Forecast Combination, Non-linear Dynamics, and the Macroeconomy

Christopher G. Gibbs*
UNSW Australia

January 5, 2016

Abstract

This paper introduces the concept of a Forecast Combination Equilibrium to model boundedly rational agents who combine a menu of different forecasts in a way that mimics the behavior of actual forecasters. The equilibrium concept is consistent with rational expectations under certain conditions, while also permitting multiple, distinct, self-fulfilling equilibria, many of which are stable under least squares learning. The equilibrium concept is applied to a Lucas-type monetary model and to a Fisherian monetary model with a Taylor rule. The existence of multiple equilibria is shown to depend on the aggressiveness of monetary policy in both models. In the latter, a more aggressive response to inflation is required in the Taylor rule than is typically found in this class of model to ensure a unique and learnable equilibrium. Real-time learning simulations with a constant gain illustrate some appealing properties of this approach including time-varying volatility and sharp movements in inflation, similar to actual data, while assuming only i.i.d. random shocks.

JEL Classifications: E31, C52, C53, D83

Keywords: Forecast Combination, Adaptive Learning, Expectations, Dynamic Predictor Selection, Inflation, Forecast Combination Puzzle.

*Contact email: christopher.gibbs@unsw.edu.au. This paper is a distillation of the key results and an extension to the third chapter of my dissertation: "Forecast Combination in the Macroeconomy." I wish to thank George Evans, Bruce McGough, Jeremy Piger, James Morley, Mikhail Anufriev, Valentyn Panchenko, the University of Oregon Macro Group; participants of the 2013 Expectations in Dynamic Macroeconomic Models Conference hosted by the Federal Reserve Bank of San Francisco; seminar participants at Washington State University, Bank of Canada, U.S. Naval Academy, Trinity College, Franklin and Marshall College; and multiple anonymous referees for valuable insights and comments on previous versions of this paper. All remaining errors are my own.

1 Introduction

A key feature of modern macroeconomic models is the forward-looking agent who makes current period decisions based on expectations of the future. The standard modeling assumption for agents' expectations is of course rational expectations (RE), which assumes that agents understand the true structure of the model and use this knowledge to form model consistent expectations. However, the actual experience of econometricians, policymakers, firms, and consumers of forming expectations involves significant model uncertainty, where there exists many suitable models to forecast any variable of interest.

The model uncertainty problem is evidenced in professional forecasting and policy making by the common use of combined forecasts that aggregate many different forecasts together. Prominent examples of these types of forecasts include the Federal Reserve's Greenbook consensus forecasts, the Survey of Professional Forecasters, or the Blue Chip Economic Indicators consensus forecasts.[1] The model uncertainty problem is also observed at the individual level in laboratory experiments where people are asked to forecast in a controlled environment. For example, Anufriev and Hommes (2012a,b) show that experimental data on forecasting asset prices in a laboratory is explained by participants coordinating on a set of heuristic forecasting rules and then switching among those rules over time.[2]

Model uncertainty is also widely studied in the forecasting literature where two types of solutions are typically proposed: either a forecaster adopts a fitness criterion to distinguish and select among forecasts or a strategy is used to aggregate and combine forecasts. The forecast combination solution is often found to be the most effective. Forecast combination captures important information from many different forecasts while lowering the risk of choosing the worst forecast.[3] This paper proposes a framework to model boundedly rational agents who face model uncertainty and employ forecast combination strategies, instead of model selection, to mimic the forecasting behavior of actual professional forecasters.

Despite the dominance of forecast combination in the forecasting literature, theoretical models that have studied agents with model uncertainty overwhelmingly model agents

[1]See Robertson (2000) for a detailed description of common central bank forecasting practices. Wieland and Wolters (2011) also provides empirical evidence for significant diversity in professional forecasts around NBER recession dates.

[2]Surveys of the experimental Learning-to-Forecast literature are found in Hommes (2011) and Hommes (2013).

