

THE THEORY OF RATIONALLY HETEROGENEOUS EXPECTATIONS: EVIDENCE FROM SURVEY DATA ON INFLATION EXPECTATIONS*

William A. Branch

Previous work with survey data on inflationary expectations casts doubt on the Rational Expectations Hypothesis. In this paper, we develop a model of expectation formation where agents form their forecasts of inflation by selecting a predictor function from a set of costly alternatives whereby they may rationally choose a method other than the most accurate. We use this model to test whether survey data exhibit rationally heterogeneous expectations. Maximum likelihood is applied to a new discrete choice setting where the observed variable is continuous and the latent variable is discrete. The results show there is dynamic switching that depends on the relative mean squared errors of the predictors.

The failure of the Rational Expectations Hypothesis to account for survey data on inflationary expectations is well documented.¹ One widely cited explanation for this result is that agents lack the requisite sophistication to form expectations rationally (Evans and Honkapohja, 2001). Recent work by Evans and Ramey (1992, 1998) and Brock and Hommes (1997, 1998, 2000) show that with information costs it may be *rational* for agents to select methods other than rational expectations. In these models, each period agents update their previous expectation by weighing the costs and benefits of predictor choice. This paper examines survey data on inflation expectations for evidence of this type of dynamic predictor selection.

The criticism of rational expectations cited by Sargent (1993) and Evans and Honkapohja (2001), among others, is that it requires agents to possess too much knowledge. To form rational expectations agents must know the true structure and probability distribution of the economy. Evans and Honkapohja (2001), in justifying the use of a least squares updating rule instead of rational expectations, note that even econometricians must approximate the true structure of the economy. Given the inability of econometricians to estimate the economic model perfectly, it is unrealistic to expect agents to have such an ability. The traditional approach to modelling boundedly rational expectations assumes agents form their expectations by using an adaptive updating rule. In effect, agents behave *as if* they were econometricians.

* This paper has benefited tremendously from comments and suggestions by George Evans. I would also like to thank Tom Sargent, Bruce McGough, the Editor, three anonymous referees and seminar participants at the University of Oregon, University of Virginia and Cleveland Federal Reserve Bank for helpful comments.

¹ See Caskey (1985), Croushore (1993, 1997), Evans and Gulamani (1984), Frankel and Froot (1987), Jeong and Maddala (1996), Souleles (2002), Struth (1984), and Ulrich and Wachtel (1984), among others. An important exception to these papers is Keane and Runkle (1990). They take into account data revisions and forecast error correlation across individuals and fail to reject the Rational Expectations Hypothesis.

Evans and Ramey (1992) first considered how agents may form their expectations optimally but still not reach the rational expectations operator. In their model, agents choose whether to update their previous predictions subject to a revision cost. Brock and Hommes (1997) extend this work with their model of Adaptively Rational Equilibrium Dynamics (ARED).² In the ARED agents choose a predictor function from a set of alternative predictors that has a cost ordering which increases in a predictor's sophistication. The probability any predictor is chosen depends on its relative net benefit. They focus on a cobweb model with rational versus naive expectations and find that predictor switching may become complicated if agents have a high 'intensity of choice' between predictors.³ Branch (2002) extends this paper to a choice set that includes adaptive expectations.⁴

This paper extends the models of rationally heterogeneous expectations of Brock and Hommes (1997) and Branch (2002). We assume agents forecast 12-step-ahead annual inflation rates by choosing a predictor function from a set of costly alternatives. In the empirical examination these predictor functions must be econometrically inferred since they are not directly observed. For this reason agents are identified with a small number of forecast methods. The three methods used are a VAR forecast, adaptive expectations and naive expectations. These were chosen because of their long history and prominence in dynamic macroeconomic research.⁵ The proportion of agents that select a predictor are a function of its relative mean squared error (MSE). We further assume that when agents report their expectations they report the value returned by their forecast method plus an idiosyncratic shock. This shock accounts for many of the idiosyncrasies documented in Bryan and Venkatu (2001*a, b*) and Souleles (2002). Our approach is also related to Ball (2000) who shows that optimal univariate expectations in a multivariate New-Keynesian model can explain some of the empirical irregularities exhibited by New-Keynesian models under rational expectations.

Our approach extends the econometric literature on discrete choice by applying maximum likelihood to a new setting where the observed variable is continuous and the latent variable is discrete. Agents choose a predictor function but report to the surveyor a continuously valued expectation. We develop an approach which assesses the likelihood an individual survey response is made by a specific predictor function given the reported idiosyncrasies. Maximum likelihood is then applied to determine the dynamic properties of predictor selection.

By using survey data from the Survey Research Center at the University of Michigan we analyse the dynamics of predictor selection and provide the first empirical test of an adapted ARED approach to expectation formation. Moreover, this paper is the first application of a first-principled model of heterogeneous beliefs applied to survey data on inflation expectations. The focus of the paper,

² Brock and Durlauf (2001) consider a more general model where choices are made based on expectations of other agents' choices.

³ Brock *et al.* (2001) and Hommes (2001) apply this work to financial markets in order to explain two stylised facts (volatility clustering and long memory).

⁴ Goeree and Hommes (2000) extend Brock and Hommes (1997) to the case of nonlinear demand and supply in the cobweb model.

⁵ The VAR forecast is intended to be a boundedly rational method 'in the spirit' of rational expectations.

though, is closely related to Carroll (2003). Carroll tests whether the median expectation in the Michigan survey is a distributed lag of some measure of a sophisticated forecast. He finds that information dissipates through the economy producing time-varying expectations. The results here provide another interpretation for expectation dynamics.⁶

Although there are other surveys of inflation expectations (e.g. Survey of Professional Forecasters, Livingston survey), this paper uses the Michigan survey because it most closely matches the theoretical model we set forth. One of the main advantages to the Michigan survey is that it consists of households rather than professional forecasters. All households make financial decisions, many of them requiring forward-looking expectations of inflation. That a fraction of them may appear misinformed is completely in line with the theory that, at times, agents may not find it worthwhile to forecast well.

The results show agents do, in fact, dynamically select predictor functions. The proportion of agents that use each predictor varies inversely with the predictor's MSE. As the mean squared error of one predictor increases *vis-à-vis* the other predictors, the proportion of the sample that uses that predictor will decrease. When compared to alternative models of homogeneous and heterogeneous expectation formation, the model here provides a better prediction of the survey data. These results suggest that the rejection of the rational expectations hypothesis when applied to survey data is not because agents blindly follow an *ad hoc* rule. Rather, at times it is not worthwhile for them to invest the effort to use more complicated predictor functions. Agents are *rationally heterogeneous* in the sense that each predictor choice is optimal for themselves. Strictly speaking, their expectations are boundedly rational and consistent with optimising behaviour.

We also find that agents have inherent predispositions to use one predictor over another. Agents switch between methods as the forecast accuracy warrants, however, there is inertia to these switches. This is because forecast error must pass some threshold before agents will abandon their previously selected prediction mechanism. On average, agents use the VAR method more often than adaptive or naive.

This paper proceeds as follows. Section 1 introduces the theory of rationally heterogeneous expectations. Section 2 discusses the estimation of predictor functions. Section 3 presents the likelihood function for a discrete choice setting where the observed variable is continuous and the latent variable is discrete. The results are presented in Section 4. Finally, Section 5 concludes.

1. Rationally Heterogeneous Expectations

The motivation for rationally heterogeneous expectations is the Adaptively Rational Equilibrium Dynamics (ARED) of Brock and Hommes (1997).⁷ Agents

⁶ Evidence of heterogeneous expectations in empirical models can also be found in Baak (1999) and Chavas (2000).

⁷ Sethi and Franke (1995) examine predictor choice in an evolutionary model with strategic complementarities.

are boundedly rational in the sense that they choose from a variety of methods with which to form expectations, rational expectations being only one choice. Agents' choices of predictors depend on the relative costs and successes of each predictor and preferences over a measure of relative success. This Section presents the model.

1.1. General Model of Predictor Selection

Consider a simple model of the economy where agents must forecast inflation for the upcoming year.⁸ Each agent i chooses in time t their expectation of inflation, $\pi_{i,t}^e$, from a set of predictor functions

$$\mathcal{H}(\pi^t) = \{H_1(\pi^t), H_2(\pi^t), \dots, H_K(\pi^t)\}, \quad (1)$$

where $\pi^t = (\pi_{t-1}, \pi_{t-2}, \dots, \pi_0)$ and $H_j: \mathbb{R}^t \rightarrow \mathbb{R}$ for $j = 1, \dots, K$. Each predictor carries a non-negative cost which agents must pay to use the function. *A priori*, we would expect there to be a hierarchical cost structure that depends on the 'sophistication' or accuracy of the predictor. If costs are interpreted as calculation time, then the more sophisticated technique should be more expensive. The framework is sufficiently general, though, to allow for other cost ordering possibilities.⁹ We will not impose a cost structure since we seek to estimate the actual hierarchy.

The ARED assumes the probability an agent selects a certain predictor is determined by a discrete choice model.¹⁰ There is an extensive literature that models individual decision making as a discrete choice in a random utility setting (Manski and McFadden, 1981). The model here is consistent with that literature. The proportion of agents in the population that use a certain predictor is increasing in its relative net benefit (i.e. success minus cost), and depends positively on the intensity with which agents react to increases in relative net benefit.

Let $U_{j,t}$ denote the relative net benefit of predictor j at time t . The proportion of agents using predictor j in time t , $n_{j,t}$, follows

$$n_{j,t} = \frac{\exp(\beta U_{j,t})}{\sum_{l=1}^K \exp(\beta U_{l,t})} \quad \text{for } j = 1, 2, \dots, K \quad (2)$$

where $\beta \in \mathbb{R}_+$ is the 'intensity of choice' parameter. The form of (2) is called a multinomial logit (MNL) and is well known in the discrete choice and ARED literature. Larger values of β lead agents to more quickly switch predictors for small changes in relative net benefit. The neoclassical case would have $\beta = +\infty$ and $n_{j,t}$ would become a point mass on the predictor with the highest return. A finite 'intensity of choice' could be an explanation for the rejection of rationality in survey data on inflation expectations. Agents may not fully respond to changes in

⁸ This section closely follows Brock and Hommes (1997) and Branch (2002).

⁹ The model set forth by Brock and Hommes (1997) assumes costs are exogenously determined. Future work will consider a more general framework where costs and prediction proportions are jointly determined.

