers are given in terms of the direction of change, i.e., if the variable is expected to increase, decrease or remain unchanged. We assume that agents' answers deal with the next period and that all agents have the same constant indifference interval $(-a, b)$ with $a = b = 5$. If $y_{it}^e \geq 5$, agent i answers that Y_{it} will increase; if $y_{it}^e \leq -5$, i expects Y_{it} to decrease; while the agent will report no change if $-5 < y_{it}^e < 5$. With these answers, qualitative variables R_{it} and F_{it} can be constructed. R_{it} (F_{it}) takes the value 1 (0) whenever the agent expects an increase (decrease) in Y_{it}. R_t and F_t are then constructed by aggregation.

(iv) The fourth step of the simulation experiment consists of using different quantification methods to trace back the series of actual changes of the generated quantitative variable, y_t, from the qualitative variables. We will refer to these expectations as estimated expectations in order to distinguish them from those that are unobservable. With the aim of analysing the performance of the different proxy series, we use the last 100 generated observations. Keeping the series of actual changes fixed, the experiment of generating the rational expectations series as well as the proxy series is replicated 1000 times.[5]

To test the robustness of the results and to check the effect on the measurement error introduced by the various procedures as respondents diverge from 'rationality', we repeat the simulation experiment introducing a bias of 0.5 in the expectations formation process. This variation allows us to check the possible influence on the results of introducing an expectational error.[6]

3.2. Evaluation of the estimated expectations

In order to evaluate the relative performance and the forecasting accuracy of the different quantification procedures, we keep the series of actual changes fixed and we replicate the experiment of generating the rational expectations series as well as the qualitative variables R_t and F_t 1000 times. The specification of the quantification procedures is based on information up to the first 100 periods; models are then re-estimated each period and forecasts are computed in a recursive way. In each simulation, forecast errors for all methods are obtained for the last 100 periods. In order to summarise this information, the Root Mean Squared Error (RMSE), the Mean Error (ME) and the three components of the Mean Square Error (MSE) are calculated.

According to Theil (1971), the MSE may be decomposed to yield:

$$MSE = \frac{1}{T}\sum_{t=1}^{T} e_t^2 = (\bar{\hat{y}} - \bar{y})^2 + (\hat{\sigma} - r_{y\hat{y}}\sigma)^2 + (1 - r_{y\hat{y}}^2)\sigma^2 \qquad (18)$$

where $\bar{\hat{y}}$ and $\bar{y}$ are the mean of the forecast and actual values respectively, $\hat{\sigma}^2$ and σ^2 are the estimated variances of the forecasts and actual values, and is the correlation between the forecasts and the actual values. After dividing by MSE, and denoting the resulting ratios by *U*1, *U*2 and *U*3:

$$1 = U1 + U2 + U3 \tag{18}$$

where *U*1 is the proportion of the MSE due to bias in the forecasts (the mean error or bias proportion), *U*2 is the proportion of the MSE due to the difference between variances (the regression error or variance proportion) and may be interpreted as the proportion of error in the MSE which is unexplained (the disturbance error or covariance proportion). Their values provide useful information for analysing the forecasting performance of each method. One would therefore expect a good forecast to have low *U*1 and *U*2, and high *U3*. Table 1 shows the results of an off-sample evaluation for the last 100 periods.

Table 1: Forecast evaluation of simulated series

	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
RMSE	12.37	1.33	3.13	2.03	1.66	4.55	2.29	0.29	0.22
ME	-4.57	0.15	-0.01	-0.01	0.31	-2.33	1.28	0.00	0.01
% *U*1	14	3	0	0	6	32	41	0	1
% *U*2	86	42	92	53	55	58	42	6	34
% *U*3	0	55	8	47	39	10	17	94	65

Notes: (i) RMSE = root mean square error; (ii) ME = mean error; (iii) %U1 = percentage of mean error (bias proportion of the MSE); (iv) %U2 = percentage of regression error (variance proportion of the MSE); (v) %U3 = percentage of disturbance error (covariance proportion of the MSE).

Table 1 shows that in State Space models (*SS1* and *SS2*), both RMSE and ME are considerably lower. At the other end, we find the balance statistic (*BAL*), with very high RMSE and ME, together with a very high proportion of regression error. Relaxing the assumption of symmetric response thresholds (*ACP*) seems to improve the forecasting performance of the Carlson-Parkin (*CP*) estimator. Anderson's regression (*AND*) also outperforms Pesaran's regression (*PES*), both in terms of RMSE and ME and regarding the distribution of the MSE components.