[3]The seminal paper demonstrating the efficacy of forecast combination is Bates and Granger (1969), who showed that weighted averages of competing forecasting methods consistently outperforms any of the individual forecasts considered. Surveys of the literature are found in Clemen (1989), Timmermann (2006), and Wallis (2011).

that select, rather than combine forecasts. Some prominent examples are Brock and Hommes (1997, 1998), Chiarella and He (2003), Branch and Evans (2006, 2007, 2011a), Branch and McGough (2008, 2010), Brock et al. (2009), De Grauwe (2011), and Anufriev et al. (2013). The agents in these models select forecast rules by a process called Dynamic Predictor Selection, where agents use a fitness measure to distinguish and select among the rules.

One explanation for the preference in the literature for model selection, rather than combination, is that from an aggregative perspective the two strategies appear to be the same. The aggregate expectation in both cases is a linear combination of the menu of forecast rules. Therefore, it would be natural to conclude that there is nothing to gain by explicitly exploring forecast combination. However, Evans et al. (2013) show that agents who engage in Bayesian model averaging over small deviations of the RE forecast can in simulation converge to a non-rational self-confirming equilibrium. This suggests that forecast combination may result in distinct and useful model predictions beyond those obtained in Dynamic Predictor Selection. In this paper, I propose a new equilibrium concept based on forecast combination to explore this possibility and compare model predictions to the outcomes obtained under Dynamic Predictor Selection.

1.1 Equilibrium Concept and Findings

I introduce the concept of a *Forecast Combination Equilibrium* (FCE), which posits that agents face model uncertainty and consider a menu of different forecast rules. Expectations are formed as a linear combination of the forecasts generated by the rules. The concept is an extension of the Restricted Perception Equilibrium concept used to study dynamic optimizing agents that possess limited information as in Sargent (2001), Evans and Honkapohja (2001), Branch (2004), and McGough (2006).

The menu of forecast rules the agents consider consists of different underparameterizations of the true data generating process. The use of underparameterized forecast rules mimics the standard practice in the forecasting literature of using parsimonious forecasting models. This is because parsimonious models avoid data overfitting that can lead to substantial losses in out-of-sample forecasting accuracy.[4] The use of a menu of parsimonious forecast rules is also the standard approach used in Dynamic Predictor Selection.

The equilibrium concept is developed in a general reduced form macroeconomic model. Combined forecasts are constructed by taking a weighted average of forecast rules with

[4]Empirical examples of the efficacy of using parsimonious forecasts from the forecasting literature are Atkeson and Ohanian (2001), Stock and Watson (2004), and Ang et al. (2007).

weights chosen to minimize the expected squared forecast error of the combined forecast. In the base case explored, there exists a *Fundamental FCE* that is observationally equivalent to the RE solution. However, this FCE is not unique. There also exist multiple self-confirming equilibria, which arise due to positive feedback of expectations and agents' optimally chosen weights. The intuition for the self-confirming equilibria is that agents suffer from a form of the Lucas Critique. The combination weights chosen by the agents are partly based on correlations that are not fundamental, but an artifact of the forecasting strategy. Positive feedback of expectations in the economy reinforces these beliefs making them self-sustaining.

The equilibrium concept is applied in two models: a Lucas-type monetary model and a frictionless dynamic stochastic general equilibrium model with monetary policy following a Taylor rule. The latter model is referred to as a Fisherian model because inflation dynamics are governed by the Fisher equation and monetary policy. Least squares learning of the individual forecast rule parameters and combination weights is used as an equilibrium selection mechanism. The analysis reveals that there exist multiple, learnable equilibria with the number and stability of equilibria dependent on monetary policy. In particular, the response to inflation in the Taylor rule must be larger than what is typically required for determinacy under rational expectations or expectational stability for a range of learning algorithms. Bullard and Mitra (2002) and Preston (2005) argue that expectational stability of the equilibrium outcomes induced by policy rules are an important consideration for policymakers. Expectational stability is typically more restrictive than determinacy under rational expectations and provides a criterion by which to judge a policymaker's ability to coordinate market actions when market participants make reasonable deviations from rationality.

