Heterogeneity in Consumers' Learning about Inflation

Robert D. J. Anderson [a]

Ralf Becker [b]

Denise R. Osborn [b]

[a] Business School - Economics, Newcastle University

[b] Centre for Growth and Business Cycle Research, Economics, University of Manchester

October 2010

Corresponding author:
Robert Anderson
Business School – Economics
Newcastle University
Newcastle-upon-Tyne
NE1 7RU
UK
Email: robert.anderson@newcastle.ac.uk

Acknowledgements: The authors are grateful for helpful suggestions made by Martyn Andrews, Lynne Evans, Jeremy Smith, members of the Econometrics and Applied Economics research area group at the University of Manchester and also members of the Economics division of the Business School at Newcastle University. We are also obliged to the UK Data Archive for their assistance in obtaining the data used. The usual disclaimer applies that all errors and omissions are entirely the responsibility of the authors. The third author would also like to acknowledge financial assistance from the Economic and Social Research Council (ESRC), under award RES-062-23-1351.

Abstract

This paper provides an empirical analysis of learning by individual consumers in the context of US inflation expectations. By exploiting the short panel dimension of the Michigan survey data, the paper demonstrates that agents overall improve the accuracy of their forecasts at the second interview compared to the first, and hence demonstrate adaptive learning. Further, the extent of this learning, as measured by the reduction in an individual's absolute forecast error for inflation, is associated with their socioeconomic and demographic characteristics. However, heterogeneity in forecast accuracy is less marked at reinterview than at the initial interview, implying that heterogeneity is reduced by learning.

JEL codes: C53, D84, E31, E37

Keywords: Inflation expectations, subjective expectations, adaptive learning, heterogeneity

1. Introduction

As succinctly put by the current governor of the US Federal Reserve: "an essential prerequisite to controlling inflation is controlling inflation expectations" (Bernanke, 2004). The effectiveness of monetary policy therefore requires that the central bank, firstly, understands how consumers form their expectations of future inflation and, secondly, is able to influence these expectations. However, individuals and households have differing experiences and have available different information sets, leading to a growing literature that finds heterogeneity in consumers' inflation expectations formation; see, for example, Branch (2004), Bryan and Venkatu (2001a, 2001b), Lombardelli and Saleheen (2003), Pfajfar and Santoro (2009a, 2009b), and Souleles (2004). In order to understand inflation expectations, therefore, the central bank needs to be aware of the factors driving this heterogeneity. Perhaps surprisingly, socioeconomic and demographic characteristics are important in the formation of individual inflation expectations, with the findings of Bryan and Venkatu (2001a, 2001b), Pfajfar and Santoro (2009b) and Soules (2004) indicating that age, gender and race play roles, together with education and income.

However, consumers not only form expectations, but they also revise these over time. In order to control inflation expectations, the central bank aims to influence this learning process. However, the literature on learning by agents in a monetary policy context (such as Evans and Honkapohja, 2001, 2008) largely ignores heterogeneity in this process. Two exceptions are the theoretical analysis of Berardi (2009) and the empirical study of Pfajfar and Santoro (2009a), both of which are concerned with the possibility that the learning process for groups of consumers can be captured through different statistical updating models. Nevertheless, these studies do not shed light on the whether the nature of this learning depends on the observed characteristics of consumers.

Using individual data from the University of Michigan's Survey of Consumer Attitudes and Behavior and, in particular, the short panel dimension of this data, the current paper investigates the processes generating consumer inflation expectations. More specifically, we focus on learning, with our first finding being that inflation expectations are substantially more accurate at the second interview than the first, which supports the adaptive learning hypothesis (see, for example, Evans and Honkapohja, 2001)

Our more substantive analysis, however, focuses on whether observed characteristics of the individual are associated with learning. Although Bryan and Venkatu (2001a, 2001b), Lombardelli and Saleheen (2003), Pfajfar and Santoro (2009b) and Souleles (2004) consider the role of demographic characteristics for the formation of inflation expectations, our study is the first to examine whether observed heterogeneity, and specifically demographic and/or socioeconomic characteristics, also play a role in learning. Since we find these to be highly statistically significant in explaining learning, our results indicate that groups of the population can be identified where efforts by the monetary authority to influence expectations may be particularly fruitful.

The structure of this paper is as follows. Section 2 reviews the data available from the Michigan survey, while Section 3 then discusses the methodology employed in our empirical analysis. Substantive results are presented and discussed in Section 4, with a final section concluding.

2. Data

This section first provides some background about the University of Michigan survey and its questions relating to inflation expectations, before detailing the other information we employ in our analysis.

2.1 Michigan Survey Inflation Expectations

Since the mid-1940s, and at a monthly frequency since 1978, the Survey Research Centre (SRC) at the University of Michigan has recorded information relating to key economic variables from around 500 adult US consumers, as summarised by Curtin (1982). Along with a changing range of other questions, the telephone interviews[1] record agents' expectations of one year-ahead inflation in the economy, together with a wide range of characteristics about the interviewee, the household in which they live, and the interview itself.

[1] Designed to be representative of the telephone-owning mainland US population.

This SRC data has been widely used in empirical analyses, typically in relation to issues of bias or rationality in inflation expectations, but almost invariably such analyses have employed average (either median or sample mean) forecasts computed over the survey participants in a particular month or quarter. Nevertheless, Branch (2004), Bryan and Venkatu (2001a, 2001b), Pfajfar and Santoro (2009a, 2009b)[2] and Souleles (2004) exploit its rich information content at the individual level to study different aspects of how consumers form inflation expectations. Of most relevance for the present study, the analyses of Bryan and Venkatu (2001a, 2001b), Pfajfar and Santoro (2009b) and Souleles (2004) establish the importance of observed individual characteristics as explanatory variables for the level of inflation expectations. While the analyses of Branch (2004) and Pfajfar and Santoro (2009a) are concerned with learning, they do not consider the role of individual characteristics in this process.

As already noted, the Michigan survey has a distinctive short rotating panel design, with respondents contacted a second time six months after the initial survey. In practice, around 40 percent of a typical monthly sample are re-contacts from six months earlier. Since it controls for individual characteristics, the extent to which inflation expectations change from the first to the second interview provides a potentially powerful source of information about learning by individual consumers, which has not been exploited in any previous study[3].

Year-ahead inflation expectations are captured by two survey questions, the first of which asks a directional question on prices and the second quantifies the (percentage) amount of expected change. However, following the reasoning of Curtin (1996), we censor inflation expectations at +50% and -10%, to counter the possibility that extreme responses could unduly affect our estimated models. This rule censors less than 1% of all responses in each tail. We favour such a broad censoring rule as it allows learning from agents whose responses are initially extreme to be captured in our analysis.