¹⁰ Specifically, a multinomial logit model.

relative net benefit and so *rationally* choose a method other than rational expectations.

1.2. An Operational Model of Predictor Selection

This subsection places bounds on the predictor choice set in order to facilitate empirical analysis.¹¹ Branch (2002) examines predictor selection under the three methods most commonly used in macroeconomics: rational expectations; adaptive expectations; and naive expectations. Adding many more predictors extends the analysis in a way that strays from the main point: there are varied degrees of sophistication in expectations with the heterogeneity depending upon benefits and preferences over those benefits. Thus, we construct three methods: a multivariate method; a univariate method based on a good choice of a smoothing parameter that weights past inflation heavily; and, a naive univariate method based on the most recent inflation. Each method is designed to be a representative of these larger classes of predictor functions.

The model under consideration departs significantly in four ways from the model of Brock and Hommes (1997). First, we consider a stochastic environment where agents forecast inflation. Second, noise will be introduced in terms of small trembles to forecasts after predictor selection is made. Third, agents' set of available predictors is restricted to a VAR forecast, adaptive expectations, and naive expectations. Fourth, relative net benefit is defined as the mean square error of forecast accuracy net of predictor costs. Each of these will be discussed below.

The set of predictor functions is assumed to be

$$\mathcal{H}(\pi^t) = \{\pi_{t-1}, AE_t, VAR_t\} \quad (3)$$

where each predictor is defined as follows. The naive predictor π_{t-1} sets next period's inflation forecast as the previous period's inflation rate. The adaptive expectations predictor, AE_t , is calculated as

$$AE_t = AE_{t-1} + \gamma(\pi_{t-1} - AE_{t-1}) = \gamma\pi_{t-1} + (1 - \gamma)AE_{t-1} \quad (4)$$

for some $\gamma \in [0, 1]$. Finally, the VAR predictor VAR_t is the linear projection from the VAR

$$\mathbf{y}_t = \alpha + \Omega_1 \mathbf{y}_{t-1} + \Omega_2 \mathbf{y}_{t-2} + \cdots + \Omega_p \mathbf{y}_{t-p} + \mathbf{u}_t \quad (5)$$

where \mathbf{y}_t is the $(n \times 1)$ vector of inflation and other relevant variables that produce the best forecast of inflation, and $\mathbf{u}_t \stackrel{iid}{\sim} N(\mathbf{0}, \mathbf{Z})$. The next section discusses the variables included in the VAR estimation.

A brief note on these predictor methods is useful. Our definition of naive expectations follows the learning literature and highlights that the most naive forecasting rule assumes that recent observations are a continuing trend. Adaptive expectations in the form of (4) has a venerable history dating back to Cagan (1956), Friedman (1957), and Nerlove (1958). It was routinely used in pre-rational

¹¹ Brock and de Fountnouvelle (2000) considers the case where the number of predictors is infinite and predictor selection of a certain predictor is a random draw from a distribution over all predictors.

expectations era models and has found a resurgence in the adaptive learning literature (Sargent, 1999). Ideally, our VAR forecast would correspond to rational expectations. However, in a world of heterogeneous information sets and an unknown true distribution for the economy, rational expectations is impossible to observe. The VAR predictor is a boundedly rational predictor that is ‘in the spirit’ of rational expectations. This approach is consistent with Carroll (2003) who assumes the ‘rational’ forecast to be the median response from the Survey of Professional Forecasters. Presumably, professional forecasters use some form of a multivariate model.

From the survey data we observe the continuously-valued expectation. However, we are interested in the agents’ choice of which predictor function to use. The observed expectations of individuals are assumed to be given by the following

$$\pi_{i,t}^e = H_j(\pi_t) + v_{i,t}, \quad (6)$$

where $H_j(\pi^t) \in \{\pi_{t-1}, AE_t, VAR_t\}$, and $v_{i,t} \stackrel{iid}{\sim} N(0, \sigma_v^2)$.¹² The stochastic term in the individual observed expectations can be interpreted as small trembles to agents’ belief formation, perhaps from a lack of full attention to the data or to forecasts made from professionals. After purchase of a given predictor, agents make an adjustment to the data and report an expectation that is their perception of future inflation.¹³ These mistakes are a mean zero process that are independent of the predictor an agent chooses.¹⁴ These trembles are crucial in identifying which predictor an individual is using at any given time. Note (2)–(6) define a testable model of expectation formation.

After using a predictor in time t , agents observe the actual inflation rate and assess the success of their predictor choice by weighing the benefits and costs.¹⁵ The benefits are defined as a low mean square error,

$$MSE_{j,t} = (1 - \delta)MSE_{j,t-1} + \delta(\pi_{i,t}^e - \pi_t)^2 \quad (7)$$

where π_t is the observed inflation rate in time t . If $\delta = 1$, $MSE_{j,t}$ corresponds to the squared error for predictor j in time t only. For $0 < \delta < 1$, $MSE_{j,t}$ is a geometrically weighted average of past forecast errors. Different values of δ correspond to different degrees of myopia when assessing forecast accuracy. We will discuss below how we determine the value of δ in our empirical procedure. The costs for the

¹² Fische and Idson (1990) also assumed expectations were normally distributed.

¹³ The distinction between an expectation derived from economic data and a perception was highlighted by Bryan and Venkatu (2001*a, b*). They note that one explanation for wide dispersions in survey data on inflation expectations is because agents have different perceptions of what the inflation rate is in their personal market basket. This paper formally models an expectation as the output from a forecasting function plus a perception adjustment made by way of a stochastic term.

¹⁴ The interpretation given to these shocks is they account for consumer idiosyncrasies. We do not allow for these shocks to be predictor dependent. If one thinks of these shocks as corrections for differences in market baskets then the noise will be the same regardless of predictor choice

¹⁵ The success of their predictor choice is relative to the other alternatives. Implicitly, agents know the success of the other predictors to be able to calculate the relative success of their own choice.

VAR, adaptive, and naive predictors are C_V , C_A , C_N , respectively.¹⁶ The assumption of the theoretical model is that the costs follow the hierarchy $C_V \geq C_A \geq C_N \geq 0$.¹⁷

The success of a given predictor is¹⁸

$$U_{j,t} = -(MSE_{j,t} + C_j). \quad (8)$$

We assume this form because it is the most tractable measure of success and is consistent with the measure employed in Brock and Hommes (1997). It is important to note that agents respond adaptively to past observed forecast errors. In $U_{j,t}$ we are measuring costs in MSE units. Our model, and the ARED, do not allow for costs (or benefits) to vary across individuals, though, we intend to account for this possibility in future research. In this model we can justify homogeneous costs across predictors since the random utility assumption of the multinomial logit term accounts for unobserved individual characteristics. The probability an individual selects a predictor j is given by,

$$n_{j,t} = \frac{\exp\{\beta[-(MSE_{j,t} + C_j)]\}}{\sum_{k \in \{N,A,V\}} \exp\{\beta[-(MSE_{k,t} + C_k)]\}}. \quad (9)$$

Equations (6) and (9) will be used to estimate the parameters β and C_j , $j \in \{N,A,V\}$. Of course, it is also clear that costs can impose inertia on the ARED. If the costs in (9) are large relative to the mean squared errors then there will be no dynamics in n_j . The ARED then is a theory which shows the relationship between optimal switching decisions and costs or inertia. In Brock and Hommes (1998) it is shown that in some models interesting predictor selection dynamics may exist even when $C_j = 0, \forall j$. In the empirical work we consider specifications with and without costs.

2. Predictor Functions

The Survey of Consumer Attitudes and Behavior asks survey participants to form twelve-month-ahead forecasts of annual inflation. Using the seasonally adjusted Consumer Price Index for all urban consumers, annual inflation is defined as

$$\pi_t = 100 \times \ln\left(\frac{CPI_t}{CPI_{t-12}}\right). \quad (10)$$

This annualised rate, though, is not the most widely published inflation rate. Most newspapers report the monthly (annualised) inflation rate. In fact, the Labor Department's own press releases and charts that are reported by major media outlets primarily cite the month to month changes in the CPI. Since monthly inflation is the most widely reported we will use it as the input to the predictor functions. Monthly inflation at an annual rate is calculated as

¹⁶ Future research will consider a more general cost structure that allows for random cost shocks.

¹⁷ The actual cost may be different across individuals if some people have higher calculation costs. That is, some people may have greater cognitive ability that allows them to form a sophisticated expectation at a lower cost. Heterogeneous cognitive abilities could be introduced through stochastic costs.

¹⁸ MSE in the form of (7) as a measure of fitness immediately follows from a quadratic loss function of the form $\delta \sum_{k=0}^t (1 - \delta)^k (\pi_{j,t-k}^e - \pi_{t-k})^2$.

$$\pi_t^m = 1200 \times \ln \left(\frac{CPI_t}{CPI_{t-1}} \right). \quad (11)$$

We assume all forecast methods map past realisations of monthly inflation to a forecast of future annual inflation. Note that because the CPI is seasonally adjusted (11) is as well.

2.1. Naive Predictor

With the naive predictor agents expect that the most previous realisation of inflation will continue again in subsequent periods. Because the simplest way to forecast inflation is to recall the latest, easily accessible inflation data, and based on the way the Bureau of Labor Statistics (BLS) releases data, we estimate naive expectations as the annualised monthly inflation rate. Naive expectations are defined and estimated as $\pi_{t+12}^{e,N} = \pi_t^m$ where π_t^m is defined as in (11). At first glance it may appear that naive expectations are ‘too naive’. Our formula is designed to seek out what proportion of agents take the most widely reported (and most noisy) data and assume it will persist into the future. Whether this definition is too simple is an empirical question. We will test for a non-zero proportion of naive agents rather than impose it *a priori*.

2.2. Adaptive Expectations

Although adaptive expectations are more ‘sophisticated’ than naive expectations, they should still be treated as a relatively simple predictor.¹⁹ The annualised monthly inflation series is used to estimate adaptive expectations for the same reasons as for the naive predictor. The recursion in (4) is estimated using the monthly inflation data over the period 1947.01–1999.04. We calculate γ , by simulating (4) for various values of γ setting $\pi_0^{e,A} = \pi_{1947.01}^m$. The γ actually used in the predictor is determined by

$$\gamma = \min_{\gamma} [MSE] = \min_{\gamma} \left\{ \frac{1}{T} \left[\sum_{t=0}^T (\pi_t^{e,A} - \pi_{t+12})^2 \right] \right\} \quad (12)$$

where T is the total number of time periods in the sample. We estimate γ over this sample, even though, it begins prior to and ends subsequent to the survey sample, to highlight that adaptive expectations is the best recursion over all sample periods.²⁰ A grid search over $\gamma \in [0, 1]$ revealed that $\gamma = 0.216$ was the value that minimises the mean square error.