Finally, the real-time learning dynamics of forecast combination are studied under constant gain learning and compared to the dynamics generated under Dynamic Predictor Selection (DPS). Branch and Evans (2007) show that multiple equilibria exist under DPS and that with constant gain learning inflation dynamics in a simple monetary model exhibit endogenous time-varying volatility consistent with observed volatility in inflation data. Endogenous weight forecast combination replicates this result while also generating sharp endogenous breaks to trend inflation. The endogenous breaks to trend inflation resemble bouts of fast rising inflation or deflation similar to the escape dynamics described by Cho et al. (2002) and McGough (2006). The primary source of the endogenous movements of inflation in the forecast combination case is the existence of unstable equilibria in close proximity to the stable equilibria of the model. These two sources of non-linear dynamics are sufficient to replicate some key properties of U.S. inflation in a model that is subject only to i.i.d. shocks.

The remainder of the paper proceeds as follows. Section 2 develops the equilibrium concept in a general reduced form macro model. Section 3 applies the equilibrium concept in the Lucas-type monetary model and the Fisherian model, demonstrates the existence of multiple equilibria, and studies the stability properties of the equilibria under least squares learning. Section 4 compares the dynamics generated by forecast combination in the Lucas-type monetary model to DPS and to actual inflation. Section 5 concludes.

2 A General Framework

In this section I develop the equilibrium concept in a reduced form macroeconomic model that has a unique Rational Expectations Equilibrium (REE). I explore issues of existence and uniqueness and describe conditions under which the equilibrium concept can nest the REE of the underlying model.

2.1 The Reduced Form Economy

Following Evans et al. (2013), I consider a reduced form economy described by a self-referential stochastic process driven by a vector of exogenous shocks. The model takes the following form,

$$y_t = \mu + \alpha E_{t-1} y_t + \zeta' x_{t-1} + w_t, \tag{1}$$

where y_t is a scalar, x_{t-1} is a $n \times 1$ vector of exogenous and observable shocks that follows a stationary process with zero mean, and w_t is white noise. The model is a reduced form version of two well-known macroeconomic models depending on the value of α. For $\alpha < 0$ the model is the reduced form of the Cobweb model of Muth (1961), while for $0 < \alpha < 1$ it is the reduced form of a Lucas-type aggregate supply model of Lucas (1973). The reduced form model has a unique Rational Expectations Equilibrium (REE), which can be written as

$$y_t = (1-\alpha)^{-1} \Omega' z_{t-1} + w_t, \tag{2}$$

where $\Omega = (\mu,\ \zeta')'$ and $z_{t-1} = (1,\ x'_{t-1})'$.[5]

[5]The timing of expectations in this model differs from most DSGE and asset pricing models. The timing in these models is usually $E_t y_{t+1}$. This case is less analytically tractable so I proceed by analyzing the $E_{t-1} y_t$ case as done in Evans et al. (2013) to develop the general concept. The Fisherian model explored in Section 3 provides an explicit example of the equilibrium concept under the alternative

2.2 Misspecified Models

Instead of rational expectations, the agents in the economy are assumed to have uncertainty over the correct specification of the data generating process for y_t. The agents consider k different underparameterized laws of motion for the economy that each omit one or more of the exogenous observables in z_{t-1}. For the general case, it is assumed that the agents' information set contains all of z_{t-1}.[6] Let the $k \geq 2$ possible underparameterized laws of motion or forecast rules be denoted as

$$\hat{y}_{i,t} = \phi_i' u_i z_{t-1} \tag{3}$$

for $i = 1, 2, ...k$, where u_i is an $m_i \times n + 1$ selector matrix that picks out the included exogenous variables for each rule, and ϕ_i is an $m_i \times 1$ vector of parameter beliefs.[7]