[2] Our analysis has elements in common with Pfajpar and Santoto (2009a, 2009b), whose studies were apparently conducted at the same time as ours. However, these papers do not exploit the short panel nature of the SRC data to examine learning.

[3] Souleles (2004) makes use of this feature of the Michigan survey, but not to examine learning.

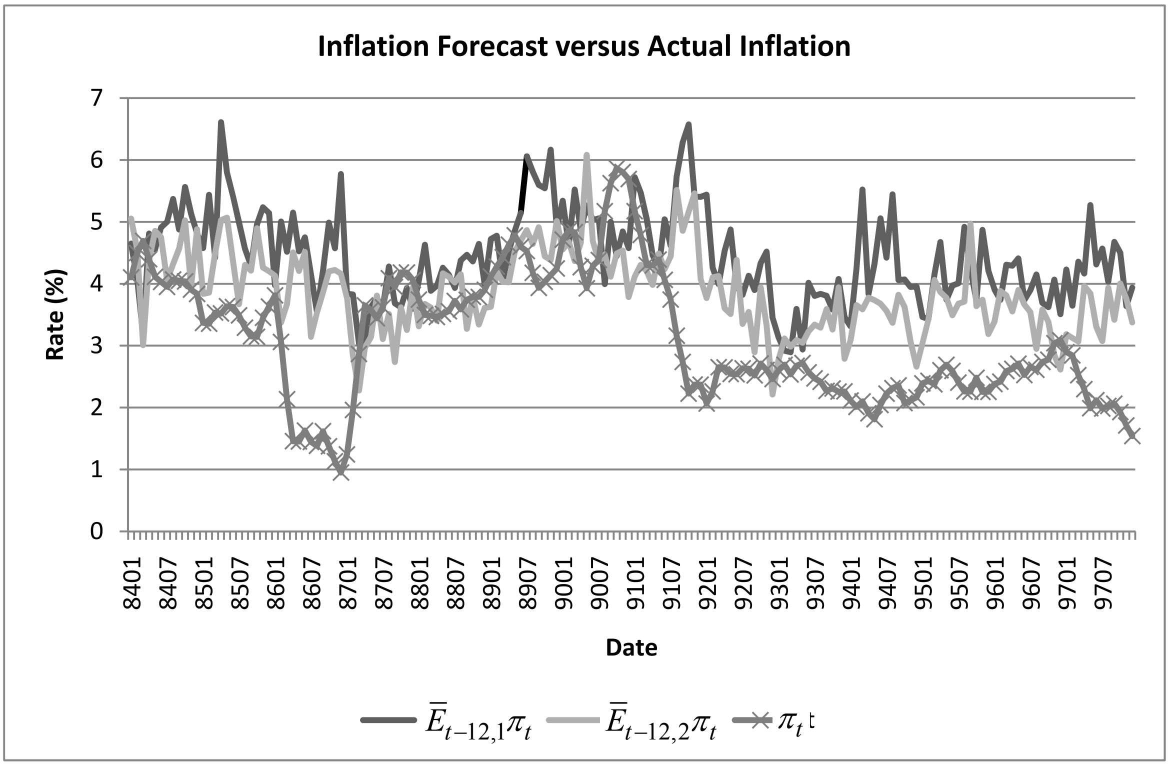
We denote the individual's year-ahead expected inflation, as recorded in the survey for month t as $E_{i,t}\pi_{t+12}$, where i indicates the individual $i = 1,\ldots,N$. As the interview stresses the requirement for a non-personal 'general' expectation response, we follow other studies (Bryan and Venkatu, 2001a, 2001b, Pfajfar and Santoro, 2009b, Souleles, 2004) in assuming that agents are forecasting national, recorded inflation, as measured by the consumer price index (namely, CPI-U).

Due to the availability of consistent individual characteristics information from the SRC data, and also to represent a period of relatively stable inflation, the sample period we analyze extends from January 1983 (8301) until December 1996 (9612)[4]. However, since there are no matching first or second interviews within the overall sample period, second interview responses are discarded for January to June 1983, and similarly first interview responses are discarded for July to December 1996. The total sample contains 168 distinct survey months and observations relate to 46,920 distinct individual respondents. Around 70% of this sample are interviewed twice, yielding a total of 80,159 inflation expectations that are subject to analysis.

Figure 1 compares the prevailing rate of annual inflation, π_t, with the average of first interview expectations from one year earlier, denoted $\overline{E}_{t-12,1}\pi_t$, and the average of second interview expectations, $\overline{E}_{t-12,2}\pi_t$. Although both averages typically exceed the actual value, this comparison indicates that second interview expectations tend to be closer to the actual inflation rate than those from the first interview. Nevertheless, Figure 1 provides only descriptive evidence that Michigan survey respondents learn about inflation between interviews and, in any case, an aggregate analysis cannot shed light on the role of individual heterogeneity.

[4] December 1996 was the most recent period for which the Michigan survey was available at the time this analysis was undertaken. The reference for this data is: "University of Michigan, Survey Research Center, Economic Behavior Program. Survey of Consumer Attitudes and Behavior, ICPSR version. Ann Arbor, MI: University of Michigan, Survey Research Center [producer], 1996. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2000". Available in the UK via reciprocal download rights through the UK Data Archive.

Figure 1: Actual versus Forecast Inflation for First and Second Interviewees



Notes: $\overline{E}_{t-12,1}\pi_t$ and $\overline{E}_{t-12,2}\pi_t$ are average inflation expectations for the year ending in month *t* from those interviewed in the Michigan survey for the first and second times, respectively, while π_t is actual CPI-U inflation over that year.

2.2 Demographic and Interview Characteristics

Table 1 shows the observed characteristics employed in our analysis. The interview and interviewer characteristics included in this table are not considered to be factors that directly influence inflation expectations. However, as discussed in the next section, these may affect the probability that a respondent will agree to a second interview and hence are used to model attrition. Similarly, marital status and household head status are used to explain attrition but not inflation expectations[5].

To facilitate analysis of the role of demographics, some categories on which data are recorded are collapsed in order to ensure adequate observations in all categories, with Table 1 showing those we employ.

[5] Household head status is not included as a potentially relevant demographic variable in previous micro studies of inflation expectations, while Souleles (2004) finds marital status to be insignificant in a model including other demographic variables similar to those of Table 1.

Income data in the Michigan survey is recorded as either current income or in an income band. In either case, we require a real income measure for comparability over time and this is generated using household income distribution information published for each year by the US Census Bureau[6]. In order to retain all income information available in the survey, including that only available in a banded form, while also reflecting both high and low income groups, the individual income response information is converted to a categorical variable, as top 20%, middle 60% and bottom 20% of the household income distribution for the relevant year.