Subsection 2.4 discusses other possible forms of the adaptive predictor. Our equation (4) with $\gamma = 0.216$ is intended to be an approximation of all adaptive models and we will identify all agents that use predictors ‘close’ to (4) as adaptive.

¹⁹ In fact, for $\gamma = 1$ adaptive expectations and naive expectations are equivalent.

²⁰ Results are robust to a predictor which is instead a recursion over the sample period as the survey data.

In order to estimate an econometrically tractable model we need a metric for choosing γ . The most reasonable metric, in our view, is the mean square error across all periods.

Given this iterative estimation technique, the adaptive predictor, as will be used later in identifying the survey expectations, is

$$\pi_t^{e,A} = 0.216\pi_{t-1}^m + 0.784\pi_{t-1}^{e,A}. \quad (13)$$

This specification reflects that predictors map from past realised monthly inflation into an annual forecast. It should be particularly noted that (13) is measured as an annual rate. Moreover, as implied by (12), equation (13) is used to forecast π_{t+12} . At time t agents use all available information to update (13) and then use the new estimate as a forecast of π_{t+12} .

2.3. VAR Predictor

The traditional definition of a rational expectation is the mathematical expectation conditional on all available information. Conceivably, each individual in an economy could have a unique expectation if information sets are disparate. Because of this we examine a VAR forecast as a boundedly rational alternative to rational expectations.²¹ In estimating the VAR forecast, we will estimate an agent's best forecast of inflation using all relevant information. This approach is consistent with Carroll (2003).

The overriding consideration for each specification was whether it produced a better forecast than the naive and adaptive predictors. A VAR model using inflation and other relevant information provides the most accurate forecast of inflation and is inspired by most reduced-form linearised models of the economy. The VAR model includes monthly inflation, the monthly unemployment rate, the monthly growth rate of M1, and the 3-month Treasury bill rate.²² One could argue that another model is preferable on theoretical or econometric grounds. However, for the purposes of this study having a good representative of all multivariate forecasting methods suffices. The optimal lag length, as defined by the minimum Akaike Information Criterion, is twelve months. Table 1 details results from this VAR specification.²³

Pesaran and Timmermann (1995) present a recursive approach to model selection that is somewhat related to our model. In their paper agents choose one (possibly misspecified) model with which they forecast excess stock returns. They

²¹ We do allow for a specification that includes rational expectations. There we define rational expectations as actual inflation plus a serially uncorrelated disturbance. We demonstrate that the empirical results are consistent with this specification. The specification here is preferred since each predictor is consistent with the same data generating process.

²² Again, monthly inflation is used over annual inflation because it is the most widely available in press releases.

²³ We also ran a specification with annual rather than monthly inflation. Additionally, to allow for time-varying parameters we used Recursive Least Squares. This specification, surprisingly, does not do better than adaptive expectations. This VAR forecast also is no worse than other popular methodologies. For example, the Bayesian Prior estimation of Litterman (1986) does not fair better than this simple VAR. This Litterman technique is useful for dealing with over parameterisation and degrees of freedom problems by specifying a prior belief about each parameter.

Table 1

*Estimated VAR Expectations Predictor Function Tests and MSEs*MSE = 2.6, $R^2 = 0.815639$, Sample 1947.01–1999.04

Hypothesis	F-Statistic	Probability
Coefficients on lagged unemployment jointly zero	1.612613	0.08512
Coefficients on lagged unemployment jointly zero	2.62539	0.002115
Coefficients on lagged t-bill rate jointly zero	5.17201	<0.000001

assume that the parameters are unknown and agents must estimate the parameters each period. In this paper we assume agents use fixed parameter models that are the best forecasts possible if all information is available. As Pesaran and Timmermann (1995) note, this is a dubious assumption indeed. We make the assumption to keep the analysis as simple as possible.²⁴ We remark that we ran a specification of the model where the parameters are updated recursively (recursive least squares) and where the parameters are time-varying (Kalman Filter) and the qualitative results are robust in each case.²⁵ In the next Section we briefly comment on the results when the VAR has time-varying parameters.

Using the results from the VAR estimation we forecast inflation over the next twelve months as follows. Each period, the VAR model is used to produce twelve-step ahead forecasts, which are then summed up to produce the annual forecast. That is, using the VAR for each t we forecast $\hat{\pi}_{t+1}^m, \hat{\pi}_{t+2}^m, \dots, \hat{\pi}_{t+12}^m$. Then the forecast of annual inflation over the next twelve months is calculated as

$$\hat{\pi}_{t+12} = \sum_{i=1}^{12} \hat{\pi}_{t+i}^m. \quad (14)$$

Comparing these forecasts to 12-month ahead annual inflation gives the desired notion of accuracy. With a MSE of 2.6 this VAR betters adaptive expectations and so forms the specification we use.

2.4. Discussion of Predictor Functions

We do not claim that our set of predictor functions is exhaustive. As the first attempt to characterise expectations in this manner we must make modelling choices consistent with the underlying theory. This subsection is a defence of these modelling decisions and a brief account of robust alternatives. We do acknowledge, however, that there are other predictors and alternative means to estimating these predictors. We do not pursue them in this paper since our central message about expectation formation is unchanged with alternative predictors.

²⁴ In models of the US economy the question of drifting coefficients and volatilities is open (Cogley and Sargent, 2003). In future research it would be interesting to conduct our approach in a model where learning about the parameters matters and the dominance of particular predictors only holds in subsamples.

²⁵ The robustness in the recursive case bolsters our choice of VAR model since it betters adaptive expectations in out of sample forecasting.

The predictor functions, as we have defined them, map past monthly inflation to an annual inflation forecast. We made this assumption for three reasons. First, inflation data are primarily reported as monthly inflation at an annual rate. Second, our approach to bounded rationality assumes agents use the most widely available data as forecasting inputs. It is reasonable, in our view, for agents to use recent monthly trends as a barometer for the total change in prices over a year. That naive agents assume the most recent monthly price changes will continue over the next twelve is quintessential naive. Third, it is necessary to base forecasts on the noisier monthly inflation series in order for the three classes of predictors to return significantly different values.

It is appropriate, however, to wonder how sensitive the results are to our particular modelling choices. For example, would functions which turn past annual data into annual expectations change the results? To check for robustness we re-estimated the predictors using annual inflation as an input. We find that the qualitative results of this paper are robust to these predictors.

By careful examination of (3)–(6) it is clear that under appropriate restrictions the VAR can reduce to adaptive or naive expectations. We tested the hypothesis that the parameters on lagged money growth, interest rates, and unemployment are jointly zero. We reject this hypothesis. Moreover, the mean squared errors produced by naive forecasts are distinct from the adaptive predictor. Even though the VAR is on average a better predictor, there are periods where the adaptive or naive predictor dominates.

One might argue that the problem is not monthly inflation as an input but adaptive expectations as the weighted average of past monthly annualised inflation rates. We argue that our approach is entirely consistent with the spirit of adaptive expectations as a univariate predictor distinct from a multivariate predictor such as the VAR. Our predictors then are consistent with those of Ball (2000).

One could replace adaptive expectations with a more sophisticated univariate model such as an ARMA model. In this case the same sum of twelve one-step ahead forecasts as in (14) could be used with an ARMA to form an expectation. Such a predictor, though, produces an MSE remarkably close to the VAR forecast and is not in the spirit of adaptive expectations. Though if one replaces the VAR with the ARMA the qualitative results still hold.

It is important to emphasise that the VAR may not correspond to rational expectations.²⁶ Without knowing the structure of the model it is quite conceivable that our model could be under parameterised. Instead its best to think of the VAR as ‘in the spirit’ of rational expectations. This approach is similar to the restricted perceptions equilibria of Evans and Honkapohja (2001) or self-confirming equilibria of Sargent (1999) where a misspecified model’s parameters are optimal and agents can not detect their misspecification. Our approach assumes that agents use models with optimally chosen parameters.

²⁶ This is particularly true in a heterogeneous world, as VAR agents do not account for the non-rational agents. In this sense the VAR expectations behave like the fundamentalists in Brock and Hommes (1998).

Much of the learning literature emphasises, however, that agents do not know these parameters and must statistically learn their values. Along such a learning path the economy's parameters will be time-varying. If the learning process converges then the parameters will be constant. In order to be certain that our results are robust both to where agents know the structure of the economy and do not, we estimated the model with time-varying VARs. We report on these results in the next Section.

It may also appear that we are neglecting rational expectations entirely. However, even though the VAR is not strictly rational it is in the spirit of rational expectations as the best multivariate forecasting strategy. As an empirical question we do consider a specification, as an alternative, which models rational expectations as actual future annual inflation plus noise as a predictor.²⁷ In this case, the cost estimate for the rational predictor should be significantly different from the VAR and adaptive predictors.

3. Empirical Model

3.1. *The Survey of Consumer Attitudes and Behavior*

The data come from a monthly survey of approximately 500 households conducted by the Survey Research Center (SRC) at the University of Michigan. The results are published as the Survey of Consumer Attitudes and Behavior. The raw coded data are available from the Inter-University Consortium for Promoting Social Research (ICPSR) over the period 1977.01–1993.12.²⁸

Each survey respondent is interviewed once and then reinterviewed six months later. In any given monthly survey there are approximately 55% new respondents, and 45% prior respondents. This paper is concerned with expectations of future inflation, and so the two relevant questions are:

1. During the next 12 months, do you think that prices in general will go up, down, or stay where they are now?²⁹
2. By about what percentage do you expect prices to go (up/down) on the average, during the next 12 months?

If respondents expect prices will not change (i.e. inflation is zero) then question 2 is coded as 0%. If respondents expect prices to go up, then question 2 is coded as the exact percentage that respondents provide. If respondents expect prices to go down, then question 2 is coded as the negative of the percentage that respondents provide. Agents are also allowed to respond that they don't know to

²⁷ This follows from the observation that under rational expectations forecast errors follow a martingale difference sequence.

