Heterogeneity in Consumers' Learning about Inflation

Dr Robert Anderson
Newcastle University Business School – Economics

Prof Denise Osborn and Dr Ralf Becker
Centre for Growth and Business Cycle Research, Economics,
University of Manchester

Motivation and Objectives

- US monetary policy responsive to changes in (consumer) expected inflation
- Related inflation expectations literature generally ignores the influence of demographic characteristics on forecasting processes.
- SRC dataset on US year-ahead consumer expectations and demographics has received very little attention by academic community
- SRC dataset contains 6-month short-rotating panel dimension– unique aspect to investigate learning

Newcastle University

Related Literature

Heterogeneity in expectations (data driven):

Branch (2004), Bryan and Venkatu (2001a, 2001b), Lombardelli and Saleheen (2003), Pfajfar and Santoro (2009a, 2009b), Souleles (2004)

Economic learning:

Evans and Honkapohja, (2001, 2008)

Berardi (2009), Pfajfar and Santoro (2009a) – different models for forecast update.

Newcastle University

Data

Survey Research Centre (SRC) survey:

- Representative sample of approx. 500 US consumers per-month
- Telephone interviews (typically 30mins duration)
- Inflation expectations elicited by:

  "During the next 12 months, do you think that prices in general will go up, or do down, or stay where they are now?"

  "By about what percent do you expect prices to go (up/down) on the average, during the next 12 months?"
- Monthly data available January 1978 to December 1996
- Short rotating panel design – 40% reinterviewed from 6-months previously (in both cases for a 12-month forward expectation)
- Demographic and interview characteristics recorded along with other sentiment and financial situation related questions

Newcastle University

Data issues:

- Inconsistent reinterview horizons – truncate sample to post 1980
- Individual identifier miscodes and data integrity – truncate sample to post 1983
- Question evolution and non-response (particularly income) – transform responses in income bands
- Extreme inflation expectation responses – censor at +50% and -10%
- Non-availability of post-1996 individual response data

Newcastle University

Characteristics employed (number of sub-groups):

- interviewee
  - Age (3), income (3), race (2), gender (2), number of adults (2) and children in household (2), region of residence (4), education (4), marital status (5), household head status (2)
- interview and interviewer
  - Length (2), whether interrupted or contains break-off (3), number of calls made (4), initial refusal (2), interviewer ex-post experience (15)

Note not all characteristics employed in all models estimated.

Newcastle University

Final dataset:

- 168 distinct survey months (first interview Jan 1983 to Jun 1996)
- 46,920 individual respondents
- 80,159 inflation expectations

Modelling Learning and Forecast Accuracy

Relating learning and forecast accuracy to demographic characteristics:

$$\left|\pi_{t+12} - E_{i,t}\pi_{t+12}\right| = \gamma_0 + \gamma_1 \mathbf{demog}_i + \gamma_2 \text{survey2}_{i,t} + \gamma_3\left(\mathbf{demog}_i \times \text{survey2}_{i,t}\right) + \gamma_4 \mathbf{time}_t + u_{i,t}$$

$$\text{for } i = 1,\ldots,N, \quad t = 8301,\ldots,9612$$

Investigating adaptive learning:

$$\left|\pi_{t+12} - E_{i,t}\pi_{t+12}\right| = \beta_0 + \beta_1 \text{survey2}_{i,t}$$