Table 2 shows the results of the comparison for the last 100 periods after relaxing the assumption of rationality. Comparison of the results in Table 2 with those in Table 1 highlights several differences, in particular regarding the distribution of the MSE and the magnitude of the ME. After introducing a 0.5 bias, the ME and the proportion of systematic error (*U1*) show an increase in all methods, with the exception of the two State Space models (*SS1* and *SS2*). Although this may be a consequence of the recursive nature of the

Kalman filter, State Space models are less affected by expectational errors and therefore more adequate as agents diverge from rationality.

Table 2: Forecast evaluation of simulated seriesintroducing an expectational error

	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
RMSE	29.00	2.02	2.60	2.77	2.26	3.98	2.22	0.23	0.25
ME	-26.58	1.16	1.76	1.36	1.46	2.39	1.43	0.00	0.02
% *U1*	84	46	61	30	55	40	53	0	1
% *U2*	16	9	2	39	9	49	14	11	19
% *U3*	0	45	37	31	36	11	33	89	80

Notes: (i) RMSE = root mean square error; (ii) ME = mean error; (iii) %U1 = percentage of mean error (bias proportion of the MSE); (iv) %U2 = percentage of regression error (variance proportion of the MSE); (v) %U3 = percentage of disturbance error (covariance proportion of the MSE).

Though it is impossible to eliminate completely the measurement error introduced when converting qualitative data on the direction of change into quantitative estimations of agents' expectations, state-space models obtain lower measurement errors and better forecasts. These results suggest that introducing asymmetric and stochastic response thresholds in a more general model may improve quantification methods with forecasting purposes. In Section four we analyse whether an asymmetric and stochastic indifference interval produces a lower measurement error and more accurate forecasts for fourteen EU countries and the euro area.

4. Empirical Results

4.1. Data

The *Joint Harmonised EU Industry Survey* is conducted in all EU member countries on a monthly basis. It started in 1962 within the framework of the Joint Harmonised EU Programme, which included all six original member countries of the European Economic Community. The results are presented by the Commission in Supplement B of *European Economy,* which is usually published 30 days after the survey was conducted. Of the six monthly questions of the survey, we use the one referring to price expectations:

- 'Selling price expectations for the months ahead: up, unchanged, down?'

Survey responses on price expectations for the manufacturing industry have been used to track the evolution of the Producer Price Index (PPI) in fourteen EU countries and the euro area. Our sample period is 1991:1-2000:12.

In Tables 3 and 4 we present the results of the analysis of survey results for the euro area, which cover most of the EU member countries for the period in question. Table 3 shows some descriptive statistics for series R_t

and F_t for the euro area, being R_t the percentage of respondents expecting selling prices to go up and F_t the percentage of respondents expecting selling prices to go down.

Table 3: Descriptive statistics of series R_t and F_t

	Maximum	Minimum	Mean	Standard error	Variation coefficient
R_t	6.5	39.9	16.5	7.02	0.43
F_t	4.9	17.7	9.9	3.18	0.32

Table 4 presents the results of applying the Augmented Dickey-Fuller (ADF) unit root test to series R_t and F_t. Although R_t seems slightly less dispersed than F_t, both series are stationary. With the purpose of testing if both series R_t and F_t are integrated of second order, we first test the null hypothesis $H_0{:}I(2)$, with the alternative hypothesis $H_1{:}I(1)$. Then we test the null hypothesis $H_0{:}I(1)$, with the alternative hypothesis $H_1{:}I(0)$. As expected from the theory, in both cases the null hypothesis is rejected, therefore both series R_t and F_t can be considered stationary.

Table 4: Augmented Dickey-Fuller (ADF) unit roots test of series R_t and F_t

$D(R_t)$	-6.09***	R_t	-3.35**
$D(F_t)$	-4.84***	F_t	-3.00**

Notes: *** Indicates that the null hypothesis can be rejected at the 1% significance and ** indicates that the null hypothesis can be rejected at the 1% significance level.

4.2. Forecasting performance

In order to evaluate the relative forecasting accuracy of the different quantification procedures, all models are estimated until 1997:12 and forecasts for one period ahead are computed, therefore assuming that firms' answers deal with the next month. The specification of the models is based on information up to 1997:12 and models are then re-estimated each month and forecasts computed. Forecast errors for each method are computed in a recursive way from 1998:1 up to 2000:12. In order to summarise this information, the Root Mean Squared Error (RMSE), the Theil coefficient (TC) and the three components of the mean square error (MSE) for each estimated series are computed.