²⁸ There are some missing months in 1977, and so the sample is restricted only to continuous periods. In this case, the sample goes from 1977.11–1993.12. The raw data is available through the ICPSR with a lag of about 6 years. We chose the sample 1977.11–1993.12 because it covers a sufficiently diverse spectrum of inflation volatility. There are no significant differences in survey responses over the period 1994.1–1996.12 as this is a period of relative price-level stability.

²⁹ Telephone operators are instructed to ask a clarifying question if respondents answer they expect prices to stay where they are now. That question is: Do you mean that prices will go up at the same rate as now, or that in general they will not go up during the next 12 months?

questions 1 and 2. Fische and Idson (1990) examine the implications of the 'don't know' response. In their analysis, the probability of answering don't know is different for question 2 than question 1, and depends on their answer to question 1. These responses are a very small proportion of the total sample and so were excluded from the study. However, because their numbers are so small any bias will be minimal.

This leaves a sample of 129,950 observations covering 187 time periods.³⁰ The mean response was 6.9550 with a standard deviation of 12.7010. The large standard deviation is accounted for by a few outliers that expect inflation to be greater than 40%. Excluding these responses does not change the qualitative results.

The Michigan survey is not the only ongoing survey of inflation expectations. Two notable alternatives are the Survey of Professional Forecasters and the Livingston survey. For several reasons these other surveys are not ideally suited to the theory of rationally heterogeneous expectations. First, these surveys are of professional forecasters and economists. It seems unlikely that any professional forecaster would ever use a univariate prediction method. The Survey of Professional Economists serve as proxies for 'rational' forecasts in papers such as Carroll (2003). Also, the Michigan survey focuses on households which is a segment of the economy whose behaviour is clearly more important than professional economists. Second, these other surveys are panel data sets. One might think that a panel is preferable but for the purpose of testing and estimating a version of the ARED of Brock and Hommes (1997) it is not. This is because the Brock-Hommes approach hypothesises that the proportion of agents that choose different forecast methods evolves with their relative successes, and makes no statement whatsoever about how individuals will change over time. A repeated cross-section is, therefore, ideal for examining their approach.

3.2. Empirical Model

Because of individual specific trembles the actual survey response differs from the estimated predictor value. This paper develops a new method to determine predictor types. This will be done by developing a likelihood function for individual survey responses under the assumption that the trembles are normally distributed.

Recall, that the actual observed survey response was given in (6)

$$\pi_{i,t}^e = H_j(\pi^t) + v_{it}, \quad (15)$$

where $H_j(\pi^t) \in \{\pi_{t-1}, AE_t, VAR_t\}$, and $v_{it} \stackrel{iid}{\sim} N(0, \sigma_v^2)$. The probability of actually using the j th predictor was given by the theoretical model as a MNL:

$$\Pr(j|U_{j,t}) = n_{j,t} = \frac{\exp\{\beta[-(MSE_{j,t} + C_j)]\}}{\sum_{k \in \{N,A,V\}} \exp\{\beta[-(MSE_{k,t} + C_k)]\}} \quad (16)$$

Since v_{it} is distributed normally, we can calculate the density of $\pi_{i,t}^e$ as

³⁰ It may seem odd that the predictors are estimated over 1947.01–1999.04 while the survey sample is 1977.11–1993.12. However, agents do not just live in 1977.11–1993.12. Also, by estimating predictors over the entire sample we construct the 'best' approximations to agents' true models.

$$P\left(\pi_{i,t}^e | MSE^t\right) = \sum_{l \in \{N,A,V\}} n_{l,t} P\left(\pi_{i,t}^e | j = l\right) \tag{17}$$

where

$$MSE^t = \{MSE_{N,t}, MSE_{A,t}, MSE_{V,t}\}$$

and

$$P\left(\pi_{i,t}^e | j\right) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left\{-\frac{1}{2} \left[\frac{\pi_{i,t}^e - H_j(\pi^t)}{\sigma_v}\right]^2\right\}.$$

Define $\mathcal{H}_t(\pi^t) = \{H_N(\pi^t), H_A(\pi^t), H_V(\pi^t)\}$. Since the sample changes each period, the probability of observing the sample is given by the following density function:

$$P\left(\pi_{i,t}^e, i = 1, \dots, N, t = 1, \dots, T | MSE^t, \mathcal{H}_t(\pi^t), t = 1, \dots, T\right) \tag{18}$$

$$= \prod_t \prod_i P\left(\pi_{i,t}^e | MSE^t\right) \tag{19}$$

$$= \prod_t \prod_i \left[\sum_{l \in \{N,A,V\}} n_{l,t} P\left(\pi_{i,t}^e | j = l\right) \right]. \tag{20}$$

Using this density function the log-likelihood function is

$$\mathcal{L} = \ln P\left[\pi_{i,t}^e, i = 1, \dots, N, t = 1, \dots, T | MSE^t, \mathcal{H}_t(\pi^t), t = 1, \dots, T\right] \tag{21}$$

$$\begin{aligned} &= \sum_t \sum_i \ln \sum_{l \in \{N,A,V\}} \frac{\exp\{\beta[-(MSE_{l,t} + C_j)]\}}{\sum_{k \in \{N,A,V\}} \exp\{\beta[-(MSE_k + C_k)]\}} \\ &\times \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left\{\frac{-1}{2} \left(\frac{\pi_{i,t}^e - H_l(\pi^t)}{\sigma_v}\right)^2\right\}. \end{aligned} \tag{22}$$

The log-likelihood function for the sample (21) demonstrates that the likelihood of observing a value $\pi_{i,t}^e$ depends on the relative mean squared errors of each of the predictors, the cost of each predictor, and a random shock that is distributed around the predicted values. The parameters β and σ_v , as well as the costs C_j for $j \in \{N, A, V\}$, must be estimated.

The parameter β is the most important parameter for our test of rationally heterogeneous expectations. Our theoretical model predicts that the proportion of agents using a certain predictor will adjust dynamically to changes in the relative mean squared error of a predictor. Hence, we expect the sign of β to be positive.

To a lesser extent, the estimation of the C_j s are important. Intuitively, we would expect the costs to be structured according to the ‘sophistication’ of the predictors; VAR expectations should be the most costly, followed by adaptive expectations and naive expectations. This conventional interpretation, though, assumes that there is not any inherent bias in predictor selection. The bounded rationality literature argues that very few agents have the capability, regardless of cost or mean squared error, to form rational expectations. Indeed, the vast literature on testing survey data rejects the rational expectations hypothesis. Conversely, those in the rational expectations camp would argue that most people do

form rational expectations, or at least expectations that result from a learning rule that shares the same structure as rational expectations. In the latter case, as Evans and Honkapohja (2001) frequently argue, agents act as econometricians. If there is a natural tendency for agents to behave like econometricians then we might expect to see a larger proportion of agents using the VAR predictor than can be justified by mean squared error. We call this a predisposition effect, since agents are predisposed to using a predictor in a way that does not depend on accuracy. This bias would be reflected in the cost term. Such an effect may seem counter to the theory but it is not. Brock and Hommes (1997) assume that costs increase in sophistication as a way to generate interesting dynamic behaviour. There is nothing in the foundations of the theory to assume *a priori* such an ordering. As Brock and Hommes (1997) clearly demonstrate the existence of costs imposes inertia or a predisposition effect. In their model it is that this inertia is the result of sophistication. We address this issue as an empirical question.

We are also interested in how well the model fits the data. Aside from the likelihood value itself, there are no useful descriptive statistics for ‘goodness of fit.’³¹ In order to conclude that the ARED model fits the data we propose alternative models and test for whether they fit the data better in a maximum likelihood sense.

4. Empirical Results

4.1. Empirical Procedure

The purpose of this paper is to estimate the parameters $\beta, C_N, C_A, C_V, \sigma_v$ in (21). The data set consists of individuals making inflation expectations in time t as well as characteristics of the predictor functions. In the previous section we derived the likelihood function that will be used in estimation.

The empirical problem to solve is

$$\begin{aligned} \max_{\beta, \{C_j\}_{j \in \{N,A,V\}}, \sigma_v} \mathcal{L} = & \sum_t \sum_i \ln \sum_{l=1}^3 \frac{\exp\{\beta[-(MSE_{l,t} + C_j)]\}}{\sum_{k \in \{N,A,V\}} \exp\{\beta[-(MSE_{k,t} + C_k)]\}} \\ & \times \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left\{\frac{-1}{2} \left(\frac{\pi_{i,t}^e - H_l(\pi^t)}{\sigma_v}\right)^2\right\}. \end{aligned} \tag{23}$$

It is well-known that the vector of parameters $[\beta, C_N, C_A, C_V]'$ may not be identified. To guarantee identification in the empirical analysis we will normalise one of the costs to zero. In the reported results we document the estimates for each possible normalisation.³²

³¹ The information matrix is used to construct standard errors of the maximum likelihood estimates.

³² We can establish identifiability more formally. Let $\theta^* = \max_{\theta} \mathcal{L}$. A necessary and sufficient condition for identification is $\nexists \theta \neq \theta^* \text{ s.t. } \mathcal{L}(\theta) = \mathcal{L}(\theta^*)$ (McFadden, 1984). Define $\theta^* = [\beta_1, \beta_2, C_V, C_A, C_N]'$ and consider an arbitrary θ . It follows

$$\Pr(\pi^e = \pi_V^e | \theta^*) = \frac{\exp(\theta^* x_1)}{\sum_j \exp(\theta^* x_j)} = \frac{\exp(\theta x_1)}{\sum_j \exp(\theta x_j)} = \frac{\exp[(\theta + \theta^*)x_1]}{\sum_j \exp[(\theta + \theta^*)x_j]}$$

for appropriately defined x_j . Without an additional constraint such as $C_j = 0$ for some j , there are many such θ .