I adopt the term "parameter beliefs" for the ϕ_i's to denote that they represent the agents' perception of how $u_i z_{t-1}$ relates to y_t and the terms "expectation" or "forecast" to denote the implied value of y_t given by $\hat{y}_{i,t}$. Agents treat each considered forecast rule as if it described the true data generating process and form parameter beliefs $\{\phi_i\}_{i=1}^{k}$ as optimal linear projections of y_t on $u_i z_{t-1}$. This implies that each ϕ_i satisfies the following orthogonality condition:

$$E u_i z_{t-1}(y_t - \phi_i' u_i z_{t-1}) = 0. \tag{4}$$

The consideration of a list of misspecified models is a familiar setup in the DPS literature. The standard way to proceed at this point is to assume that agents choose a fitness criteria, such as past mean squared forecast error, and select a single rule to make a forecast. I deviate from this structure and instead assume that the agents follow standard practices in the forecast combination literature to create a single prediction by employing a weighted sum of all the individual forecast rules:

$$E_{t-1}^* y_t = \sum_{i=1}^{k} \gamma_i \phi_i' u_i z_{t-1}, \tag{5}$$

expectation timing assumption and illustrates that the key conclusions are unchanged. Select definitions for the $E_t y_{t+1}$ case are provided in the appendix.

[6] This assumption is not crucial for most of the analysis and the agents' information set could contain only a portion of z_{t-1}.

[7] The m_i's correspond to the number of parameters included in each forecast rule so that $m_i \leq n$ for all i. The $m_i \times n+1$ selector matrices are $n+1 \times n+1$ identity matrices with rows that do not correspond to included exogenous variables deleted. An example is given in the appendix. Similar notation is used in Branch and Evans (2006).

where $\gamma_i \in \mathbb{R}$ is the weight given to i^{th} forecast rule and E^*_{t-1} denotes a combined expectation. The model is then closed by specifying how agents choose weights.

2.3 Combination Weights

I consider two ways to choose weights. The first is to assume exogenous weights that do not change with the evolution of the economy. This case is a catchall that nests all possible forecast combination strategies, since in equilibrium any endogenously chosen weights are fixed. The case also nests one of the most common and effective ways to create a combined forecast, which is to take a simple average of the forecasts. The simple average acts as a hedge against model uncertainty by remaining agnostic about the best forecast rule. The strategy is often a key part of any discussion of forecast combination because it is routinely found to outperform more theoretically justified combination strategies. This finding is often referred to as the forecast combination puzzle.[8]

The principle case I consider is an endogenous weight case where agents pick weights to minimize the expected squared forecast error of the combined forecast:

$$\min_{\{\Gamma\}} E[(y_t - \Gamma' Y_t)^2], \tag{6}$$

where $Y_t = (\hat{y}_{1,t},\ \ldots\ , \hat{y}_{k,t})'$ and $\Gamma = (\gamma_1, ..., \gamma_k)'$. I refer to this case as the *optimal weights* case. The optimal weights case is similar to the weights proposed by Bates and Granger (1969) and Granger and Ramanathan (1984) for use in the actual empirical practice of forecasting.

Consistent with the bounded rationality assumption imposed throughout this paper, the "optimal" weights are of course not optimal in a number of real world situations. In particular, the weights derived from Equation (6) do not account for the relative sizes of the different forecast rules or the feasibility of implementing the solution on actual data.[9] However, the weights do capture the key objective of model averaging in a tractable way to provide intuition for how an endogenous weighting strategy interacts with expectations and the economy. In addition, the findings by Evans et al. (2013) with regard to forecast combination with Bayesian model averaging in a self-referential environment suggest the insights developed here carry over to other combination strategies. Alternative combination strategies or different forecast objective functions represent an interesting topic for

[8]See Hendry and Clements (2004) for a more detailed discussion of the forecast combination puzzle.