The results of the forecasting performance of the different methods in terms of RMSE and TC values are reported in Tables 5 and 6, where forecast errors are obtained from an off-sample evaluation from 1998:1 to 2000:12. As regards the forecast accuracy of the different methods, the state-space approach (SS1 and SS2) outperforms the other procedures for all fourteen countries. The third lowest RMSE is obtained with the TVP in seven countries. These results suggest that a stochastic indifference interval may improve the forecasting performance of quantification methods. When comparing the forecast accuracy shown by regression methods, the AND shows a better performance than the PES.

Table 5: Forecast evaluation in terms of RMSE

RMSE	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Austria	7.44	2.04	2.56	16.63	2.02	16.63	2.00	0.95	0.93
Belgium	7.27	4.36	5.91	1.86	4.34	1.86	4.04	1.59	1.74
Finland	10.76	4.89	5.67	3.46	4.97	2.33	4.82	1.59	1.81
France	13.04	1.56	1.98	5.05	1.54	5.29	1.34	0.57	0.59
Germany	7.13	1.34	1.72	1.12	1.32	0.84	1.36	0.54	0.61
Great Britain	17.20	1.29	1.87	10.49	1.06	10.92	1.06	0.38	0.40
Greece	9.48	3.88	3.75	3.32	3.69	4.17	2.95	1.35	1.32
Ireland	10.07	3.04	3.01	7.10	2.88	6.75	2.58	1.12	1.16
Italy	9.32	1.72	2.52	1.32	1.71	1.50	1.61	0.86	0.85
Luxembourg	18.40	3.93	5.11	3.90	4.02	4.26	3.76	1.78	1.88
Netherlands	5.19	5.34	7.16	6.04	5.33	4.45	4.58	2.10	2.14
Portugal	9.37	12.04	12.04	11.37	11.90	8.79	11.45	3.80	3.64
Spain	5.28	2.43	2.85	4.63	2.36	5.13	2.29	0.80	0.82
Sweden	8.54	1.34	2.69	2.11	1.32	2.03	1.38	1.00	1.09
Euro Area	5.77	2.17	2.72	1.74	2.13	1.21	2.12	0.75	0.82

Note: RMSE = root mean square error.

Tables 7 to 9 show the proportion of the MSE due to each one of its components: the bias proportion (*U1*), the variance proportion (*U2*) and the covariance proportion (*U3*). As mentioned above, the lower the *U1* and *U2*, and therefore the higher the *U3*, the better. The state-space approach (*SS1* and *SS2*) shows low values of *U1* and the lowest *U1* in two countries, the lowest *U2* in all countries but five, and the highest *U3* in all countries with the exception of Great Britain.

In addition, if we look at the descriptive statistics presented in Table 10 as a summary of the results of Tables 5 to 9, *SS1* and *SS2* are the methods that present the lowest standard deviation for all three components of the MSE. The distribution of MSE for the other methods, apart from being less desirable, shows enormous differences between countries.

Table 6: Forecast evaluation in terms of the Theil coefficient

TC	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Austria	0.44	0.34	0.58	0.80	0.33	0.80	0.32	0.03	0.03
Belgium	0.15	0.28	0.80	0.03	0.28	0.02	0.22	0.02	0.02
Finland	0.24	0.41	0.63	0.13	0.42	0.04	0.39	0.02	0.02
France	0.64	0.20	0.97	0.83	0.20	0.85	0.19	0.03	0.03
Germany	0.43	0.20	0.31	0.12	0.18	0.05	0.19	0.02	0.02
Great Britain	0.95	0.12	0.16	0.96	0.11	0.96	0.12	0.01	0.01
Greece	0.24	0.09	0.09	0.07	0.08	0.09	0.08	0.01	0.01
Ireland	0.51	0.38	0.41	0.59	0.33	0.57	0.23	0.02	0.03
Italy	0.33	0.07	0.14	0.04	0.07	0.03	0.06	0.01	0.01
Luxembourg	0.53	0.48	0.76	0.19	0.51	0.21	0.46	0.04	0.05
Netherlands	0.08	0.29	0.75	0.41	0.28	0.15	0.17	0.02	0.02
Portugal	0.17	0.59	0.58	0.49	0.56	0.20	0.47	0.02	0.02
Spain	0.22	0.19	0.25	0.19	0.16	0.21	0.16	0.01	0.02
Sweden	0.40	0.06	0.25	0.10	0.06	0.09	0.08	0.04	0.04
Euro Area	0.23	0.23	0.38	0.13	0.22	0.05	0.22	0.01	0.02

Note: TC = Theil coefficient.