There is a paradox between theory and reality that dictates how forecast accuracy is defined. We assume that agents input monthly inflation data to generate their forecast output because that is the way the media reports inflation. However, the theory claims that agents are concerned with the annual forecast and so should discount the very noisy monthly data. We accommodate these two assumptions by assuming that agents react both to the most recent observable forecast error and a dampening of past forecast errors. The following was chosen because it produces the most precise results

$$\begin{aligned} MSE_{j,t} &= \left[MSE_{j,t-1}^{\delta=1}, MSE_{j,t-3}^{\delta=0.9} \right] \\ &= \left[\left(\pi_{t-1}^m - \pi_{t-2}^j \right)^2, .1MSE_{j,t-3}^{\delta=0.9} + 0.9 \left(\pi_{t-2}^m - \pi_{t-3}^j \right)^2 \right] \end{aligned}$$

and $\beta = [\beta_1, \beta_2]'$.³³ Performance measures with inertia such as these are considered in Brock and Hommes (1997), Branch and McGough (2003), and Branch and Evans (2003). Under this specification $C_j = [C_j^1, C_j^2]$. To simplify the estimation procedure we assume that $C_j = \beta_1 C_j^1 + \beta_2 C_j^2$. This approach allows for two different δ 's in (7). With such a specification rational expectations still dominates adaptive in forecast error. Here $MSE_{j,t-1}^{\delta=1}$ is the forecast error model j produces based on the most recent monthly inflation observation. $MSE_{j,t-3}^{\delta=0.9}$ is a geometrically weighted average of all past forecast errors. Thus, we allow for agents to adjust their choices to new information but not neglect the past as well.

The dependent variable, $\pi_{i,t}^e$, is the survey response to the question on inflation expectations over the next 12 months. The model was estimated over the period 1977.11–1993.12.³⁴ Note also that the variables $\pi_{i,t}^e$, $MSE_{j,t}$, $H_l(\pi^l)$ are all measured as percentages. The units are chosen to ensure a well-conditioned log-likelihood function.

4.2. Empirical Results

4.2.1. Baseline model 1: dynamic predictor selection with no costs

This subsection focuses on a baseline model which simultaneously estimates predictor proportion and the parameters β_1 and β_2 but ignores costs. We now solve the problem (23) numerically under the assumption that

$$U_{j,t} = -\beta_1 MSE_{j,t-1}^{\delta=1} - \beta_2 MSE_{j,t-3}^{\delta=0.9}.$$

This is equivalent to the hypothesis of zero costs; that is, $H_0 : C_N = C_A = C_V = 0$. The discussion of the full model will present a test of this hypothesis.

³³ All of the qualitative results are robust to alternative choices of δ and lags. This specification was chosen because it provides the best fit of the data and is suggested by theory.

³⁴ An issue in previous empirical models, that is not a concern here, is the occurrence of overlapping forecasts. Because an agent's forecast in $t + 1$ overlaps with another agent's time t forecast, the errors may be serially correlated. In tests of rational expectations this may produce a spurious rejection of the Rational Expectations Hypothesis. This is not relevant here since our hypothesis tests do not concern the presence of serial correlation. Moreover, the estimator's covariance matrix is robust to MA disturbances.

Table 2 gives the results of the maximum likelihood estimation for the model with all three predictors and no cost parameters.³⁵ With $\beta_1 = 0.4594$ and $\beta_2 = 0.0356$ agents respond more strongly to recent forecast errors than to dampened past errors; the effect of the most recent error is over 10 times the size of the effect for the dampened error. This fits into the theoretical notion of bounded rationality introduced by Brock and Hommes (1997): agents rationally choose the appropriate predictor with which to form their expectations, however, they only consider its most recent success. In a stochastic environment, random shocks could make a naive predictor more accurate for an occasional period but, on average, the naive predictor does worse. Agents are myopic and base their decisions on the most recent inflation realisation, so if the naive predictor does well in one period a large number of agents will switch to the naive predictor the next period.

Figures 1–3 plot the estimated predictor proportions over the sample period. As the figures demonstrate there is considerable volatility in the predictor proportions over time. For each predictor there are multiple periods where either all, or no agents used that predictor.

4.2.2. Baseline model 2: static predictor selection

Another possible explanation for observed heterogeneity is it is static. Traditional macroeconomics would argue that there may be heterogeneity in

Table 2

Maximum Likelihood Results for Baseline Model 1 (costs normalised to Zero)

$$U_{j,t} = -\text{Beta}_1 \text{MSE}_{j1}(-1) - \text{Beta}_2 \text{MSE}_{jpt} 9(-3)$$

Beta ₁	Beta ₂	Sigma	Log Likelihood value
0.4594 (1.05e-04)	0.0356 (2.82e-04)	12.4483 (1.87e-05)	-367120

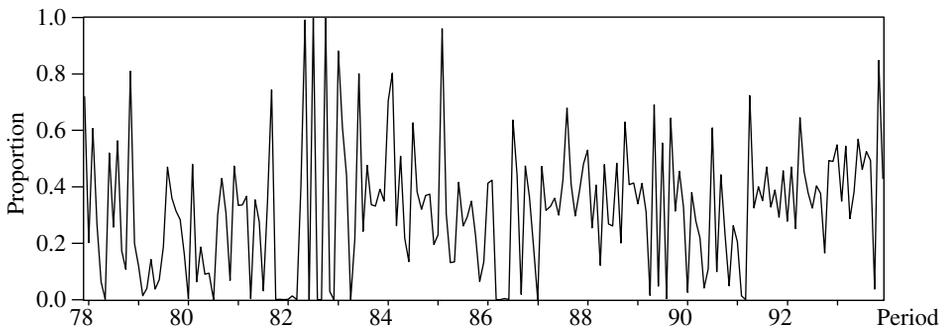


Fig. 1. *Predicted Proportion of VAR in Baseline Model 1*

³⁵ The log-likelihood values in Table 2 may seem large, but for a data set of this size they are a plausible magnitude.

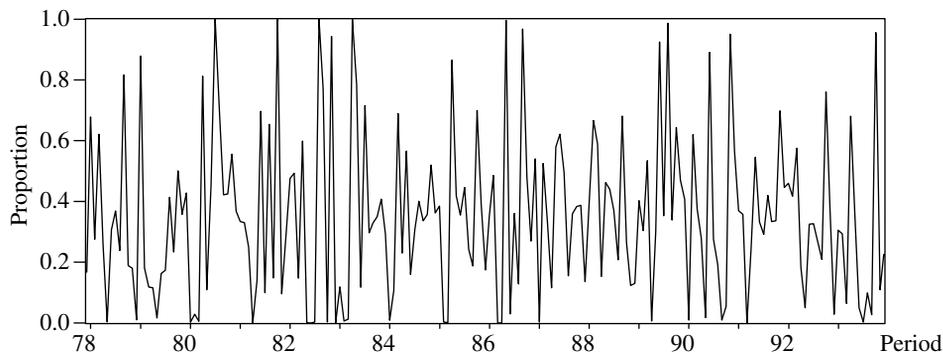


Fig. 2. *Predicted Proportion of Adaptive in Baseline Model 1*

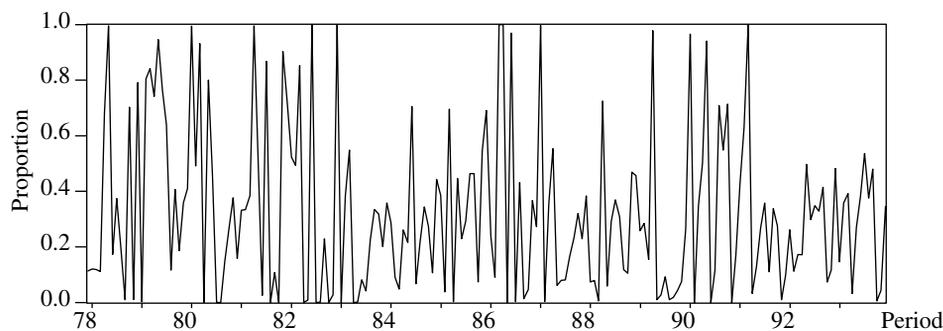


Fig. 3. *Predicted Proportion of Naive in Baseline Model 1*

expectation formation, but this is an innate heterogeneity. Some agents were born as ‘VAR’ agents, while others were born ‘adaptive’. The Baseline Model 2 considers this point of view. We re-estimate the full model with 3 predictors but with costs and no mean squared error. It differs from Baseline Model 1 by allowing for costs and disallowing dynamic effects. We estimate (23) with the utility function:

$$U_{j,t} = -C_j.$$

Results for the maximum likelihood estimation of the Baseline Model 2 are presented in Table 3. The table presents the estimates for the costs C_N , C_A , C_V under three different cost normalisations. In each normalisation one of the predictors was set to zero in order to increase the degrees of freedom. Because of this normalisation the numerical values for the costs vary.³⁶ However, the

³⁶ The properties of the exponential function cause a wide range of values for C to return a similar likelihood value. Even though there are multiple maxima, they have the same qualitative ordering. This occurs in the full model with 3 predictors as well. Correlation coefficients for predicted proportions between these different values for C range from 0.81 to 1, with most having values of approximately 0.999. These large, positive correlations imply that although there are quantitative differences between these estimates they have very similar empirical implications.

Table 3
Maximum Likelihood Estimates for Baseline Model 2 (no dynamic mse effects) $U_t = C_j$

Cost Normalization	C_V	C_A	C_N	Sigma	Log Likelihood value	n_V	n_A	n_N
Rational Expectations	0	0.0731 (1e-08)	1.9064 (4.5e-05)	12.4002 (1e-08)	-366980	0.4812	0.4473	0.0715
Adaptive Expectations	-0.0748 (8.6e-07)	0	1.8337 (7.0e-08)	12.401 (1.00E-090)	-366980	0.4816	0.4469	0.0714
Naïve Expectations	-1.9114 (6.1e-04)	-1.8382 (3.7e-02)	0	12.3994 (7.0e-08)	-366980	0.4814	0.4474	0.0712

qualitative hierarchy remains the same for each normalisation: $C_N > C_A > C_V$. The naive predictor carries a greater cost than the adaptive predictor which costlier than the VAR predictor. Since we are estimating constant predictor proportions in the sample, this result is equivalent to saying that the greatest proportion of agents use the VAR predictor followed by the adaptive and naive, respectively. Table 3 also details precisely what the estimated predictor proportions are for each normalisation. The VAR predictor accounts for about 48% of the sample, adaptive accounts for about 44%, and the naive predictor has a 7% share of the sample. The low fraction of naive as compared to VAR and naive, is consistent with the criticism of naive expectations raised by macroeconomists in the 1960s. In fact, the proportions follow the same patterns as the development of expectation models in dynamic macroeconomics.