[9]It is a common finding that optimal weights underperform relative to equal weights due to imprecision from estimating the weights as shown by Smith and Wallis (2009) or by Yang (2004), which is essentially the forecast combination puzzle mentioned previously.

future research.

2.4 Exogenous Weights

A combined forecast with exogenous weights $(\gamma_1, \ \ldots \ , \gamma_k)' = \Gamma \in \mathbb{R}^k$ takes the form of equation (5). An equilibrium is a set of parameter that satisfies the orthogonality condition, equation (4), for each rule simultaneously given that expectations are formed using a combination of all the rules. The combined forecast functions as a perceived law of motion (PLM) for the economy. The PLM can be substituted in for $E_{t-1}y_t$ in equation (1) to yield the actual law of motion (ALM) for the economy

$$y_t = (\Omega' + \alpha \sum_{i=1}^{k} \gamma_i \phi_i' u_i) z_{t-1} + w_t. \tag{7}$$

The ALM describes the actual evolution of y_t given agents' parameter beliefs $\{\phi_i\}_{i=1}^k$. Substituting the ALM into equation (4) and simplifying yields the equilibrium condition for the parameter beliefs for each of the considered forecast rules

$$\phi_i = [(1 - \alpha\gamma_i) u_i \Sigma_z u_i']^{-1} (u_i \Sigma_z \Omega + \alpha \sum_{j \neq i} \gamma_j u_i \Sigma_z u_j' \phi_j) \tag{8}$$

where $E z_{t-1} z_{t-1}' = \Sigma_z$.

Definition 1: *An Exogenous Weight Forecast Combination Equilibrium (EW-FCE) is a set of parameter beliefs $\{\phi_i^*\}_{i=1}^k$ and weights Γ that satisfies the system of equations given by (8) for all i.*

In equilibrium, the ALM of the economy and the i parameter beliefs can be written as

$$y_t = D' z_{t-1} + w_t \text{ and } \phi_i = (u_i \Sigma_z u_i')^{-1} u_i \Sigma_z D.$$

this implies that an EW-FCE must satisfy

$$D' = \Omega' + \alpha \sum_{i=1}^{k} \gamma_i D' \Sigma_z u_i' (u_i \Sigma_z u_i')^{-1} u_i. \tag{9}$$

Remark 1: *There exists a unique Exogenous Weight Forecast Combination Equilibrium if and only if* $\Delta(\Gamma)$ *is invertible, where*

$$\Delta(\Gamma) = I - \alpha \sum_{i=1}^{k} \gamma_i \Sigma_z u_i' (u_i \Sigma_z u_i')^{-1} u_i$$

and I *is an* $n+1 \times n+1$ *identity matrix.*

The same condition given in Remark 1 is derived by Branch and Evans (2006) to show the existence of a Restricted Perceptions Equilibrium in a model with heterogeneous agents. They show that the condition of Remark 1 is always met if Γ is in the unit simplex and α is sufficiently small. Under these assumptions, $\Delta(\Gamma)$ is a diagonal dominant matrix and has a non-zero determinant. Thus, for the aforementioned special case of equal weights, there always exists an α small enough such that an EW-FCE exists.

Remark 2: *There exist multiple equilibria if* $\alpha \sum_{i=1}^{k} \gamma_i = 1$.

This remark is due to the fact that if each individual forecast rule has at least an intercept belief in common, then any intercept belief becomes completely self-confirming when $\alpha \sum_{i=1}^{k} \gamma_i = 1$. This is because any intercept belief is perfectly reflected in the realizations of the data generating process.

The EW-FCE definition also nests an equilibrium that is observationally equivalent to the REE of the model. This FCE is of particular interest because it represents the conditions under which the REE predictions are robust to the boundedly rational behavior proposed in this paper. I call this equilibrium the Fundamental FCE.