Table 1: Characteristics Employed in the Analysis

Type	Characteristic	Categories
Interviewee Demographics	Age	18-34, 35-54, 55-97
	Income	Bottom 20%, middle 60%, top 20% of US household income distribution
	Race	White, non-white
	Gender	Male, female
	Adults in household	1 (survey respondent), 2+
	Children in household	0, 1+
	Region of residence	North Central (Mid-West), North East, South, West
	Education	No high school diploma, high school diploma, some college, college degree
	Marital status	Married, separated, divorced, widowed, never married
	Household head status	Household head, non-household head
Interview & Interviewer	Interview length	Over 45 minutes, under 45 minutes
	Interview interruption	Interview interrupted and required one or more call-backs, not interrupted
	Interview break-off	Incomplete interview (break-off), complete interview (no break-off)
	Number of calls	1, 2-4, 5-9, 10+
	Initial refusal	No, yes
	Interviewer	Experience 1 (< 0.05% of total interviews), Experience 2 (0.05 to 0.2%), Experience 3 (0.2 to 1%), Interviewer #1 to Interviewer #16 (each > 1% of interviews)

Table notes: The base category used in later analysis is underlined in each case. Interview interruptions have only been recorded since the February 1984 survey. Prior to this date, this variable is set to no-interruption. Interview break-offs were not recorded in the March 1983, June 1983, December 1983, or March 1984 surveys. For these surveys, this variable is set to no-break-off.

[6] Historical information for our sample period is obtained from Table IE-1, Selected Measures of Household Income Dispersion: 1967 to 2001 at www.census.gov/hhes/www/income/histinc/ie1.html.

Demographic characteristics are those provided at the individual's first interview[7]. However, any missing first interview demographic characteristic is substituted by the second interview one. If the latter is also missing, the individual is dropped from the sample.

3. Methodology

This section discusses a number of methodological issues relating to empirically modeling learning about inflation.

3.1 Learning Model Specification

For inflation over the year from the survey in month t, the error in the inflation forecast by individual i is given by $\pi_{t+12} - E_{i,t}\pi_{t+12}$. However, our analysis focuses on the absolute value of the forecast error, that is $\left|\pi_{t+12} - E_{i,t}\pi_{t+12}\right|$, since we wish to capture whether learning improves accuracy, in the sense of the absolute forecast error being closer to zero.

A parsimonious specification that allows us to examine whether learning is present can then be written as

$$\begin{gathered}\left|\pi_{t+12} - E_{i,t}\pi_{t+12}\right| = \beta_0 + \beta_1 \text{survey2}_{i,t} + u_{i,t} \\ \text{for } i = 1,\ldots,N,\ \ t = 1,\ldots,168\ \left(8301,\ldots,9612\right)\end{gathered} \qquad (1)$$

where $\text{survey2}_{i,t}$ is a dummy variable taking the value unity when the interview for individual i in month t is a second interview.

Equation (1) provides an empirical test, at the individual level, for whether consumers exhibit learning, the presence of which is important for the conduct of monetary policy (Gaspar, Smets and Vestin, 2006, Orphanides and Williams, 2005). If the coefficient β_1 is significant and negative, then, on average, consumers have a lower forecast error on reinterview than at the first interview. Viewing the reinterview as a stimulus which serves

[7] As such, people who move, for example, are treated as if resident in the region recorded for their first interview. Therefore, a change in demographic group between interviews is implicitly assumed to have no effect on inflation forecast behaviour.

to heighten awareness of inflation, and the initial prediction as a reference point, such a learning process could be described as adaptive learning[8].

It is worth pointing out that (1) is used in preference to examining directly the change in the forecast error for individual *i* between first and second interviews. This is because our principal focus is modelling this change and, since most available explanatory variables are time invariant, such an approach would difference out the very effects we wish to examine. Further, this would effectively reduce the analysis to available-case analysis and would implicitly involve weighting responses from both interviews. This reduces efficiency relative to our approach that uses all available first interview responses in conjunction with second interview responses weighted to reflect attrition (see the following subsection).

To explore the role of demographic characteristics in both explaining the magnitude of inflation forecast errors and also learning about inflation, the specification used is

$$\begin{aligned}\left|\pi_{t+12} - E_{i,t}\pi_{t+12}\right| &= \gamma_0 + \boldsymbol{\gamma}_1\mathbf{demog}_i + \gamma_2\text{survey2}_{i,t} + \boldsymbol{\gamma}_3\left(\mathbf{demog}_i \times \text{survey2}_{i,t}\right) \\ &\quad + \boldsymbol{\gamma}_4\mathbf{time}_t + u_{i,t} \qquad (2) \\ &\text{for } i = 1,\ldots,N \;\; t = 1,\ldots,168 \;\; (8301,\ldots,9612)\end{aligned}$$

where $\mathbf{demog}_i$ contains demographic and socioeconomic characteristics indicator variables. The individual level variables employed are the interviewee demographics listed in Table 1 (excluding household head and marital status indicators), together with a gender and race interaction variable[9]. **time** is a vector of time dummy variables, one for each survey month. The time dummy variables in (2) take account of changing macroeconomic conditions, as well as the information available to all consumers in the form of policy and other announcements[10]. For reasons of parsimony and ease of interpretation, the effects of these characteristics are assumed not to vary over time[11].

---

[8] Since the second interview is six months after the first, while it is annual inflation which is being forecasted, consumers do not entirely observe their initial forecast error. Nevertheless, the observed course of inflation provides substantial information on the accuracy of this initial forecast.

[9] This interaction variable arises from an initial exploratory analysis that permitted interaction effects in more restricted models.

[10] Hence the time dummy variables also take account of serial correlation that would otherwise arise from the overlapping one-year ahead inflation forecast horizons.

[11] It is not feasible to interact all combinations of demographic characteristics with the survey date dummies, as each such interaction would involve 167 additional regressors. The same consideration applies to the implicit assumption in (1) that learning is time-invariant.

Note that (2) includes a constant, which represents the forecast error for an individual in the base category for all variables: being a white male in mid-age (35-54), in a single adult household with no children and income in the central 60% of the US income distribution, who holds a high school diploma (but no higher educational qualification) and is being interviewed for the first time. October 1995 is treated as the time base period.

The effects of observed heterogeneity on forecast accuracy for first-time interviewees are captured by the coefficients in $\boldsymbol{\gamma}_1$, with these interpreted relative to the base category[12]. In the context of (2), a coefficient in $\boldsymbol{\gamma}_1$ quantifies the improvement (or deterioration) in forecast accuracy relative to the base category for an individual possessing the indicated characteristic, with a significant negative coefficient indicating greater accuracy. Therefore, this specification allows demographic and socioeconomic characteristics to influence inflation expectations, in line with the findings of Bryan and Venkatu (2001a, 2001b), Pfajfar and Santoro (2009b) and Souleles (2004) for the SRC inflation expectations data.