A brief note on existing related results is warranted. Baak (1999) and Chavas (2000) present empirical evidence for the presence of heterogeneous price expectations in the US beef market. Similar to the approach of the restricted model in this subsection, they estimate the proportion of 'boundedly rational' agents. Chavas (2000) finds a significant number (81.7%) of linear univariate and naive agents. Baak (1999) finds that about 40% of the agents are boundedly rational. In particular, Baak's results are quantitatively close to those here. One possibility is that each of his forecast functions return similar forecasts as adaptive and VAR do for inflation.

The result that VAR is the least expensive seems paradoxical. Common sense suggests that cost should be ordered based on sophistication: the most sophisticated predictor should be the most expensive. If this were true then our results suggest that the naive predictor is the most sophisticated, and the VAR predictor is the least sophisticated. The adaptive predictor would also be slightly more sophisticated than the VAR. These results suggest that our initial interpretation of cost is not right. In the next Section we will present an alternative interpretation.

4.2.3. Full model with 3 predictors

The model in (23) includes the two baseline models as special cases. The full model considers both the dynamic and static effects. In this sense, it incorporates the traditional expectation formation models and the more recent rationally heterogeneous expectations models.

The maximum likelihood estimates for the full model with 3 predictors are presented in Table 4. As in the Baseline Model 2, the results are segmented by normalisation. The mean squared effects decrease from Baseline Model 1. The estimates for β_1 and β_2 are virtually identical for the cases when VAR and adaptive are normalised. They are significantly more when the naive predictor is normalised. The cost estimates again follow the same structure as in Baseline Model 2. The naive predictor is the most expensive, the adaptive predictor is in the middle, and the VAR predictor is the cheapest.³⁷

³⁷ When a VAR is estimated assuming the parameters follow an AR(1), the results are robust. For example, typical estimates are $\beta_1 = 0.0201$, $\beta_2 = 0.1047$, $C_V = 0$, $C_A = 0.9337$, $C_N = 3.0342$. We conjecture that these qualitative results will be robust to any predictor specifications with the same forecast accuracy ordering.

Table 4
Maximum Likelihood Estimates for the full Model with 3 Predictors (standard errors in parentheses)
 $U_{jt} = -\text{Beta}_1 \text{MSE}_{jt}(-1) - \text{Beta}_2 \text{MSE}_{jpt} 9(-3) - C$

Cost Normalization	Beta_1	Beta_2	C_v	C_A	C_N	Sigma	Log Likelihood value
VAR Expectations	0.0079 (5.86e-02)	0.0137 (5.7e-02)	0	0.2396 (7.74e-02)	2.0483 (1.38e-03)	12.3857 (1.76e-02)	-3666920
Adaptive Expectations	0.0072 (1.0e-08)	0.0135 (2.0e-07)	-0.115 (1.3e-05)	0	1.93 (8.4e-06)	12.4001 (1.7e-05)	-3666920
Naive Expectations	0.0113 (1.024e-07)	0.0158 (3.8537e-12)	-2.5479 (3.562e-07)	-2.3462 (2.273e-07)	0	12.3947 (8.0e-10)	-3666920

Here cost represents a predisposition to use a certain predictor. The Baseline Model 2 was introduced as the traditional macroeconomic argument that agents are born with a given predictor. The theory on rationally heterogeneous expectations argues this is not the case. However, there may be evidence for both arguments. Agents may be predisposed to using a VAR predictor, but depending on the forecast error may find it worthwhile to switch methods. The VAR predictor certainly involves a greater calculation cost. In times where the naive predictor returns as good or better forecasts, even those most likely to use a VAR predictor at all times may choose to switch. A similar argument holds for those more likely to use a naive predictor.

Thus, cost in the model acts as a threshold with which forecast errors must cross in order to induce a switch of methods. We argue this is a reasonable and natural interpretation of expectation formation. In fact, this argument has been advanced by Keane and Runkle (1990) as support for investigating the Survey of Professional Forecasters. They describe a situation where a survey respondent under forecasts by 1% an actual inflation rate of 9%. They argue this person does not trade in the bond market based on the expected lower inflation rates because 'when he thinks about it, 9 percent seems reasonable.' (Keane and Runkle, 1990, p. 715). In our context, the agent will not switch his prediction method because the forecast error is 'reasonable'. Only forecast errors above a threshold will induce changes in predictor choice. Agents are inclined to use one predictor over another, *ceteris paribus*, but when the forecast error for a predictor becomes larger they are more likely to switch to a more accurate forecast method.

This result is in line with the findings of an endowment or status quo effect in the economic psychology literature (Kahneman *et al.*, 1991). In economic experiments it is sometimes found that agents make decisions based on changes from a reference point. Given an endowment, agents are reluctant to make changes. In our interpretation of the predisposition effect agents are inherently biased and switch predictors when past forecast errors exceed this bias. The inherent predisposition then is like the reference point in the economic psychology literature.

It should also be noted that we tested for whether the costs and 'intensities of choice' are statistically significantly different from each other. In particular, even though the VAR and adaptive models return quantitatively similar costs, they are statistically significantly different. They return cost estimates that are close because they return similar forecasts. In Section 3.3 we replace VAR with rational expectations and find that the costs are now an order of magnitude different.

Although the values for β_1 and β_2 differ across normalisations they return essentially the same predicted proportions.³⁸ Note that the log-likelihood values are identical across all normalisations. This is expected since normalisations should not change the nature of the estimates. The properties of the exponential function make it so that small changes in parameter values, in a neighborhood of the maximum, return even smaller changes in likelihood value. It turns out that

³⁸ Despite the existence of local maxima, all estimates return roughly the same predicted proportions.

the best fit for the naive normalisation has larger values for β_1 and β_2 than in the other two cases. This difference in values does not affect the role mean square error plays in the estimated proportions. Table 5 lists correlation coefficients for predicted proportions across normalisation. The first line gives the correlation between predicted proportions of VAR for the VAR and adaptive normalisations. These correlation coefficients are all close to or equal to 1, indicating that the estimated proportions are the same for each normalisation. Thus, a value of β_1 of 0.0079 has the same implications as a value of 0.0113 depending on the normalisation.

Figures 4–6 plot the predicted proportions of VAR, adaptive, and naive predictors over the sample for the estimated parameter values. There is considerably less volatility in the proportions in this case than in Baseline Model 1. The proportions vary around the means that were seen in Baseline Model 2. This illustrates the threshold or predisposition effect discussed earlier. The proportions vary inversely of each other since as the mean square error for one increases relative to the other predictors, then the proportion for that one will decrease and increase for the others. Table 6 shows descriptive statistics for the estimated predictors across normalisations. These results are similar to what was seen in Baseline Model 2 where the largest fraction of agents use the VAR predictor, a slightly smaller proportion use the adaptive predictor, and a much smaller proportion use the naive predictor.

Table 5
Correlation Coefficients for Predicted Proportions across Normalisations

Cost Normalisations	VAR	Adaptive	Naïve
VAR-Adaptive	0.9974	0.997	0.9953
VAR-Naïve	0.9974	0.997	0.9953
Adaptive-Naïve	0.9987	0.9968	0.9939

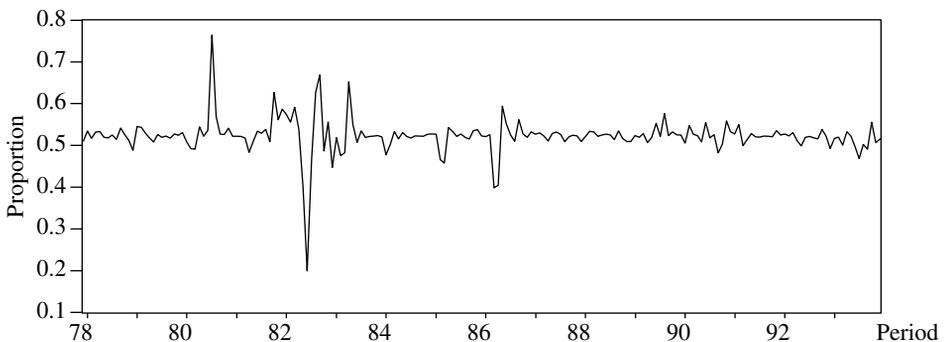


Fig. 4. *Predicted Proportion of VAR under VAR Normalisation in Full Model with 3 Predictors.*

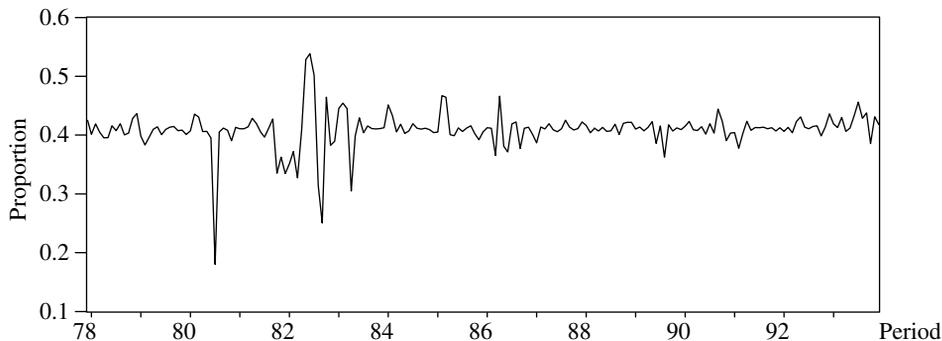


Fig. 5. *Proportion of Adaptive under VAR Normalisation in Full Model with 3 Predictors.*

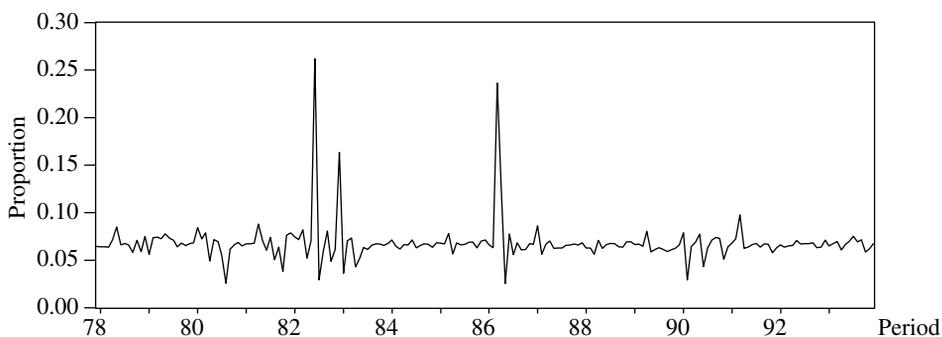


Fig. 6. *Proportion of Naive under VAR Normalisation in Full Model with 3 Predictors.*

The full model with 3 predictors also allows us to test the two baseline models, and whether either of the two competing theories are superior. The theory of rationally heterogeneous expectations, or Baseline Model 1, occurs under the hypothesis that the C_j s are jointly equal to zero. Table 6 also presents results of the likelihood ratio tests for this hypothesis. In each normalisation a χ^2 test statistic of 200 was calculated, which clearly rejects the hypothesis that costs are jointly equal to zero. Thus, the purely dynamic model is not superior to a model that includes both dynamic and static effects. Baseline Model 2, or the traditional static model, is equivalent to the hypothesis that the parameters on the MSEs are jointly equal to zero. Again, the χ^2 test statistic of 60 rejects this hypothesis.³⁹ These results indicate that the model which blends both of these approaches is superior to a purely dynamic or static approach.⁴⁰ That the β 's are positive and statistically significant provides support for the ARED model.