Table 7: Forecast evaluation in terms of % *U1*

%U1	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Austria	39.79	11.11	2.55	10.47	9.83	11.48	10.46	4.41	8.27
Belgium	0.42	18.50	15.75	11.76	17.72	15.00	18.04	3.77	5.31
Finland	2.98	8.78	7.27	18.30	7.33	0.11	9.88	4.34	5.76
France	36.48	49.96	0.29	3.48	49.51	6.55	20.07	2.96	3.34
Germany	53.01	1.10	0.14	0.85	0.04	6.51	0.03	2.97	5.07
Great Britain	84.90	11.03	71.39	82.22	0.10	81.79	6.35	6.90	7.66
Greece	49.72	18.25	18.24	5.52	36.47	27.46	6.65	3.98	4.58
Ireland	39.17	20.52	27.57	0.03	19.46	0.00	16.88	3.64	4.08
Italy	41.77	1.29	2.38	0.01	5.33	33.07	1.50	8.67	8.95
Luxembourg	2.61	13.23	4.35	5.72	14.45	0.36	3.34	0.09	0.12
Netherlands	57.04	10.63	13.41	11.06	8.64	1.74	7.68	1.84	3.49
Portugal	0.40	13.03	13.76	14.49	11.51	7.76	19.61	5.15	7.15
Spain	5.28	0.42	0.16	4.44	5.08	7.83	2.10	5.70	7.41
Sweden	0.04	3.42	17.37	7.45	1.98	0.27	2.23	1.28	1.41
Euro Area	7.73	8.50	0.97	20.98	6.93	4.90	7.53	4.90	6.84

Note: *%U1* = percentage of mean error (bias proportion of the MSE).

Table 8: Forecast evaluation in terms of % *U2*

%U2	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Austria	47.61	70.85	88.80	53.06	72.76	52.32	10.46	3.33	7.31
Belgium	79.75	76.57	80.65	24.85	77.07	1.31	18.04	3.59	2.33
Finland	79.79	87.36	86.61	52.90	88.45	28.27	9.88	3.24	1.55
France	60.50	26.43	69.94	11.88	30.87	12.12	20.07	3.10	3.18
Germany	42.19	83.70	84.11	73.94	87.35	47.64	0.03	1.82	2.91
Great Britain	11.54	4.16	9.50	8.33	7.06	8.76	6.35	2.27	0.39
Greece	23.26	0.26	0.27	2.99	1.74	14.64	6.65	0.01	0.37
Ireland	46.25	62.81	58.15	9.27	75.96	6.75	16.88	0.72	0.65
Italy	52.34	64.53	76.50	28.31	64.34	7.79	1.50	3.75	3.76
Luxembourg	79.57	66.69	39.15	0.18	67.40	1.49	3.34	1.93	0.30
Netherlands	1.52	81.55	81.13	75.61	83.21	65.75	7.68	0.90	0.95
Portugal	32.51	78.47	79.07	71.18	78.58	61.13	19.61	0.75	0.90
Spain	67.72	66.51	70.75	64.73	74.27	66.81	2.10	1.40	1.79
Sweden	90.05	0.53	59.22	36.11	5.50	47.86	2.23	2.78	1.06
Euro Area	84.27	80.91	90.38	63.76	83.49	47.69	81.96	2.46	2.91

Note: *%U2* = percentage of regression error (variance proportion of the MSE).