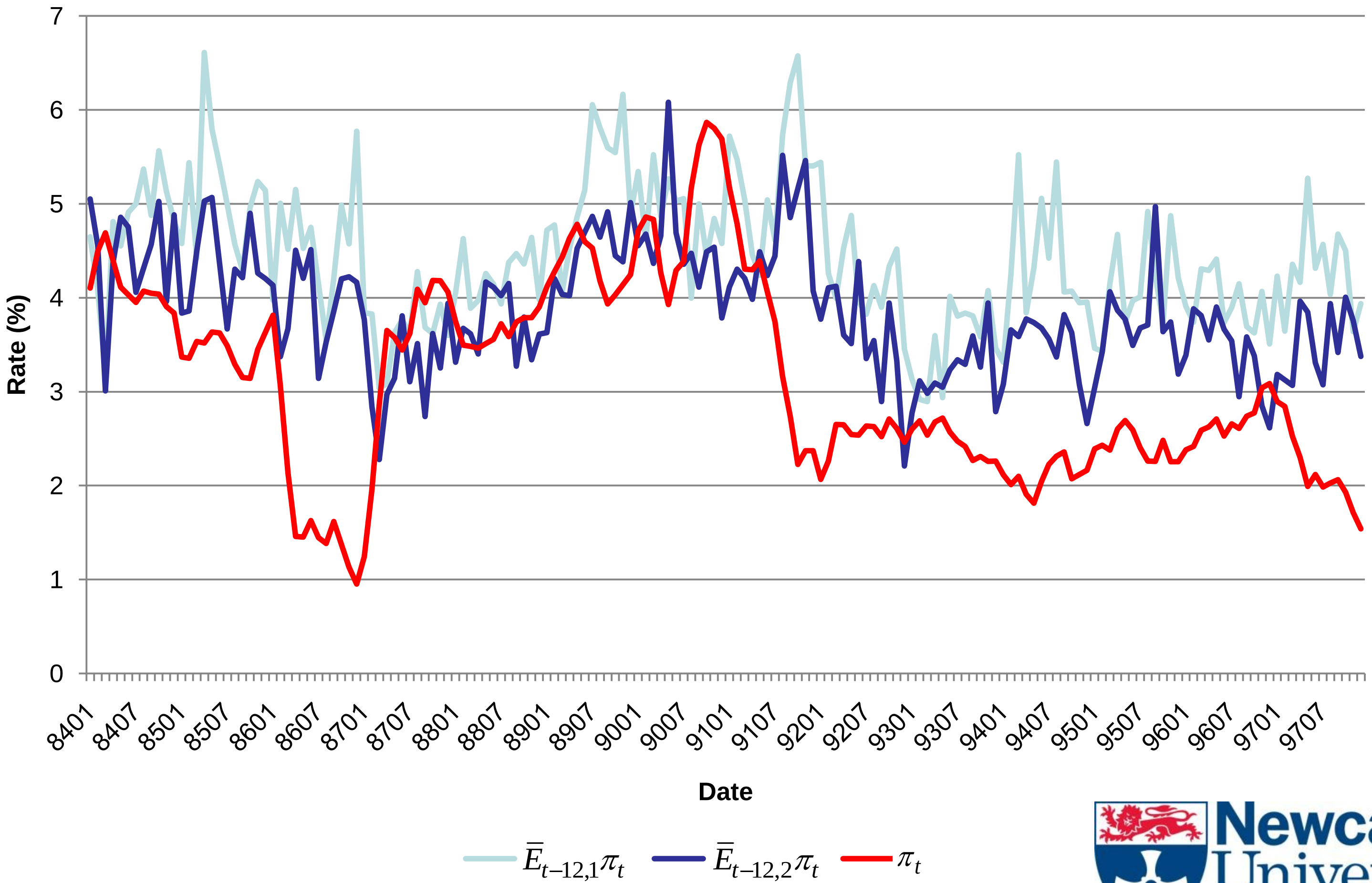
$$\text{for } i = 1,\ldots,N, \quad t = 8301,\ldots,9612$$

Newcastle University

In all cases:

- π_t is taken to be the CPI-U rate (calculated using methods at the time of release)
- 6-month reinterview horizon will mean that forecast errors will not be observed at time of reinterview
- Demographic characteristics are fixed at first-interview

Inflation Forecast versus Actual Inflation



Estimation Issues

Potential issues:

1. Second interview attrition – 30% dropout
2. Unobserved heterogeneity – innate biases and misconceived experiences
3. Forecast error distribution non-Normality

Particular problem in dealing with these issues simultaneously!

Newcastle University

1. Attrition

Assuming 'missing at random' (Rubin 1976) mechanism, two methods appropriate:

- Heckman (1979) correction

  Sample selection model estimated and used to produce 'Mills ratio' which is included to correct coefficient estimates. Standard errors corrected for induced heteroskedasticity. Problem of applicability with other estimation issues.

- Inverse Probability Weighting (Robins, Rotnitzky, Zhao 1994 and Wooldridge 2002)

  Sample selection model estimated and used to weight variables by inverse of this probability. Those cases less likely to be reobserved receive more weight. No further correction to standard errors required.