³⁹ In particular, we are imposing 2 restrictions. After correcting for degrees of freedom the critical statistic at a 5% significance level is 5.99.

⁴⁰ In an earlier version we also tested a restricted model with only VAR and adaptive expectations. Under the hypothesis $H_0 : C_N = \infty$ we reject that the proportion of naive agents is zero. We also rejected tests of homogeneous expectations. The rejection of these restricted models provides support for the full model with costs.

Some might argue these results cast doubt on the ARED model. Brock and Hommes' (1997) model emphasises the comparison of the relative performance of predictors and assuming a 'finite' intensity of choice is a simple way to incorporate inertia. That we find a finite 'intensity' of choice and inertia stemming from the costs might lead some to conclude that the results here are not consistent with the ARED. However, it is our contention that our results support central elements of the ARED. We do find that the share of agents' who use a predictor evolves according to past performance. We also find that agents are somewhat slow to react. However, we also find that this inertia may be a result of their 'inattentiveness' as Brock and Hommes emphasised, or because of a predisposition towards a particular predictor. These findings uphold the ARED model but also indicates that it supports a rich variety of behaviour.

It would be fruitful for future research to examine a 'richer' dynamical structure than the model at hand. One possible accounting for the inertia finding is that a predictor choice in period t is not independent from the choice in $t - 1$. The serial correlation could be the result of time-varying random costs or random utility shocks. Such analysis, however, is beyond the scope of this paper. The inertial finding in the cost estimates then could be interpreted as adjusting to bring the misspecified model (MNL) closer to the 'true' model with serial dependence.

4.2.4. *Alternative models*

This Section considers alternative hypotheses to the rationally heterogeneous expectations model. The contention of this paper is that this model of dynamic predictor selection will outperform alternative models. This subsection shows that, in terms of log-likelihood, dynamic predictor selection fits the Survey of Consumers data better than the alternatives. We will list results below of comparisons between this model and two simple alternatives. We conclude from these tests of 'goodness of fit' that the ARED model is supported by the data.

The simplest alternative model is where all agents are randomly distributed around the mean forecast. In this alternative we will assume that inflation expectations are normally distributed with mean and variance equal to the sample mean and variance. Next we will consider where the means and variances are time-varying. The sample mean across all time periods and all agents is $\bar{x} = 6.9550$ and a standard deviation of $\sigma_S = 12.7010$. This approach makes no pretenses about how it is that agents form their expectations. It just assumes that over the population these heterogeneous expectations form a normal distribution.

The density function for any observation in the sample is given by

$$f(\pi_{i,t}^e) = \frac{1}{\sqrt{2\pi}\sigma_S} \exp\left[-(1/2)\left(\frac{\pi_{i,t}^e - \bar{x}}{\sigma_S}\right)^2\right]. \quad (24)$$

Table 6

Descriptive Statistics and Tests for Proportion Predictions from the Full Model with 3 Predictors

Cost Normalisations	Mean	Median	Std. Deviation
VAR Expectations			
VAR	0.5228	0.5227	0.0427
adaptive	0.4088	4.11E-01	0.0332
naïve	0.0683	6.66E-02	0.0225
Adaptive Expectations			
VAR	0.496	0.4954	0.0407
adaptive	0.4392	4.42E-01	0.0329
naïve	0.0649	0.0634	0.02
Naïve Expectations			
VAR	0.5285	0.5283	0.0513
adaptive	0.4288	4.31E-01	0.0426
naïve	0.0427	4.07E-02	0.0212
Hypothesis Tests			
	$H_0 = u_{jt} = u_j$	$H_0 = C_v = C_A = C_N = 0$	
	$\chi^2 = 60$	$\chi^2 = 200$	
	P < 0.0001	P < 0.0001	

Table 7

Maximum Likelihood Estimates and Tests for the Full Model with Rational, Adaptive, and Naïve Expectations $U_{j,t} = -\text{Beta}_1 \text{MSE}_{j1}(-1) \text{Beta}_2 \text{MSE}_{jpt} 9(-3) - C_j$

Cost Normalisation	Beta ₁	Beta ₂	C _{RE}	C _A	C _N	Sigma _v	Sigma _e	Log Likelihood value
Naïve Expectations	0.0057	0.0071	-3.6255	-5.793	0	3.8594	35.4123	-301820
Mean Predicted Proportions:								
Rational	0.1121							
Adaptive	0.8851							
Naïve	0.0027							
Random Model 1:								-725910
Random Model 2:								-697940

Then the likelihood function is

$$\mathcal{L} = \prod_t \prod_i f(\pi_{i,t}^e). \tag{25}$$

Table 7 reports the likelihood value for this alternative model. With a log-likelihood value of -725910 the full-model with 3 predictors would be chosen over this alternative model in a maximum likelihood procedure.

A drawback to this simple alternative is that inflation is time-varying and our alternative should include time-varying means. In the second alternative approach, we examine where agents are again normally distributed around the sample mean forecast, but now that sample mean is taken across individuals and not time. We examine the log-likelihood value for a model where agents in each time period are distributed normally around the mean and variance of that time period. Now the likelihood function becomes

$$\mathcal{L} = \prod_t \prod_i \frac{1}{\sqrt{2\pi}\sigma_{S,t}} \exp \left[(1/2) \left(\frac{\pi_{i,t}^e - \bar{x}_t}{\sigma_{S,t}} \right)^2 \right]. \quad (26)$$

Table 7 list the log-likelihood value for this second alternative model. Here the log-likelihood has increased to 697940, which is still significantly less than the log-likelihood value under the 3-predictor model. This indicates that it is, in fact, the dynamic predictor selection mechanism that yields a better fit.

4.3. *The Full Model with Rational Expectations*

One criticism of our approach so far is that the VAR predictor is not sufficiently sophisticated and that we are excluding rational expectations as an alternative. Heretofore we have estimated the VAR model as an approximation to all multivariate models in order to keep all predictors consistent with the same model of bounded rationality: agents behave like economic forecasters by using an array of techniques to map past monthly inflation into annual forecasts. However, it is possible to replace the VAR predictor with rational expectations. We briefly comment on this possibility and the consistency of its results with prior analysis. The main drawback to including rational expectations is that there is no mechanism specified by which people form these expectations. This is why we focused on a VAR model instead.

Under rational expectations the difference between actual inflation and expected future inflation (i.e. $\pi_{t+1} - E_t\pi_{t+1}$) is an unforecastable error term. The standard test for rational expectations is whether these errors are correlated with other explanatory variables. Usually such tests reject the null hypothesis of rational expectations. By replacing the VAR predictor with rational expectations in the full-model we consider the possibility that a subset of the population is rational. We assume that agents choice set is now $\{\pi_{t-1}, AE_t, \hat{\pi}_{t+1}\}$ and $\hat{\pi}_{t+1} = \pi_{t+1} - \hat{\epsilon}_{t+1}$ where $\hat{\epsilon}_{t+1}$ is distributed *iid* normal. Note in particular that $\{\hat{\epsilon}_t\}$ is a martingale difference sequence and $\hat{\pi}_t$ are rational expectations. The likelihood function in this case is

$$\mathcal{L} = \ln P \left[\pi_{i,t}^e, i = 1, \dots, N, t = 1, \dots, T \mid MSE^t, \mathcal{H}_t(\pi^t), t = 1, \dots, T \right] \quad (27)$$

$$= \sum_t \sum_i \ln \left[\frac{\exp\{\beta[-(MSE_{j,t} + C_j)]\}}{\sum_{k \in \{N,A,RE\}} \exp\{\beta[-(MSE_k + C_k)]\}} \right] \quad (28)$$

$$\begin{aligned} & \times \frac{1}{\sqrt{2\pi}\sigma_v} \exp \left\{ \frac{-1}{2} \left[\frac{\pi_{i,t}^e - H_t(\pi^t)}{\sigma_v} \right]^2 \right\} \\ & + \frac{\exp\{\beta[-(MSE_{RE,t} + C_{RE})]\}}{\sum_{k \in \{N,A,RE\}} \exp\{\beta[-(MSE_k + C_k)]\}} \\ & \times \frac{1}{\sqrt{2\pi}\sigma_{\hat{\epsilon}}} \exp \left\{ \frac{-1}{2} \left(\frac{\pi_{i,t}^e - \hat{\pi}_{t+1}}{\sigma_{\hat{\epsilon}}} \right)^2 \right\}. \end{aligned} \quad (29)$$

The empirical problem is to find β , C_j , σ_V , σ_ε which maximises (27). Table 7 summarises the results. For brevity we report only the naive-normalisation results. As before the results extend to the other normalisations as well. Note that the cost to adaptive is still significantly less than naive. Now the rational predictor is more expensive than adaptive but less than naive. This result is consistent with earlier results because of the close connection between VAR and adaptive. A significant fraction of agents tend to be 'sophisticated'. Computing actual rational expectations, though, is too costly and only a small percentage will be rational. This means that the vast majority of agents are actually adaptive. Table 7 also details the mean proportions of agents assigned to each predictor. Naive agents continue to be the least used predictor as it is quite myopic. Adaptive expectations is a good forecaster of inflation and perfect foresight plus noise is too costly. The excessively high estimate for the standard error of the uncorrelated error term in the rational predictor also suggests that this specification is unrealistic. The lower standard error for the other white noise shock indicates that much of the volatility in survey responses comes from the 'forward-looking' agents. We have included these results in order to demonstrate the flexibility of our approach.