Table 9: Forecast evaluation in terms of % *U3*

%U3	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Austria	47.61	70.85	88.80	53.06	72.76	52.32	10.46	3.33	7.31
Belgium	79.75	76.57	80.65	24.85	77.07	1.31	18.04	3.59	2.33
Finland	79.79	87.36	86.61	52.90	88.45	28.27	9.88	3.24	1.55
France	60.50	26.43	69.94	11.88	30.87	12.12	20.07	3.10	3.18
Germany	42.19	83.70	84.11	73.94	87.35	47.64	0.03	1.82	2.91
Great Britain	11.54	4.16	9.50	8.33	7.06	8.76	6.35	2.27	0.39
Greece	23.26	0.26	0.27	2.99	1.74	14.64	6.65	0.01	0.37
Ireland	46.25	62.81	58.15	9.27	75.96	6.75	16.88	0.72	0.65
Italy	52.34	64.53	76.50	28.31	64.34	7.79	1.50	3.75	3.76
Luxembourg	79.57	66.69	39.15	0.18	67.40	1.49	3.34	1.93	0.30
Netherlands	1.52	81.55	81.13	75.61	83.21	65.75	7.68	0.90	0.95
Portugal	32.51	78.47	79.07	71.18	78.58	61.13	19.61	0.75	0.90
Spain	67.72	66.51	70.75	64.73	74.27	66.81	2.10	1.40	1.79
Sweden	90.05	0.53	59.22	36.11	5.50	47.86	2.23	2.78	1.06
Euro Area	84.27	80.91	90.38	63.76	83.49	47.69	81.96	2.46	2.91

Note: *%U3* = percentage of disturbance error (covariance proportion of the MSE).

Table 10: Forecast evaluation: descriptive statistics

RMSE	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Mean	9.89	3.51	4.20	5.35	3.46	3.51	3.23	1.32	1.36
Min.	5.19	1.29	1.72	0.84	1.06	1.29	1.06	0.38	0.40
Max.	18.40	12.04	12.04	16.63	11.90	12.04	11.45	3.80	3.64
Std. Dev.	3.95	2.83	2.83	4.32	2.82	2.83	2.68	0.87	0.84
TC	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Mean	0.38	0.26	0.48	0.35	0.25	0.30	0.23	0.02	0.02
Min.	0.08	0.06	0.09	0.03	0.06	0.02	0.06	0.01	0.01
Max.	0.95	0.59	0.97	0.96	0.56	0.96	0.47	0.04	0.05
Std. Dev.	0.23	0.16	0.28	0.32	0.16	0.34	0.14	0.01	0.01
% U1	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Mean	29.54	12.95	13.90	14.28	13.39	12.95	8.92	3.98	5.19
Min.	0.04	0.42	0.14	0.00	0.04	0.42	0.03	0.09	0.12
Max.	84.90	49.96	71.39	81.79	49.51	49.96	20.07	8.67	8.95
Std. Dev.	27.40	12.53	18.58	21.92	14.11	12.53	7.12	2.23	2.58
% U2	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Mean	51.04	55.03	63.13	30.19	58.18	55.03	56.13	2.11	1.96
Min.	1.52	0.26	0.27	1.31	1.74	0.26	0.50	0.01	0.30
Max.	90.05	87.36	88.80	66.81	88.45	87.36	90.32	3.75	7.31
Std. Dev.	27.17	32.38	28.10	25.43	32.10	32.38	29.54	1.22	1.90
% U3	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Mean	19.41	32.02	22.97	55.53	28.43	32.02	34.95	93.91	92.86
Min.	3.02	3.86	3.60	9.45	4.21	3.86	6.83	87.58	84.43
Max.	67.09	96.05	81.48	98.15	92.84	96.05	92.85	97.98	99.59
Std. Dev.	17.45	31.54	21.83	27.20	30.93	31.54	30.94	2.75	3.87

Notes: (i) RMSE = root mean square error.
(ii) TC = Theil coefficient.
(iii) *%U1* = percentage of mean error (bias proportion of the MSE).
(iv) *%U2* = percentage of regression error (variance proportion of the MSE).
(v) *%U3* = percentage of disturbance error (covariance proportion of the MSE).

Finally, in Table 11, we rank the quantification procedures based on RMSE, TC, *%U1*, *%U2* and *%U3* values. The considerably lower RMSE and TC values obtained by the state-space representation in all countries, together with a preferable distribution of MSE components, suggest that the introduction of asymmetric and stochastic response thresholds in a more general model may be useful when using quantification methods with forecasting purposes.

However, one key aspect that should be addressed is if the reduction in RMSE is statistically significant when comparing models. With this in mind, we have calculated the measure of predictive accuracy proposed by Diebold and Mariano (1995) between the forecast coming from the ingenuous model and those coming from the nine methods used in this article.[7] We have also

Table 11. Ranking of models based on RMSE, TC, *%U1*, *%U2* and *%U3* values

	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
RMSE	0	0	0	0	0	0	0	10	4
TC	0	0	0	0	0	1	0	13	10
% U1	3	0	3	1	1	3	1	2	0
% U2	0	1	0	1	0	1	2	6	3
% U3	0	0	0	0	1	0	0	10	3

Notes: (i) RMSE = root mean square error.
(ii) TC = Theil coefficient.
(iii) *%U1* = percentage of mean error (bias proportion of the MSE).
(iv) *%U2* = percentage of regression error (variance proportion of the MSE).
(v) *%U3* = percentage of disturbance error (covariance proportion of the MSE).

compared the forecasts coming from the two methods proposed in the article to quantify expectations and those coming from the other methods found in the literature. In all cases we have calculated the S(1) measure for the two predictions using a long-run estimate of the variance of the difference series.[8] These results are shown in Tables 12, 13 and 14.