The coefficient γ_2 for the second interview indicator variable in (2) allows a test for adaptive learning by the base demographic group. Heterogeneity in such learning is then represented by the vector $\boldsymbol{\gamma}_3$, where each coefficient is again interpreted in relation to the base group. Joint significance of $\boldsymbol{\gamma}_3$ is a test of the null hypothesis that the extent of learning, as measured by the change in forecast accuracy between interviews, is constant across groups.

There is, however, a serious issue with estimation of the coefficients of (2), which arises because not all respondents to the initial survey agree to a second interview. Indeed, such second interview (self-selecting) attrition is generally around 30% for most of the sample period. Clearly, this may lead to the sub-sample on whom second interview data are available, to be a non-random sample from the initial respondents, resulting in potentially

[12] Considerations of absolute forecast accuracy would also require consideration to be taken of the survey date dummy variable coefficients in $\boldsymbol{\gamma}_4$.

biased parameter estimates in (2)[13]. The next subsection discusses these issues, before considering estimation and inference issues for (1) and (2).

3.2 Modelling Attrition

Methods to deal with attrition, or what Wooldridge (2002) describes as incidental truncation, primarily depend on what is assumed (or known) about the drop-out mechanism that causes respondents to fail to continue. Our assumption is that second interview data are *missing at random*, in the sense that known characteristics determine the drop-out propensity.

Vandecasteele and Debels (2007) identify age, gender, region, race, education, marital status, household size, income, interviewer, length of interview, labour force status and home ownership status as factors commonly related to attrition. Table 1 demonstrates that all except the final two of these indicators are available in the SRC dataset. Therefore, we model the reinterview probability using these available factors, together with survey date indicator variables. More specifically, a logit regression model is estimated using maximum-likelihood, where the model has the form:

$$\begin{gathered}\Pr\left(\text{not-reinterviewed}_{i,t}\right) = \Phi\left(\alpha_0 + \alpha_1 \mathbf{characteristics}_i + \alpha_2 \mathbf{time}_t + v_{i,t}\right) \\ \text{for } i = 1,\ldots,n,\ t = 1,\ldots,162\ \left(8301,\ldots,9606\right)\end{gathered} \tag{3}$$

where $\text{not-reinterviewed}_{i,t}$ is an indicator variable for whether the individual is not-reinterviewed six-months hence, $\mathbf{characteristics}_i$ is a matrix of demographic, interview and interviewer characteristics indicator variables, as detailed in Table 1.

3.3 Estimation and Inference

Estimation and inference for the learning models of (1) or (2) needs to confront three distinct issues, namely the possibility of attrition between the first and second interviews, the correlation of the inflation forecast errors for a specific individual over surveys and the nature of the dependent variable which, by construction, has its values truncated at zero.

[13] The small proportion of individuals who respond to both interviews and provide a year-ahead inflation expectation in the first interview, but not the second, are dropped from the sample. As the factors which determine complete attrition from the survey as opposed to question attrition may be different, these cases are not used for the estimation of attrition probability.

Following the proposal of Robins, Rotnitzky and Zhao (1994), the potential bias associated with attrition is corrected by using inverse-probability weighting, whereby the available second interview cases for the inflation forecast error models of (1) and (2) are weighted by the inverse of the (estimated) probability of that case responding to a second interview, computed from the logit regression of (3). These inverse probability weights are then used in conjunction with weighted least squares (WLS) estimation to compensate for attrition.

Individual-level correlation arises in our case because the forecast errors made by an individual at first and second interview are correlated, due to the presence of the unobserved individual level traits and misconceptions. In effect, the error term in (1) and (2) can be written as

$$u_{i,t} = a_i + \varepsilon_{i,t} \tag{4}$$

where a_i is the unobserved individual effect and $\varepsilon_{i,t}$ is an error term that is assumed to be white noise over both time and individuals.

The cluster-robust variance-covariance matrix, proposed by Rogers (1991) as an extension to the heteroskedasticity robust variance-covariance matrix of White (1980), can accommodate both an individual-level effect (resulting in a "cluster") and heteroskedasticity of unspecified form, on the assumption that the unobservable individual effect, a_i in (4), is uncorrelated with the explanatory variables in (1) or (2)[14]. The method also ensures that individual and joint tests of significance are (asymptotically) robust to possible heteroskedasticity.

For the $i = 1,\ldots,N$ clusters (individuals) and $s = 1,2$ observations (interviews) on each individual, the Rogers (1991) method employs the $(s \times 1)$ vector of (attrition corrected) WLS residuals applicable to each cluster, $\hat{u}_i$, and the $(s \times K)$ matrix X_i of the K explanatory variables relevant to this cluster. Employing the complete $(Ns \times K)$ matrix of

[14] This assumption would usually be tested by comparing a random-effects model with a fixed-effects model, using a Hausman test. However, since an individual is observed at only two periods, the fixed-effects model is identical to OLS on a first-differenced equation, which is unattractive in our context for the reasons discussed in subsection 3.1.

explanatory variables, X, the full-sample heteroskedasticity and cluster robust variance-covariance matrix for the vector of coefficients $\hat{\gamma}$ is formed as[15]

$$\hat{V}_{\hat{\gamma}}^{CR} = h\left(X'X\right)^{-1}\left(\sum_{i=1}^{N}\left(X_i'\hat{u}_i\right)\left(X_i'\hat{u}_i\right)'\right)\left(X'X\right)^{-1} \tag{5}$$

where h is a residual weighting scheme, distinct from the attrition weights, used to reduce bias caused by using residuals in the computation of this robust variance-covariance matrix. In practice we employ