5. Conclusion

This paper investigated the empirics of dynamic predictor selection as a test of the rationally heterogeneous expectations model. In this setting, agents choose between VAR, adaptive, and naive expectations when forming their beliefs about future inflation. The probability that an agent will choose a predictor depends on its relative mean squared error and also a fixed cost to use the predictor. Survey data from the University of Michigan's Survey Research Center on inflationary expectations were analysed for evidence of this type of dynamic expectation formation. Using maximum likelihood to analyse a new discrete choice setting where the latent variable is discrete and the observed variable is continuous, we find evidence that agents switch predictor use as the relative mean squared errors change.

We find that agents' predictor choices respond negatively to increases in relative mean square error. If the mean square error of the VAR predictor increases *vis-à-vis* the naive predictor in time t , then the probability that predictor will be chosen in time $t + 1$ decreases. We provide evidence that agents also have predispositions towards one predictor over another. Evidence indicates that belief formation is not purely a rational choice between competing alternatives, but there is also some inherent bias towards one predictor over another. We also show the data are not consistent with a model of static heterogeneous expectations. The model with static and dynamic effects has a higher likelihood value than the traditional heterogeneous expectations model. Finally, our results suggest, on average, more agents use VAR expectations than naive or adaptive expectations. In comparison with alternative models, the theory of rationally heterogeneous expectations fits the data best.

These empirical results suggest the following story of expectation formation. Each individual has an inherent bias towards using one predictor over another.

Perhaps this tendency is a genetic quality, or a product of learning. Individuals are not wedded to this predictor, however. Consistent with the theme of the rational expectations hypothesis, a predictor that does poorly will be used with a lower probability. Agents look back at how the predictor did recently, and to a lesser degree over the entire history, and make a boundedly rational decision on whether to use that predictor again or to switch. In periods of low inflation volatility there is little mean square error advantage to using a VAR predictor over a naive predictor and there is little switching from the naive predictor. Common sense suggests that in such times there will be switching from VAR to naive. We find, though, that because of these predispositions agents will not switch from VAR to naive. The past forecast errors must cross a threshold before agents will switch their methodology.

This paper indicates that the data are supportive of the model of rationally heterogeneous expectations, but not completely so. The theoretical model can not account for bias towards a predictor. These results indicate that a theory which extends the previous models to account for these predispositions is called for.

The theoretical model is also one of a discrete choice. Agents choose from a finite set of alternatives. The data used in this paper are continuous, and so a new approach was necessary for empirical analysis. In future research, we will extend the model to where there is individual-specific behaviour that adapts over time and examine the implications with a panel data set. We intend to conduct this analysis in a framework where costs and/or random utility are serially correlated.

University of California, Irvine

Date of receipt of first submission: October 2002

Date of receipt of final typescript: October 2003

Dataset is available for this paper: www.res.org.uk

References

- Baak, S.J. (1999). 'Tests for bounded rationality with a linear dynamic model distorted by heterogeneous expectations', *Journal of Economic Dynamics and Control*, vol. 23, pp. 1517–43.
- Ball, L. (2000). 'Near-rationality and inflation in two monetary regimes', NBER Working Paper 7988.
- Branch, W.A. (2002). 'Local convergence properties of a cobweb model with rationally heterogeneous expectations', *Journal of Economic Dynamics and Control*, vol. 27(1), pp. 63–85.
- Branch, W.A. and Evans, G.W. (2003). 'Intrinsic heterogeneity in expectation formation', mimeo.
- Branch, W.A. and McGough, B. (2003). 'Replicator dynamics in a cobweb model with rationally heterogeneous expectations', mimeo.
- Brock, W.A. and de Fontnouvelle, P. (2000). 'Expectational diversity in monetary economics', *Journal of Economic Dynamics and Control*, vol. 24, pp. 725–759.
- Brock, W.A. and Durlauf, S.N. (2001). 'Discrete choice with social interactions', *Review of Economic Studies*, vol. 68(2), pp. 235–60.
- Brock, W.A. and Hommes, C.H. (1997). 'A Rational Route to Randomness', *Econometrica*, vol. 65, pp. 1059–1160.
- Brock, W.A. and Hommes, C.H. (1998). 'Heterogeneous beliefs and routes to chaos in a simple asset pricing model', *Journal of Economic Dynamics and Control*, vol. 22, pp. 1235–74.
- Brock, W.A. and Hommes, C.H. (2000). 'Rational animal spirits', in (P. J. J. Herings, A. J. J. Talman and G. Van der Laan, eds), *The Theory of Markets*, pp. 109–37. Amsterdam: North-Holland.
- Brock, W.A., Hommes, C.H. and Wagener, F. O. O. (2001). 'Evolutionary dynamics in financial markets with many trader types', Cendef Working paper 01-01, University of Amsterdam.

- Bryan, M.F. and Venkatu, G. (2001a). 'The demographics of inflation opinion surveys', *Economic Commentary*, Federal Reserve Bank of Cleveland, October.
- Bryan, M.F. and Venkatu, G. (2001b). 'The curiously different inflation perspectives of men and women', *Economic Commentary*, Federal Reserve Bank of Cleveland, November.
- Cagan, P. (1956). 'The monetary dynamics of hyper-inflation', in (M. Friedman, ed.) *Studies in the Quantity Theory of Money*, Chicago: University of Chicago Press.
- Carroll, C.D. (2003). 'The epidemiology of macroeconomic expectations', *Quarterly Journal of Economics*, vol. 118(1), pp. 269–98.
- Caskey, J. (1985). 'Modeling the formation of price expectations: a bayesian approach', *American Economic Review*, vol. 75(4), pp. 768–76.
- Chavas, J.-P. (2000). 'On information and market dynamics: the case of the U. S. beef market', *Journal of Economic Dynamics and Control*, vol. 24, pp. 833–53.
- Cogley, T.W. and Sargent, T.J. (2003). 'Drifts and volatilities: monetary policies and outcomes in the postwar US', mimeo.
- Croushore, D. (1993). 'Introducing: the survey of professional forecasters', *Business Review*, Federal Reserve Bank of Philadelphia, November/December.
- Croushore, D. (1997). 'The livingston survey: still useful after all these years', *Business Review*, Federal Reserve Bank of Philadelphia, March/April, pp. 1–12.
- Evans, G.W. and Gulamani, R. (1984). 'Tests for rationality of the Carlson-Parkin inflation expectations data', *Oxford Bulletin of Economics and Statistics*, vol. 46(1), pp. 1–20.
- Evans, G.W. and Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*, Princeton University Press.
- Evans, G.W. and Ramey, G. (1992). 'Expectation calculation and macroeconomic dynamics', *American Economic Review*, vol. 82, pp. 207–24.
- Evans, G.W. and Ramey, G. (1998). 'Calculation, adaptation and rational expectations', *Macroeconomic Dynamics*, vol. 2, pp. 156–82.
- Fishe, R.P.H. and Idson, T.L. (1990). 'Information-induced heteroscedasticity in price expectations data', *Review of Economics and Statistics*, vol. 72(2), pp. 304–12.
- Frankel, J.A. and Froot, K.A. (1987). 'Using survey data to test standard propositions regarding exchange rate expectations', *American Economic Review*, vol. 77(1), pp. 133–53.
- Friedman, M. (1957). *Theory of the Consumption Function*, Princeton: Princeton University Press.
- Goeree, J.K. and Hommes, C.H. (2000). 'Heterogeneous beliefs and the non-linear cobweb model', *Journal of Economic Dynamics and Control*, vol. 24, pp. 761–98.
- Hommes, C.H. (2001). 'Financial markets as nonlinear adaptive evolutionary systems', *Quantitative Finance*, vol. 1(1), pp. 149–67.
- Jeong, J. and Maddala, G.S. (1996). 'Testing the rationality of survey data using the weighted double-bootstrapped method of moments', *Review of Economics and Statistics*, vol. 78(2), pp. 296–302.
- Kahneman, D., Knetsch, J.L. and Thaler, R.H. (1991). 'The endowment effect, loss aversion and status quo bias', *Journal of Economic Perspectives*, vol. 5(1), pp. 193–206.
- Keane, M.P. and Runkle, D.E. (1990). 'Testing the rationality of price forecasts: new evidence from panel data', *American Economic Review*, vol. 80(4), pp. 714–735.
- Litterman, R.B. (1986). 'Forecasting with bayesian vector autoregressions – five years of experience', *Journal of Business and Economic Statistics*, vol. 4, pp. 25–38.
- Manski, C.F. and McFadden, D. (1981). *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge, MA: MIT Press.
- McFadden, D. (1984). 'Econometric analysis of qualitative response models', in (Z. Griliches and M. D. Intriligator, eds.), *Handbook of Econometrics*, Vol. II, Amsterdam: Elsevier.
- Nerlove, M. (1958). *The Dynamics of Supply: Estimation of the Farmers' Response to Price*, Baltimore: Johns Hopkins University Press.
- Pesaran, M.H. and Timmermann, A. (1995). 'Predictability of stock returns: robustness and economic significance', *Journal of Finance*, vol. 50(4), pp. 1201–28.
- Sargent, T.J. (1993). *Bounded Rationality in Macroeconomics*, Oxford: Oxford University Press.
- Sargent, T.J. (1999). *The Conquest of American Inflation*, Princeton: Princeton University Press.
- Sethi, R. and Franke, R. (1995). 'Behavioural heterogeneity under evolutionary pressure: macroeconomic implications of costly optimisation', *ECONOMIC JOURNAL*, vol. 105, pp. 583–600.
- Souleles, N.S. (2002). 'Consumer sentiment: its rationality and usefulness in forecasting expenditure? Evidence from the michigan micro data'. forthcoming *Journal of Money, Credit, and Banking*.
- Struth, F.K. (1984). 'Modelling expectations formation With parameter-adaptive filters: an empirical application to the livingston forecasts', *Oxford Bulletin of Economics and Statistics*, vol. 46(3), pp. 211–39.
- Urich, T. and Wachtel, P. (1984). 'The structure of expectations of the weekly money supply announcement', *Journal of Monetary Economics*, vol. 13, pp. 183–94.