Table 12. Results of the Diebold-Mariano test for the Ingenuous model

Ingenuous model	BAL	AND	PES	CP	ACP	BK	TVP	SS1	SS2
Austria	-4.63	-2.57	-3.65	-3.02	-2.65	-3.02	-2.63	-1.61	-0.01
Belgium	-5.54	-2.91	-3.37	-0.61	-2.94	-0.18	-2.91	1.84	1.05
Finland	-7.05	-3.13	-3.71	-1.48	-3.33	-0.23	-3.16	3.77	-0.40
France	-9.33	-3.50	-4.21	-6.37	-3.38	-6.47	-3.71	0.56	-0.96
Germany	-4.45	-3.69	-4.28	-3.09	-3.99	-1.63	-3.94	0.91	0.00
Great Britain	-10.05	-3.82	-4.67	-8.42	-3.30	-8.28	-3.42	0.72	-1.05
Greece	-4.06	-2.64	-2.54	-2.77	-2.96	-3.34	-3.61	0.69	1.84
Ireland	-7.78	-3.37	-3.63	-6.07	-3.15	-6.19	-3.20	0.65	0.03
Italy	-4.46	-3.24	-4.25	-2.01	-3.36	-1.76	-3.40	0.77	0.90
Luxembourg	-6.00	-3.10	-4.53	-4.18	-3.13	-3.82	-3.08	0.22	-3.48
Netherlands	-3.15	-3.25	-3.66	-3.68	-3.36	-3.97	-3.21	0.99	0.37
Portugal	-4.94	-4.63	-4.56	-4.60	-4.83	-4.89	-4.64	-1.42	-0.37
Spain	-5.19	-5.05	-5.30	-5.27	-5.07	-4.91	-5.14	0.62	0.11
Sweden	-6.71	-0.95	-4.47	-2.83	-1.06	-3.45	-1.10	1.69	-0.99
Euro Area	-5.11	-3.38	-4.45	-2.36	-3.48	-2.00	-3.48	2.01	1.24

Note: *t*-statistic associated to the Diebold-Mariano test for predictive accuracy. The null hypothesis is that the difference between the two competing series is non-significant. A negative sign of the statistic implies that the absolute forecast error associated with the forecast coming from the ingenuous model is lower, while a positive sign implies the opposite.

Table 13. Results of the Diebold-Mariano test for the SS1 model

SS1 model	Naive	BAL	AND	PES	CP	ACP	BK	TVP	SS2
Austria	1.61	-4.62	-2.47	-3.59	-3.01	-2.56	-3.01	-2.52	1.12
Belgium	-1.84	-6.01	-3.03	-3.45	-0.61	-3.07	-0.75	-3.04	-2.08
Finland	-3.77	-7.26	-3.42	-3.98	-1.83	-3.62	-0.84	-3.47	-2.47
France	-0.56	-9.39	-3.45	-4.18	-6.34	-3.33	-6.46	-3.71	-0.78
Germany	-0.91	-4.46	-4.04	-4.57	-3.50	-4.40	-2.02	-4.36	-0.83
Great Britain	-0.72	-10.08	-3.84	-4.74	-8.43	-3.26	-8.29	-3.37	-1.05
Greece	-0.69	-4.10	-2.68	-2.58	-2.81	-3.01	-3.43	-3.54	0.40
Ireland	-0.65	-7.81	-3.39	-3.63	-6.09	-3.17	-6.21	-3.22	-0.73
Italy	-0.77	-4.47	-3.45	-4.39	-2.18	-3.56	-1.87	-3.62	-0.12
Luxembourg	-0.22	-5.93	-3.44	-4.97	-3.86	-3.46	-3.70	-3.47	-1.41
Netherlands	-0.99	-3.30	-3.22	-3.62	-3.62	-3.33	-3.90	-3.16	-1.09
Portugal	1.42	-4.59	-4.46	-4.39	-4.38	-4.67	-4.48	-4.40	1.66
Spain	-0.62	-5.22	-5.21	-5.47	-5.28	-5.29	-4.91	-5.31	-0.96
Sweden	-1.69	-6.80	-1.39	-5.01	-3.17	-1.59	-3.85	-1.57	-1.94
Euro Area	-2.01	-5.16	-3.49	-4.57	-2.49	-3.59	-2.32	-3.58	-2.12

Note: *t*-statistic associated to the Diebold-Mariano test for predictive accuracy. The null hypothesis is that the difference between the two competing series is non-significant. A negative sign of the statistic implies that the absolute forecast error associated with the forecast coming from the SS1 model is lower, while a positive sign implies the opposite.