$$h = \frac{N}{N-1}.\frac{Ns}{Ns-K}. \tag{6}$$

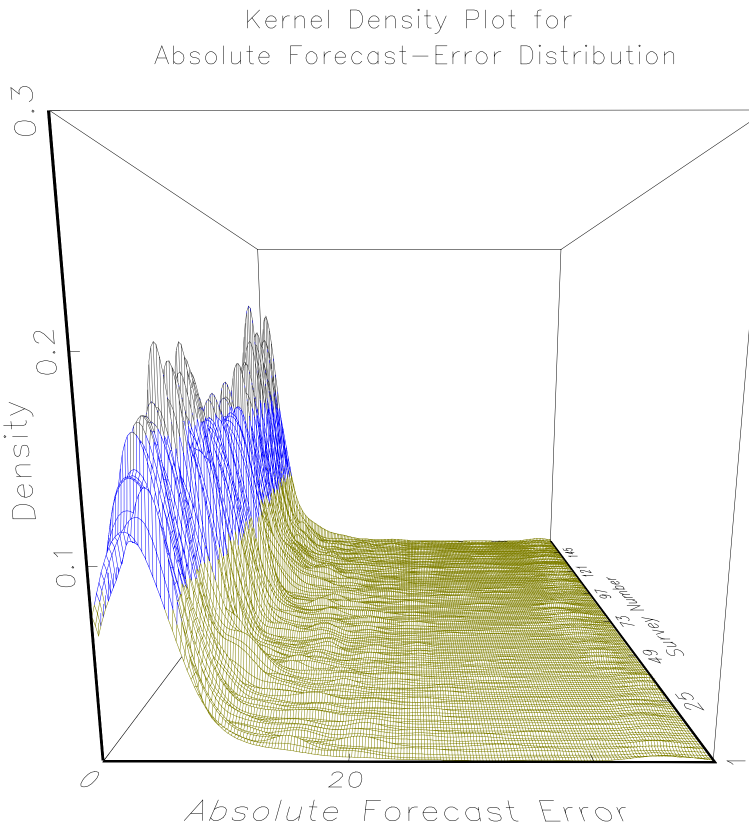
Rogers (1991) provides simulation evidence that the robust variance-covariance matrix is generally effective, except in situations where the size of any cluster is more than 5% of the total observations. This condition is not violated by the SRC dataset, since all clusters contain a maximum of two observations (an initial and possible re-interview).

Estimation and inference techniques for truncated variables often rely on the underlying (non-truncated, in this case the forecast error) distribution being normal. However, this does not apply in our case and consequently, standard truncated-normal regression techniques are inadequate.

The distribution of the absolute forecast errors, shown in Figure 2 calculated over all interviewees for each survey in our sample, is a zero left-limit truncated distribution, which suffers from excess kurtosis. Essentially, this non-normality arises from two sources. The first is the non-normality of the underlying (signed) forecast error distribution, while the second is the use of the absolute value. Positive forecast skew further compounds these issues.

[15] This is a modified version of the Rogers (1991) formula. When s=1, $\hat{u}_i$ is a scalar, and (5) collapses to the usual White heteroskedasticity robust covariance matrix, equivalent to $X'\hat{\Omega}X$ where $\hat{\Omega}$ is a diagonal matrix containing $\hat{u}_i^2$ on the main diagonal.

Figure 2: Absolute Forecast Error Distribution



Note: Survey 1 corresponds to January 1993, survey 2 corresponds to February 1993 etc.

Due to this highly non-normal distribution, asymptotic standard errors may not be reliable for inference. To account for this, we employ the cluster-robust procedure of Cameron, Gelbach and Miller (2008).

This procedure begins by drawing (with replacement) a new sample of s = 2 (first and, if relevant, second survey) observations on N individuals from the original sample. Using these bootstrap observations, and their corresponding weights that account for attrition, bootstrap coefficient estimates, denoted $\hat{\gamma}^*_{k,b}$ for coefficients k = 1, …, K are obtained by WLS with associated cluster-robust standard errors obtained as the square roots of the diagonal elements of (5), with these denoted $s^{*CR}_{\hat{\gamma}_k}$. This is repeated for 200 bootstrap samples, $b = 1,...,200$. Using these bootstrap sample statistics, a vector of 200 bootstrap t-statistics is computed for the kth regressor, where the bth element of that vector is given by

$$\hat{w}_{k,b} = \frac{\hat{\gamma}^*_{k,b} - \hat{\gamma}_k}{s^{CR}_{\hat{\gamma}^*_{k,b}}}, \tag{7}$$

which is centred on the WLS coefficient estimate $\hat{\gamma}_k$ obtained from the observed data. A bootstrap (100 – α)% t-statistic confidence interval around zero is then obtained using the $\alpha/2$ and the 1-$\alpha/2$ percentiles of the (200×1) vector of bootstrap sample t-statistics. We report this interval for α = 5%[16]. By comparing the sample cluster robust t-statistic to this interval, bootstrap inference can be conducted at the 5% level of significance. Further, comparison of the bootstrap interval to the conventional 95% t-statistic confidence interval [-1.96, 1.96] that would apply for normal variables provides an indication of the validity of conventional asymptotic inference.

4. Results

Before turning to our principal focus, namely testing for adaptive learning about inflation and the role of observed individual characteristics in such learning (subsections 4.2 and 4.3, respectively), subsection 4.1 considers our results on modelling attrition.

4.1 Attrition

The results in Table 2 establish that attrition is related to the respondent's characteristics. More highly educated individuals, together with those who are married or widowed, have a higher propensity for reinterview than the base group. Those who live in regions other than the MidWest, have below average income, are aged 18-34, have low education achievements (no high school diploma), are non-white or separated from their partners, all have a reduced propensity for reinterview. However, gender, the presence of children or other adults within the household, and status of the interviewee within the household, apparently play no role. Jointly, demographic characteristics are highly significant for attrition.

[16] The heavy computation cost of the bootstrap procedure with our sample of N = 46,920 limits the feasibility of performing a larger number of bootstrap replications and hence the feasibility of conducting bootstrap inference at tighter levels of significance.

Table 2: Logit Model for No Reinterview

	Coefficient	Std. Err.	*p*-Value	Significance
Respondent characteristics				
Low income	0.203	0.0350	0.0000	***
High income	-0.036	0.0327	0.2709	
No HS Diploma	0.173	0.0386	0.0000	***
Some College	-0.171	0.0311	0.0000	***
College Degree	-0.325	0.0305	0.0000	***
Age 34-	0.179	0.0288	0.0000	***
Age 55+	0.036	0.0350	0.3037	
Female	0.018	0.0346	0.6029	
Non-head of household	-0.006	0.0374	0.8725	
Non-white	0.327	0.0456	0.0000	***
Non-white x Female	0.032	0.0599	0.5932	
Separated	0.239	0.0670	0.0004	***
Married	-0.124	0.0374	0.0009	***
Widowed	-0.183	0.0590	0.0019	**
Divorced	-0.003	0.0464	0.9484	
North East	0.122	0.0343	0.0004	***
South	0.082	0.0302	0.0066	**
West	0.169	0.0340	0.0000	***
Interview characteristics				
Interview Breakoff	0.618	0.0933	0.0000	***
Interview Interrupt	0.261	0.0421	0.0000	***
Interview Length > 45min.	-0.165	0.0496	0.0009	***
Calls 2+	0.072	0.0349	0.0391	*
Calls 5+	0.326	0.0373	0.0000	***
Calls 10+	0.495	0.0333	0.0000	***
Initial Coversheet Refusal	0.387	0.0358	0.0000	***
Constant	-3.389	0.1379	0.0000	***
Joint Hypothesis Tests	**Statistic**	**DoF**	***p*-Value**	**Significance**
Interviewer	35.76	18	0.0076	**
Survey Month	938.31	161	0.0000	***
Demographics	700.45	20	0.0000	***
Interview characteristics (exc. interviewer)	677.78	7	0.0000	***
All coefficients (exc. constant)	2796.02	206	0.0000	***

Notes: The equation estimated is given by (3) where **characteristics** is the matrix of demographic and survey indicator variables in Table 1. Interviewer indicator variables, as defined in Table 1, are included in the regression, as are time (survey month) indicator variables, although detailed results for these are not shown. The joint tests are asymptotically χ^2 under the null hypothesis that the relevant coefficients are all zero, with indicated degrees of freedom (DoF). * denotes significance at the 5% level, ** at the 1% level, *** at the 0.1% level. A blank in this column indicates significance only at levels above 5%.