Definition 2: *A Fundamental FCE is an FCE that is observationally equivalent to the REE given by equation (2).*

A Fundamental FCE exists assuming the following three conditions:

A1: $\mu = 0$.

A2: The exogenous observables are uncorrelated, i.e. Σ_z *is a diagonal matrix.*

A3: Each forecast rule uses a mutually exclusive set of the exogenous observable shocks, *which together include all elements of x_t. This implies*

$$\sum_{i=1}^{k} u_i' u_i = \begin{bmatrix} \omega & 0 \\ 0 & I \end{bmatrix},$$

where $0 \leq \omega \leq k$ and I is an $n \times n$ identity matrix.

Proposition 1: *Assume A1, A2, A3, $\Gamma = (1, \ \dots \ , 1)'$, and $\alpha \neq \frac{1}{\omega}$, then there exists a unique EW-FCE given by $\phi_i^* = (1-\alpha)^{-1} u_i \Omega$ for all $i = 1, \dots, k$, which is the Fundamental FCE.*

The three conditions required for Proposition 1 are each consistent with standard practices in the macro theory literature or the forecasting literature, and importantly, do not place restriction on the number of forecast rules agents may consider. In particular, A1 is equivalent to redefining the reduced form model in terms of deviations from steady state. A2 eliminates the possibility of omitted variable bias in the equilibrium beliefs for the menu of considered forecast rules by requiring that all the exogenous shocks be uncorrelated. And A3 is akin to requiring that all forecast rules are non-encompassing, meaning that no forecast rule is considered that draws upon information already found in another rule. This restriction is often recommended when constructing combined forecasts (see for example Harvey and Newbold (2000)).

The $0 \leq \omega \leq k$ condition of A3 implies that each forecast rule does not need to include an intercept belief in its specifications since A1 already assumes that the process is mean zero. However, it is standard to always include an intercept and it will be useful in the next section to define a strengthened version of A3.

A3s: Each forecast rule includes an intercept and employs a mutually exclusive set of the exogenous observable shocks, which together include all elements of x_t. This implies

$$\sum_{i=1}^{k} u_i' u_i = \begin{bmatrix} k & 0 \\ 0 & I \end{bmatrix},$$

where I is an $n \times n$ identity matrix.

2.5 Optimal Weights

The first order condition for the optimal Γ from the minimization problem given by (6) can be expressed as an orthogonality condition similar to the one given for agents' beliefs,

$$EY_t(y_t - \Gamma' Y_t) = 0. \tag{10}$$

The orthogonality condition and the ALM given by (7) define a system of equations for the optimal weights as a function of the agents' parameter beliefs $\{\phi_i\}_{i=1}^k$, where the equation for the i^{th} weight is

$$\gamma_i = [(1-\alpha)\phi_i' u_i \Sigma_z u_i' \phi_i]^{-1}(\phi_i' u_i \Sigma_z(\Omega + (\alpha - 1)\sum_{j \neq i} \gamma_j u_j' \phi_j)). \tag{11}$$

Definition 3: *An Optimal Weight Forecast Combination Equilibrium (OW-FCE) is a set of parameter beliefs $\{\phi_i^*\}_{i=1}^k$ and a vector of weights Γ^* that solves the system of equations given by (8) and (11) for all i.*

The equilibrium in this case is defined by a system of polynomial equations, which may have many real solutions. The degree of the system of polynomial equations can grow with the number of forecast rules considered and with respect to the specification of the rules. Therefore, in the general case, there is often not an algebraically tractable solution.[10] However, the solution is tractable in one special case of interest: the case that permits the Fundamental FCE.