Newcastle University

Model propensity for attrition and estimate using Logit:

$$\Pr\left(\text{not-reinterviewed}_{i,t}\right) = \Phi\left(\alpha_0 + \alpha_1 \textbf{characteristics}_i + \alpha_2 \textbf{time}_t + v_{i,t}\right)$$

for $i = 1, \ldots, N, \quad t = 1, \ldots, 162 \ \left(8301, \ldots, 9606\right)$

Characteristics included guided by work of Vandecasteele and Debels (2007), augmented with survey date dummies.

Base group is middle income, 35-54 old white males, from single occupancy households without children, average educated with normal interview from the most experienced interviewer.

Newcastle University

	Coefficient	Std. Err.	Significance
Respondent characteristics			
Low income	0.203	0.0350	***
High income	-0.036	0.0327	
No HS Diploma	0.173	0.0386	***
Some College	-0.171	0.0311	***
College Degree	-0.325	0.0305	***
Age 34-	0.179	0.0288	***
Age 55+	0.036	0.0350	
Female	0.018	0.0346	
Non-head of household	-0.006	0.0374	
Non-white	0.327	0.0456	***
Non-white x Female	0.032	0.0599	
Separated	0.239	0.0670	***
Married	-0.124	0.0374	***
Widowed	-0.183	0.0590	**
Divorced	-0.003	0.0464	
North East	0.122	0.0343	***
South	0.082	0.0302	**
West	0.169	0.0340	***

Notes: *** denotes significance at 0.1% level, ** at 1% level.

	Coefficient	Std. Err.	Significance
Interview characteristics			
Interview Breakoff	0.618	0.0933	***
Interview Interrupt	0.261	0.0421	***
Interview Length > 45min.	-0.165	0.0496	***
Calls 2+	0.072	0.0349	*
Calls 5+	0.326	0.0373	***
Calls 10+	0.495	0.0333	***
Initial Coversheet Refusal	0.387	0.0358	***
Constant	-3.389	0.1379	***
Joint Hypothesis Tests	**Statistic**	**DoF**	**Significance**
Interviewer	35.76	18	**
Survey Month	938.31	161	***
Demographics	700.45	20	***
Interview characteristics (exc. interviewer)	677.78	7	***
All coefficients (exc. constant)	2796.02	206	***

Notes: *** denotes significance at 0.1% level, ** at 1% level, * at 5% level.

Characteristics which

– increase reinterview propensity:

Higher than average educated, married or widowed individuals, those with long initial interviews

–reduce reinterview propensity:

Low income, below average educated, non-white, separated, residents of regions other than North Central, interview breakoffs and interruptions.

Regression significance suggests attrition is a significant problem – bias is imposed by including second interview data in aggregate analysis!

Newcastle University

2. Individual-Level Correlations

Error term in learning model contains an ‘individual’ effect $u_{i,t} = a_i + \varepsilon_{i,t}$, which:

- Captures individual traits and misconceptions yet is unobservable
- Causes correlation between forecasts – uncorrected standard errors incorrect
- Treated as a ‘random effect’ since uncorrelated with all other ‘dummy’ regressors

Accommodation of correlation (and possible heteroskedasticity) using Rogers (1991) cluster-robust covariance matrix approach.

Newcastle University

Cluster-robust covariance matrix calculated as:

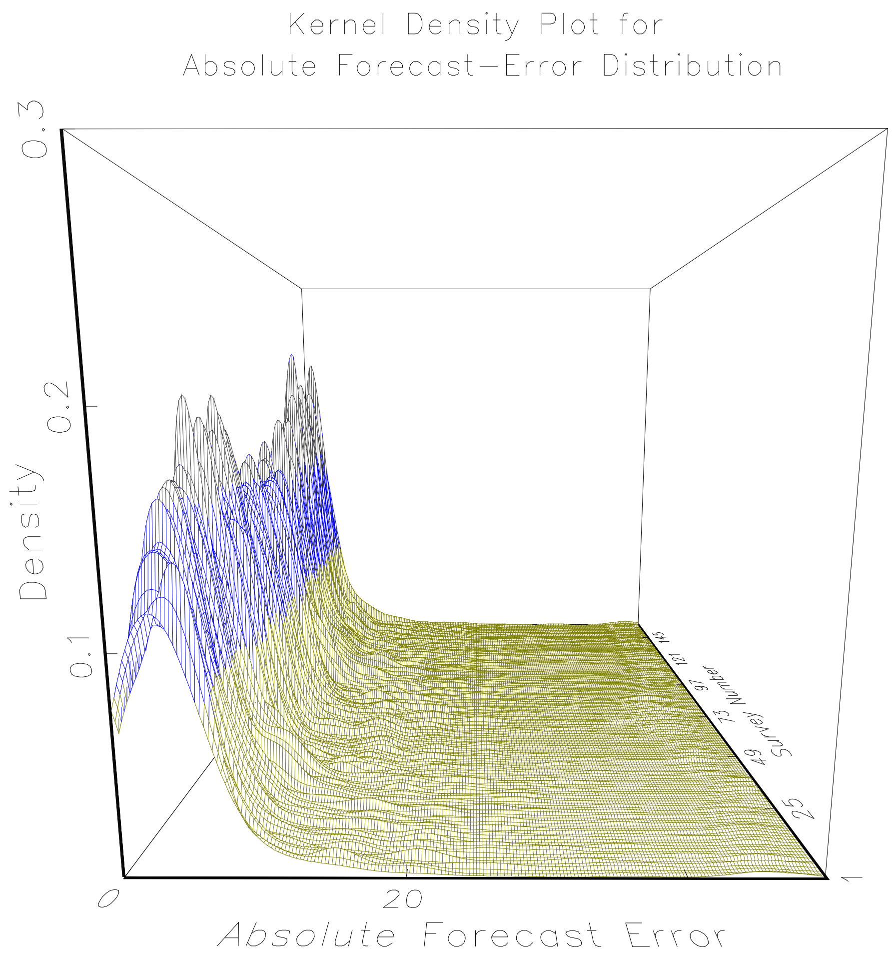
$$\hat{V}_{\hat{\gamma}}^{CR} = h(X'X)^{-1}\left(\sum_{i=1}^{N}(X_i'\hat{u}_i)(X_i'\hat{u}_i)'\right)(X'X)^{-1}$$

where *X* is the full sample regressor matrix, X_i is the regressor matrix pertaining to that cluster, and u_i are the OLS residuals from that cluster and where

$$h = \frac{N}{N-1} \cdot \frac{Ns}{Ns-K}$$

Newcastle University

3. Distribution



Non-normality resulting from:

- Positive skew in individual forecasts
- Underlying (signed) forecast error distribution
- Absolute value of forecast errors

Asymptotic *t* values will possibly be incorrect

Standard truncated-Normal regression techniques invalid.

Cluster-robust bootstrap *t*-statistics (Cameron, Gelbach and Miller 2008) possible solution:

1. Compute coefficient estimates from original weighted sample $\hat{\gamma}_k$
2. Draw a random new sample (with replacement) of *N* individuals
3. Estimate coefficient estimates, $\hat{\gamma}^*_{k,b}$ and associated standard errors $s^{*CR}_{\hat{\gamma}_k}$ using Rogers cluster-robust method
4. Repeat above 200 times and produce a vector of 200 bootstrapped *t*-statistics using $\hat{w}_{k,b} = \frac{\hat{\gamma}^*_{k,b} - \hat{\gamma}_k}{s^{CR}_{\hat{\gamma}^*_{k,b}}}$
5. Take the 2.5 and the 97.5 percentiles from this vector to compute the 95% *t*-statistic confidence interval around zero.

Newcastle University

6. Reject non-significance of regressor if observed *t*-statistic lies outside the 95% *t*-statistic confidence interval

Comparing bootstrapped *t* confidence interval with conventional 95% *t*-statistic confidence interval [-1.96, 1.96] illustrates influence of the non-Normality on inference.