Interview characteristics are also important. Initial interviews containing interruptions or breakoffs are less likely to be re-observed. Similarly, initial (coversheet) refusal and requiring multiple calls to obtain the first interview reduce the propensity for reinterview. Interestingly, when an initial interview lasts more than 45 minutes, the respondent is more likely to agree to be reinterviewed! Again, these variables are jointly significant. The impact of the particular interviewer on the reinterview probability is much less strong, but the interviewer indicator dummy variables are nevertheless jointly significant at the 5% level.

Although detailed results are not shown, attrition rates are also affected by the survey date. Jointly, these time-effects are highly significant, as is the overall regression.

4.2 Testing Adaptive Learning

The first substantive question we confront is whether consumers learn about inflation between the first and second interviews of the Michigan survey. Adaptive learning would imply that forecast errors are lower at reinterview than at the first interview. Table 3 reports the results through estimation of (1). These results are obtained weighting observations using the attrition probabilities implied by the estimates of Table 2, and with bootstrapped cluster (and heteroscedasticity) robust inference applied, as outlined in subsection 3.3.

Table 3: Overall Adaptive Learning Model Estimates

	Coefficient	Std. Error	*t*-statistic	Asymptotic *p*-value	Significance	Bootstrap 95% *t*-statistic Interval
Constant	3.766	0.026	146.253	0.000	***	[-2.153,1.788]
Survey2 indicator	-0.531	0.037	-14.360	0.000	***	[-1.963,2.042]

Notes: The equation estimated is given by (1). All *p*-values and significance tests refer to cluster adjusted asymptotic Wald tests of significance (null is equality with zero, against a two-sided alternative). The methodology for construction of the bootstrap confidence intervals for *t*-statistics is discussed in subsection 3.3. *** indicates significance at the 0.1% level

The highly significant negative *t*-statistic on the survey indicator coefficient supports the presence of adaptive learning. Overall, responses are closer to the actual year-ahead inflation rate by approximately half of a percentage point when an individual is reinterviewed, compared to the initial interview, and hence respondents (on average) substantially improve their forecast accuracy from the first to second interview.

This result is important for a number of reasons. Firstly, it supports the hypothesis that inflation expectations are adaptive, with individuals learning from their past inflation forecast errors. However, this learning applies at the level of the individual consumer, and not necessarily at the macroeconomic level. Nevertheless, it indicates that raising the awareness of inflation (and, presumably, other macroeconomic phenomena) leads to consumers having better information and, consequently, to better decision-making by both individual economic agents and policy-makers (Orphanides and Williams, 2004).

A further consequence of learning is that analysis of the "average" inflation expectation from the SRC survey is not representative of the general population, because it reflects the impact of around 40% of the total sample having more information than a typical consumer due to these people being reinterviewed. Therefore, macro-level analyses of inflation expectations that are based on SRC data could draw inappropriate conclusions in relation to expectations held by US consumers in aggregate.

4.3 Heterogeneity in Learning

The results of Table 4 allow for heterogeneity, both in the first interview response and also in the extent of adaptive learning. The base group for this analysis, represented by the constant term, is middle income households, containing no adults other than the respondent and no children, where the respondent is a white male in the 35-54 age group. Inference is performed using conventional (asymptotic) Wald tests and also through the bootstrap *t*-ratio test.

Results in the first sub-panel confirm those of Souleles (2004) and Pfajfar and Santoro (2009b) in finding that initial accuracy of inflation expectations is strongly influenced by income, age, race and gender. However, unlike Souleles (2004), we find no role for the presence of children in the household. The only demographic group for which initial interview forecasts are statistically significantly more accurate than the base group are those from high-income households, who achieve an accuracy gain of 0.7 percentage points. On the other hand, respondents from low income households are less accurate by around 1.2 percentage points, confirming that accuracy is inversely related to income.

Table 4: Heterogeneity and Learning Estimation

	Coefficient	Std. Error	*t*-statistic	Asymptotic *p*-value	Significance	Bootstrap 95% *t*-statistic Interval
Constant	3.783	0.134	28.201	0.0000	***	[-2.109, 1.872]
Low income	1.152	0.102	11.266	0.0000	***	[-1.735, 2.087]
High income	-0.683	0.050	-13.763	0.0000	***	[-2.048, 1.951]
Age 34-	0.270	0.060	4.506	0.0000	***	[-2.028, 1.955]
Age 55+	-0.014	0.067	-0.203	0.8391		[-2.228, 1.950]
Non-white	1.264	0.118	10.711	0.0000	***	[-2.132, 1.928]
Female	0.960	0.047	20.458	0.0000	***	[-2.293, 1.801]
Children in HH	-0.065	0.075	-0.865	0.3870		[-1.762, 1.666]
Multiple adults in HH	0.179	0.062	2.866	0.0042	**	[-1.976, 1.767]
Non-white x Female	1.109	0.192	5.765	0.0000	***	[-2.335, 1.777]
Survey2 indicator	-0.210	0.108	-1.945	0.0518		[-2.080, 1.775]
Low income x survey2	-0.292	0.141	-2.061	0.0393	*	[-1.953, 1.860]
High income x survey2	0.134	0.073	1.822	0.0685		[-1.931, 1.966]
Age 34- x survey2	-0.214	0.086	-2.477	0.0132	*	[-1.762, 2.113]
Age 55+ x survey2	0.044	0.099	0.444	0.6570		[-2.139, 1.892]
Non-white x survey2	-0.033	0.189	-0.177	0.8595		[-2.667, 1.721]
Female x survey2	-0.137	0.072	-1.904	0.0569		[-2.110, 1.891]
Children in HH x survey2	-0.206	0.082	-2.519	0.0118	*	[-1.846, 1.991]
Multiple adults in HH x survey2	-0.047	0.090	-0.517	0.6052		[-1.751, 1.770]
Non-white x Female x survey2	-0.682	0.288	-2.371	0.0177	*	[-1.769, 2.129]
Joint Hypothesis Tests	**Statistic**	**DoF**				
Demographics (first interview)	156.33	(9, 46911)		0.0000	***	
Demographics/survey2 interactions	5.34	(8, 46911)		0.0000	***	
Survey month indicator dummies	4.37	(167, 46753)		0.0000	***	

Notes: The equation estimated is given by (2). All *p*-values and significance tests refer to cluster adjusted asymptotic Wald tests of significance (null is equality or joint equality with zero, against a two-sided alternative). The methodology for construction of the bootstrap confidence intervals for *t*-statistics is discussed in subsection 3.3. * denotes significance at the 5% level, ** at the 1% level and *** at the 0.1% level significance. A blank in this column indicates significance only at levels above 5%. The constant term includes the effect of both the base demographic group and the base month.