If we look at the results from Table 12, we can see that the ingenuous model outperforms and is statistically better at the 5% level ($t = 2.028$) than the other procedures, with the exception of the state-space approach (SS1 and SS2). With a few exceptions, forecasts coming from SS1 and SS2 are better statistically than those coming from the rest of the quantification procedures (Tables 13 and 14). The state-space approach also outperforms the ingenuous model in most countries, although there is no significant difference with the exception of Finland.

As we can see from these results, the consideration of asymmetric and stochastic response thresholds in a more general model may improve quantification methods with forecasting purposes. These results are in line with the ones shown in Tables 10 and 11 and in Section three.

Table 14. Results of the Diebold-Mariano test for the SS2 model

SS2 model	Naive	BAL	AND	PES	CP	ACP	BK	TVP	SS1
Austria	0.01	-4.64	-2.53	-3.59	-3.01	-2.61	-3.02	-2.58	-1.12
Belgium	-1.05	-5.73	-2.90	-3.35	-0.16	-2.93	-0.36	-2.89	2.08
Finland	0.40	-7.04	-2.93	-3.51	-1.33	-3.12	-0.13	-2.95	2.47
France	0.96	-9.32	-3.47	-4.17	-6.36	-3.35	-6.46	-3.74	0.78
Germany	0.00	-4.44	-3.57	-4.19	-2.96	-3.86	-1.51	-3.81	0.83
Great Britain	1.05	-10.05	-3.81	-4.56	-8.45	-3.34	-8.31	-3.48	1.05
Greece	-1.84	-4.10	-2.70	-2.61	-2.84	-3.03	-3.42	-3.67	-0.40
Ireland	-0.03	-7.76	-3.36	-3.60	-6.05	-3.13	-6.17	-3.18	0.73
Italy	-0.90	-4.47	-3.37	-4.34	-2.15	-3.50	-1.85	-3.54	0.12
Luxembourg	3.48	-5.99	-2.79	-4.26	-3.71	-2.83	-3.44	-2.76	1.41
Netherlands	-0.37	-3.19	-3.20	-3.61	-3.62	-3.31	-3.90	-3.14	1.09
Portugal	0.37	-4.92	-4.61	-4.54	-4.57	-4.82	-4.83	-4.61	-1.66
Spain	-0.11	-5.19	-5.11	-5.37	-5.26	-5.16	-4.90	-5.19	0.96
Sweden	0.99	-6.71	-0.90	-4.46	-2.82	-1.00	-3.43	-1.05	1.94
Euro Area	-1.24	-5.10	-3.37	-4.44	-2.36	-3.47	-2.03	-3.47	2.12

Note: *t*-statistic associated to the Diebold-Mariano test for predictive accuracy. The null hypothesis is that the difference between the two competing series is non-significant. A negative sign of the statistic implies that the absolute forecast error associated with the forecast coming from the SS2 model is lower, while a positive sign implies the opposite.

5. Concluding comments

Business tendency surveys record the proportion of firms expecting an economic variable to rise, fall or remain unchanged at a given point in time. The qualitative nature of data on the direction of change has often led to quantifying survey results making use of official data, introducing a measurement error due to incorrect assumptions. Monte Carlo simulations allow isolation of the measurement error introduced by different methods used to quantify survey results as respondents diverge from rationality. We used nine different quantification procedures to convert the generated qualitative data on the direction of change into a quantitative measure of agents' expectations. The extensions of the Carlson-Parkin method that allow for asymmetric and dynamic response thresholds obtain lower measurement errors and provide more accurate forecasts of the quantitative variable generated for use as a benchmark.