Newcastle University

Issue Resolution Summary

1. Attrition – inverse probability weighting
2. Unobserved heterogeneity - Rogers (1991) cluster-robust standard errors
3. Distribution non-normality – Cameron, Gelbach and Miller (2008) cluster-robust bootstrap procedure

Newcastle University

Results

Models estimated (recap):

1. Adaptive learning:

$$\left|\pi_{t+12} - E_{i,t}\pi_{t+12}\right| = \beta_0 + \beta_1 \text{survey2}_{i,t}$$

2. Learning heterogeneity:

$$\left|\pi_{t+12} - E_{i,t}\pi_{t+12}\right| = \gamma_0 + \gamma_1 \mathbf{demog}_i + \gamma_2 \text{survey2}_{i,t} + \gamma_3\left(\mathbf{demog}_i \times \text{survey2}_{i,t}\right) + \gamma_4 \mathbf{time}_t + u_{i,t}$$

where in both cases, $i = 1, \ldots, N, \;\; t = 1, \ldots, 162 \;\; (8301, \ldots, 9612)$

1. Adaptive Learning

	Coefficient	***t*-statistic**	**Significance**	**Bootstrap 95% *t*-statistic Interval**
Constant	3.766	146.253	***	[-2.153,1.788]
Survey2 indicator	-0.531	-14.360	***	[-1.963,2.042]

Note: *** denotes significance at 0.1% level.

Second interview responses more accurate by 0.5% across all respondents.

- Learning is prevalent at the level of individual consumers
- Following Orphanides and Williams (2004), raised awareness of inflation results in better decision making.
- Agents learn from past errors (even though those errors are unobserved)

Newcastle University

2. Learning Heterogeneity

Is forecast accuracy dependent on the demographic characteristics of the individuals?

- Base group is average income, 35-54 year old, white, male individuals in single occupancy households with no children in October 1995
- Constant term does not show the general level of forecast accuracy – only for the base group in the base month
- Ranking of forecast accuracy and learning invariant to changes in base month

Newcastle University

	Coef.	*t*-statistic	Sig.	Bootstrap 95% *t*-statistic Interval
Constant	3.783	28.201	***	[-2.109, 1.872]
Low income	1.152	11.266	***	[-1.735, 2.087]
High income	-0.683	-13.763	***	[-2.048, 1.951]
Age 34 and under	0.270	4.506	***	[-2.028, 1.955]
Age 55 and over	-0.014	-0.203		[-2.228, 1.950]
Non-white	1.264	10.711	***	[-2.132, 1.928]
Female	0.960	20.458	***	[-2.293, 1.801]
Children in household	-0.065	-0.865		[-1.762, 1.666]
Multiple adults in household	0.179	2.866	**	[-1.976, 1.767]
Non-white x Female	1.109	5.765	***	[-2.335, 1.777]
Survey2 indicator	-0.210	-1.945		[-2.080, 1.775]
Low income x survey2	-0.292	-2.061	*	[-1.953, 1.860]
High income x survey2	0.134	1.822		[-1.931, 1.966]
Age 34- x survey2	-0.214	-2.477	*	[-1.762, 2.113]
Age 55+ x survey2	0.044	0.444		[-2.139, 1.892]
Non-white x survey2	-0.033	-0.177		[-2.667, 1.721]
Female x survey2	-0.137	-1.904		[-2.110, 1.891]
Children in HH x survey2	-0.206	-2.519	*	[-1.846, 1.991]
Multiple adults in HH x survey2	-0.047	-0.517		[-1.751, 1.770]
Non-white x Female x survey2	-0.682	-2.371	*	[-1.769, 2.129]

Notes: *** denotes significance at 0.1% level, ** at 1% level, * at 5% level.

Joint Hypothesis Tests	Statistic	DoF	
Demographics (first interview)	156.33	(9, 46911)	***
Demographics/survey2 interactions	5.34	(8, 46911)	***
Survey month indicator dummies	4.37	(167, 46753)	***

Significant differences in initial forecast accuracy:

- High income individuals are more accurate
- Low income, non-white, female and those in multiple occupancy households are less accurate
- Forecast differences are significant across all groups
- Significant time-variation (macroeconomic factors)

Second interview improvement in forecast accuracy:

- Significant for individuals in low income households, who are under 35, have children (initially unbiased), are both non-white and female
- Learning is jointly significant across all groups

Newcastle University

Conclusions

- Results suggest:
  - Initial forecast accuracy dependent on demographics
  - Learning (forecast improvement) significant for a range of demographic groups, not just those who are initially biased
- Implications for policy:
  - Central banks align expectations to target rate by stimulating learning
  - Information dissemination policy
- For the academic community:
  - Permit further analysis of dataset

Newcastle University

Thank you!

Any Questions!

robert.anderson@newcastle.ac.uk