Proposition 2: *Assume A1, A2, and A3, then parameter beliefs $\phi_i^* = (1-\alpha)^{-1} u_i \Omega$ for all $i = 1, ..., k$ and combination weights $\Gamma^* = (1, \ \dots \ , 1)'$ constitute a Fundamental OW-FCE.*

For tractability it is also useful to assume:

A4: $\Omega' u_i' u_i \Sigma_z u_i' u_i \Omega = C$ *for all $i = 1, ..., k$, where C is a constant scalar.*

This assumption has no economic interpretation and is imposed to simplify the derivation of the multiple equilibria result that follows. The multiple equilibria result is shown to be robust to this assumption and to assumptions A1 and A2 in the application of the equilibrium concept in Section 3.

Proposition 3: *Assume A1, A2, A3s, A4, and $\frac{1}{k} < \alpha < 1$, then there exist multiple OW-FCE.*

[10] See Sturmfels (2002) for further explanation on the complexity of systems of polynomial equations and applications to economic problems.

Proof: From Proposition 2 it follows that there exists at least one OW-FCE: the Fundamental FCE. Therefore, to prove the proposition I show the existence of other distinct equilibria under the same conditions.

One way multiplicity of equilibrium can arise is through the intercept beliefs. From Remark 2 it follows that any intercept belief is self-confirming if $\alpha \sum_{i=1}^{k} \gamma_i = 1$. Therefore, consider equilibrium parameter beliefs of the form $\phi_i^* = \{(a_i, (1-\alpha\gamma_i)^{-1}\zeta_1, ..., (1-\alpha\gamma_i)^{-1}\zeta_m)' \mid m < n\}$ for $i = 1, .., k$, where a_i is a free variable. These beliefs are obtained using equation (8) and applying assumptions A1, A2, and A3s and are the Fundamental FCE beliefs if $a_i = 0$ for all i.

The equilibrium condition for the i^{th} optimal weight is then given by

$$\gamma_i = [(1-\alpha)\phi_i' u_i \Sigma_z u_i' \phi_i]^{-1}(\phi_i' u_i \Sigma_z (\Omega + (\alpha - 1)\sum_{j \neq i} \gamma_j u_j' \phi_j)).$$

Substituting the equilibrium parameter beliefs $\{\phi_i^*\}_{i=1}^k$ into the condition and rearranging yields

$$\gamma_i(1-\alpha)\Psi = \Xi + (\alpha - 1)\sum_{j \neq i} \Lambda_j,$$

where without loss of generality

$$\begin{aligned}
\Psi &= \phi_i' u_i \Sigma_z u_i' \phi_i = a_i^2 + (1-\alpha\gamma_i)^{-2}(\zeta_1^2\sigma_1^2 + ... + \zeta_{m_i}^2\sigma_{m_i}^2) \\
\Xi &= \phi_i' u_i \Sigma_z \Omega = (1-\alpha\gamma_i)^{-1}(\zeta_1^2\sigma_1^2 + ... + \zeta_{m_i}^2\sigma_{m_i}^2) \\
\Lambda_j &= \gamma_j a_i a_j.
\end{aligned}$$

Now imposing $a_i = a_j$ for all i and j, A4, and using $\alpha \sum_{i=1}^{k} \gamma_i = 1$, it follows that the equilibrium condition for γ_i simplifies to

$$a^2 = \frac{\alpha C(1-\gamma_i)}{(1-\alpha)(1-\alpha\gamma_i)^2}.$$

One possible set of equilibrium weights is given by

$$\gamma_i = \frac{C + 2a^2(\alpha-1) \pm \sqrt{C\left(C - 4a^2(\alpha-1)^2\right)}}{2a^2(\alpha-1)\alpha}.$$

Finally, picking either solution for γ_i, imposing $\alpha \sum_{i=1}^{k} \gamma_i = 1$ for all i, and solving the equation for a yields

$$a = \pm \frac{\sqrt{k}\sqrt{C(1-k\alpha)}}{(k-1)\sqrt{\alpha-1}}. \tag{12}$$

The solutions for a are real and non-zero if $\frac{1}{k} < \alpha < 1$. Therefore, there exist multiple distinct OW-FCEs if $\frac{1}{k} < \alpha < 1$. $\square$