Also, non-white respondents and females are less accurate than the base category by 1.3 and 1.0 percentage points, respectively, with being both female and non-white compounding these individual characteristic effects by a further 1.1 percentage points. Compared to the (white male) base group, non-white females are less accurate by 3.3 percentage points. Other significant (at 1%) characteristics associated with less accurate initial forecasts, though of a smaller magnitude, are respondents in the 18-34 age group (0.3), and those in households containing further adults (0.2). While the young have less accurate inflation expectations, being aged 55 or above is not significant. The demographic characteristics are jointly highly significant for initial forecast accuracy.

Compared with the survey indicator coefficient of Table 3, the effects of learning are more than halved in Table 4. However, in the latter case, this refers only to the base group, who improve forecast accuracy by 0.2 percentage points, which is marginally significant as the computed *t*-statistic is just inside the bootstrap 95% *t*-statistic interval. Overall, however, (measured by the joint significance of the coefficients for the interactions of the demographic variables with the survey2 indicator) learning exhibits highly significant heterogeneity over demographic groups, and hence the base group effect is not indicative of consumers overall.

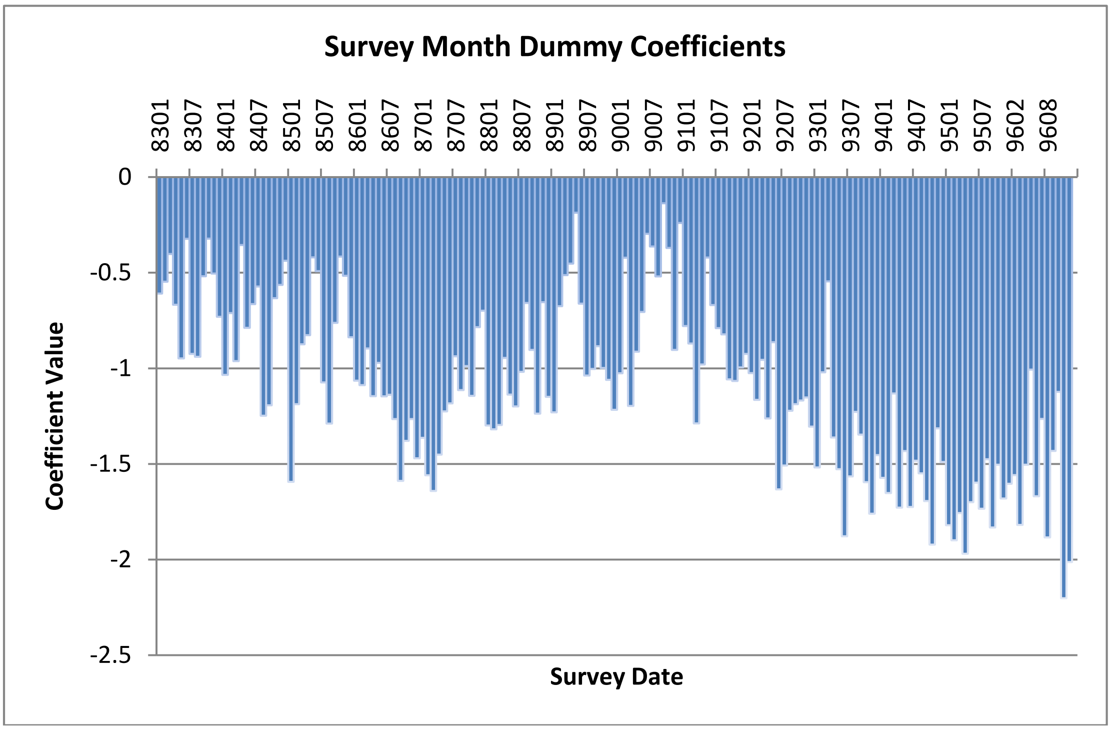
Indeed, the reinterview interaction coefficients that are significant (at 5%) are all negative, indicating that these groups improve their forecasts compared to the base group. More specifically, low-income individuals, those in the 18-34 age group and those in households with children are each found to improve their forecasts in the second interview by 0.2-0.3 percentage points more than the base group, while non-white females further increase accuracy by nearly 0.7 percentage points. Apart from the case of non-whites and females, who individually do not have significantly improved accuracy, individuals who are the least accurate initial forecasters typically improve their forecast accuracy by the time of the second interview. However, although individuals in households with two or more adults are significantly less accurate than the base group at first interview, they do not improve their relative accuracy at the second interview. Nevertheless, overall, heterogeneity in the accuracy of inflation expectations is less marked at the second interview than at the first.

For the high-income group, which was the most accurate group of initial forecasters, results suggest that this group may not learn more (or less) than the base (mid-income) group over the six-month reinterview horizon. Two further effects are worthy of comment. Firstly, households with children improve their relative forecast accuracy in the second survey wave, as do younger respondents (aged less than 35). The additional learning for the former takes place despite the group not differing from the base at first interview, while the latter effectively "catch-up" at reinterview, so that age then apparently becomes irrelevant for the accuracy of inflation expectations.

From an econometric perspective, the bootstrap cluster robust 95% *t*-statistic confidence intervals reported in Table 4, although generally shifted to the left compared with usual [-1.96,1.96] 95% confidence interval, provide evidence that the use of conventional asymptotic inference is reasonably reliable, despite the highly non-normal data distribution for the dependent variable (as discussed in Section 3).

Finally, Figure 3 plots the estimated coefficient values for the survey month dummy variables. Since these coefficients are predominantly negative, inflation forecasts are typically more accurate that the value implied by the constant in Table 4 (relating to the base survey month of October 1995). Visually, it appears that consumers' inflation forecasts have tended to become more accurate over the sample period examined here, with accuracy increasing from around 1992, which may be associated with a decline in volatility of US CPI inflation around that date (see Bataa, Osborn, Sensier and van Dijk, 2008).