In this paper we have also used business survey data from the *Joint Harmonised EU Industry Survey* to derive quantitative forecasts of the evolution of industrial prices in fourteen EU countries and the euro area. Though it is impossible to eliminate completely the measurement error introduced when converting qualitative data on the direction of change into quantitative estimations of agents' expectations, state-space models outperform the other methods. In all cases, the state-space models with dynamic threshold parameters outperform the other quantification methods and provide more accurate forecasts of the Producer Price Index. These results suggest that introducing asymmetric and stochastic response thresholds in a more general model may improve quantification methods with forecasting purposes.

Accepted for publication: 3rd March 2006

ENDNOTES

1. Department of Econometrics, University of Barcelona, Diagonal, 690, 08034 Barcelona, Spain. Corresponding author. Tel.: +34-934-021-825; fax.: +34-934-021-821 E-mail address: oclaveria@ub.edu

2. We also estimated initial conditions by OLS regression, obtaining very similar results.

3. To check the robustness of the results, we chose different values for the autoregressive parameter, ranging from 0.1 to 1.0 with an increase of 0.1 each time. As the final results did not vary significantly from one specification to the other we used the model that generates a series as similar as possible to the Producer Price Index in the euro area countries.

4. Muth (1961) assumed that rationality implied that expectations had to be generated by the same stochastic process that generates the variable to be predicted.

5. All simulations are performed with Gauss for Windows NT/95 version 3.2.38.

6. Understood as the difference between the quantitative series generated (actual changes in the variable) and the series of unobservable expectations.

7. We are grateful to an anonymous referee for this suggestion.

8. In order to estimate this long-run variance from its autocovariance function, we have used the Bartlett kernel, as it guarantees that variance estimates are positive definite, while the maximum lag order has been calculated using the Schwert criterion as a function of the sample size.

REFERENCES

Anderson O (1951) 'Konjunkturtest und Statistik', *Allgemeines Statistical Archives*, 35, 209-220.

___ (1952) 'The Business Test of the IFO-Institute for Economic Research, Munich, and its Theoretical Model', *Revue de l'Institut International de Statistique*, 20, 1-17.

Balcombe K (1996. 'The Carlson-Parkin Method Applied to NZ Price Expectations using QSBO Survey Data', *Economics Letters*, 51, 51-57.

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

Carlson J A and Parkin M (1975) 'Inflation Expectations', *Economica*, 42, 123-138.

Claveria O, Pons E and Ramos R (2004) Are Price Expectations Useful for Forecasting Purposes?, Paper presented at the 27th CIRET Conference, Warsaw.

Common M (1985) 'Testing for Rational Expectations with Qualitative Survey Data', *Manchester School of Economic and Social Statistics*, 53, 138-148.

Cooley T F and Prescott E C (1976) 'Estimation in the Presence of Stochastic Parameter Variation', *Econometrica*, 44, 167-184.

Dasgupta S and Lahiri K (1992) 'A Comparative Study of Alternative Methods of Quantifying Qualitative Survey Responses Using NAPM Data', *Journal of Business and Economic Statistics*, 10, 391-400.

Diebold F X and Mariano R (1995) 'Comparing Predictive Accuracy', *Journal of Business and Economic Statistics*, 13, 253-263.

Lee K C (1994) 'Formation of Price and Cost Inflation Expectations in British Manufacturing Industries: a Multi-sectoral Analysis', *Economic Journal*, 104, 372-385.

Löffler G (1999) 'Refining the Carlson-Parkin Method', *Economics Letters*, 64, 167-171.

Mitchell J (2002) 'The Use of Non-normal Distributions in quantifying Qualitative Survey Data on Expectations', *Economics Letters*, 76, 101-107.

Mitchell J, Smith R J and Weale M R (2004) 'Quantification of Qualitative Firm-level Survey Data', *Economic Journal*, 112, 117-135.

Muth J (1961) 'Rational Expectations and the Theory of Price Movements', *Econometrica*, 29, 315-335.

Nardo M and Cabeza-Gutés M (1999) 'The Role of Measurement Error in Rational Expectation Testing', Working Paper 451, Universitat Autònoma de Barcelona.

Pesaran M H (1984) 'Expectation Formation and Macroeconomic Modelling' in Malgrange P and Muet P A (eds) *Contemporary Macroeconomic Modelling*, Oxford: Blackwell.

___ (1987) *The Limits to Rational Expectations*, Oxford: Blackwell.

Seitz H (1988) 'The Estimation of Inflation Forecasts from Business Survey Data', *Applied Economics*, 20, 427-438.

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.

___ (1971) *Applied Economic Forecasting*, Revised Edition, Amsterdam:North-Holland.