Figure 3: Survey Month Dummy Coefficients



5. Conclusions and Summary

The previously unexploited short panel aspect of the Michigan SRC dataset offers the opportunity to examine whether individual consumers learn about inflation and whether any such learning is heterogeneous across different demographic groups. While recent theoretical models studying optimal monetary policy assume adaptive learning by consumers (for example, Gaspar, Smets and Vestin, 2006, Orphanides and Williams, 2005), the present paper is (to our knowledge) the first to explicitly test this hypothesis using survey inflation data at the individual level. Reassuringly for these models, the SRC data provides strong evidence of adaptive learning, such that year-ahead inflation forecasts are substantially more accurate when agents are reinterviewed compared with an initial interview six months earlier.

Nevertheless, the vast majority of theoretical models also implicitly assume homogeneity in both expectations and learning, whereas this paper focuses on the possibility that inflation forecast accuracy is associated with the demographic characteristics of the respondent. The SRC data is particularly valuable for this analysis, because the survey collects a rich store of individual level information, while the panel dimension enables us to control for additional (unobserved) heterogeneity when examining learning.

Our results suggest that initial forecast accuracy, compared to a mid-aged male with mid-range household income, is substantially reduced if the respondent is non-white, female, or in a low income household. Other, less strong characteristics which reduce initial forecast accuracy are the respondent being in the 18-34 age group, or if there are further adults in the household. Conversely, the highest forecast accuracy is achieved by individuals in high income households. These results support previous findings that consumers' inflation expectations are linked to their observed demographic characteristics (Bryan and Venkatu, 2001a, 2001b, Lombardelli and Saleheen, 2003, Pfajfar and Santoro, 2009, and Souleles, 2004).

However, we also show that learning is heterogeneous over demographic groups. In particular, non-white females, low-income individuals, those aged 18-34 or in households with children improve their second accuracy by the largest amounts at the second interview. Since, in general, these are also the groups which have least accurate initial forecasts, learning acts to reduce heterogeneity. This suggests that a known reinterview acts as an incentive for individuals with poor initial forecasts to notice inflation, an incentive which is not present at an initial 'cold-call' interview. Although our analysis does not include the role of macroeconomic announcements, this finding is compatible with the emphasis placed by Orphanides and Williams (2005) on the importance of effective communication of the central bank's inflation objective in order to anchor inflation expectations.

The paper also discusses estimation issues, including dealing with unobserved individual effects and allowing for possible heteroskedasticity of unknown form. Attrition is modelled and then (for the purposes of modelling learning) is corrected using inverse probability weighting. Forecast accuracy and learning are quantified through the use of the absolute value of a non-normal dependent variable, and inference is verified through application of a bootstrap procedure to the estimated

Wald *t*-test statistics. However, the usual asymptotic inference provides a satisfactory approximation to this procedure.

Our overall conclusion is that central banks wishing to anchor inflation expectation to actual inflation should consider initiatives which stimulate agents to learn about (or simply notice) inflation. Since our results provide evidence that the magnitude of learning and the level of initial forecast accuracy depend on observed demographic characteristics, information can be targeted to specific groups of the population in order to stimulate their learning and hence improve the effectiveness of monetary policy actions.

References

Bataa, E.; Osborn, D.R.; Sensier, M.; van Dijk, D. (2008) "Identifying Changes in Mean, Seasonality, Persistence and Volatility for G7 and Euro Area Inflation", Centre for Growth and Business Cycle Research Discussion Paper Series No 109, Economics, The University of Manchester.

Berardi, M. (2009) "Monetary Policy with Heterogeneous and Misspecified Expectations", Journal of Money, Credit and Banking, 41, 79-100.

Bernanke, B.S. (2004) Remarks by Governor Ben S. Bernanke at the Bond Market Association Annual Meeting, New York, New York: The economic outlook and monetary policy, April 23, 2004, available from: www.federalreserve.gov/boarddocs/speeches/2004/200404222/default.htm

Branch, W.A. (2004) "The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations", Economic Journal, 114, 592-621.

Bryan, M.F., and Venkatu, G. (2001a) "The Demographics of Inflation Opinion Surveys", Federal Bank of Cleveland Economic Commentary, November.

Bryan, M.F., and Venkatu, G. (2001b) "The Curiously Different Inflation Perspectives of Men and Women", Federal Bank of Cleveland Economic Commentary, February.

Cameron, A.C., Gelbach, J.B. and Miller, D.L. (2008) "Bootstrap-Based Improvements for Inference with Clustered Errors", The Review of Economics and Statistics, 90, 414-427.

Curtin, R.T. (1982) "Indicators of Consumer Behavior: The University of Michigan Surveys of Consumers", The Public Opinion Quarterly, 46, 340-352.

Curtin, R.T. (1996) "Proceedure to Estimate Price Expectation", University of Michigan, available from www.sca.isr.unich.edu/main.php

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

Evans, G.W., and Honkapohja, S. (2008) "Expectations, Learning and Monetary Policy: An Overview of Recent Research", University of St. Andrews Centre for Dynamic Macroeconomic Analysis Working Paper Series, CDMA08/02.

Gaspar, V., Smets, F., and Vestin, D. (2006) "Adaptive Learning, Persistence and Monetary Policy", Journal of the European Economic Association, 4, 376-385.

Lombardelli, C., and Saleheen, J. (2003) "Public Expectations of UK Inflation", Bank of England Quarterly Bulletin, Autumn, 281-286.

Orphanides, A., and Williams, J.C. (2005) "Imperfect Knowledge, Inflation Expectations and Monetary Policy". In Inflation Targeting, edited by B. S. Bernanke and M. Woodford (eds.), University of Chicago Press, Chicago.

Pfajfar, D., and Santoro, E. (2009a) "Heterogeneity, Learning and Information Stickiness in Inflation Expectations", Mimeo, University of Copenhagen.

Pfajfar, D., and Santoro, E. (2009b) "Asymmetries in Inflation Expectations across Socioeconomic Groups", Mimeo, University of Copenhagen.

Robins, J.M., Rotnitzky, A. and Zhao, L.P. (1994) "Estimation of Regression Coefficients when Some Regressors are Not Always Observed", Journal of the American Statistical Association, 89 (427-September), 846-866.

Rogers, W. (1991) "Regression Standard Errors in Clustered Samples", Stata Technical Bulletin 13, available in *Stata Technical Bulletin Reprints*, 3, 89-94.

Souleles, N.S. (2004) "Expectations, Heterogeneous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys", Journal of Money, Credit and Banking, 34(1), February, 39-72.

Vandecasteele, L., and Debels, A. (2007) "Attrition in Panel Data: The Effectiveness of Weighting", European Sociological Review, 23, 81-97.

White, H.L. (1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and Direct Test for Heteroskedasticity", Econometrica, 48(1-May), 817-838.

Wooldridge, J.M. (2002) "Inverse Probability Weighted M-Estimators for Sample Selection, Attrition and Stratification", Portuguese Economic Journal, 1, 117-139.