There are two takeaways from Proposition 3 and the OW-FCE case. The first is that the Fundamental OW-FCE coexists with other non-fundamental equilibria. Therefore, the economy may obtain equilibria that depart from the rational predictions of the model even when the Fundamental FCE is a possible outcome.[11] The second is that the range of parameter values for which multiple equilibria exist is dependent on the number of forecast rules considered. As the number of forecast rules considered increases, the parameter space for which there is a unique OW-FCE shrinks. This implies that multiple equilibria can arise in any parameterization of the model with positive feedback as long as agents consider enough different forecast rules.

In addition, the multiplicity result does not rely on agents holding some type of flawed belief. For example, none of the beliefs that parameterize the forecasts rules exhibit omitted variable bias due to assumption A2.[12] There is also no bias present in any of the individual forecasts or in the combined forecast by definition of the equilibrium. Therefore, although the agents are boundedly rational, the multiple equilibria result does not require agents to entertain beliefs that an econometrician could readily fault.[13]

3 FCE in Two Macro Models

This section applies the FCE concept to a Lucas-type monetary model and to a Fisherian model to illustrate the multiple equilibria and discuss their connection to the underlying economics of the models. To assess the plausibility of the equilibria, least squares learning is used as an equilibrium selection mechanism.

[11] This is similar to the findings of Evans et al. (2013).

[12] The existence of these biases also does not rule out the existence of multiple equilibria.

[13] A remaining restriction that an econometrician may want to impose is that the weights sum to one. This case is considered in Section 3.

3.1 Least Squares Learning

A natural equilibrium selection mechanism given the definition of a Forecast Combination Equilibrium and the assumption of bounded rationality is least squares learning of the form described by Evans and Honkapohja (2001). Least squares learning assumes that agents do not know the equilibrium parameter values of the menu of forecast rules or the combination weights, but must estimate them using past data. The estimation is assumed to occur recursively with each time period yielding a new realization of the endogenous variables of interest, which agents use to update their beliefs. The beliefs that are recovered asymptotically from this recursive estimation procedure from nearby initial beliefs are said to be "learnable" and are deemed the plausible equilibrium outcomes of the model.

Least squares learning for forecast combination with optimal weights requires agents to estimate $k+1$ regressions. A regression for each of the k forecast rules and one regression to estimate the optimal combination weights. Stacking the regressions appropriately, the estimation procedure is written recursively as

$$\begin{aligned} \Theta_t &= \Theta_{t-1} + \kappa_t \mathbf{R}_t^{-1} \mathbf{z}_{t-1}(\mathbf{y}_t - \mathbf{z}'_{t-1}\Theta_{t-1}) \\ \mathbf{R}_t &= \mathbf{R}_{t-1} + \kappa_t(\mathbf{z}_{t-1}\mathbf{z}'_{t-1} - \mathbf{R}_{t-1}), \end{aligned} \tag{13}$$

where Θ_t is a vector containing the parameter estimates for all considered models and the combination weights, $\mathbf{z}_t$ is a block diagonal matrix containing the appropriate exogenous shocks and past realization of the forecasts, and κ_t is the gain sequence that governs the weight given to new observations.[14] If $\kappa_t = t^{-1}$, then the recursion is equivalent to an ordinary least squares regression. The usefulness of least squares learning as a selection criteria for FCEs is seen by noting that $E\mathbf{z}_{t-1}(\mathbf{y_t} - \mathbf{z}'_{t-1}\Theta_{t-1})$ is exactly the equilibrium condition given in the previous section for an OW-FCE. Therefore, steady states of Equation (13) corresponds to the OW-FCEs of the underlying model.

The convergence to a given OW-FCE from nearby initial beliefs can be determined by appealing to the E-stability principle. The E-stability principle states that the stability of a fixed point of the recursive least squares algorithm is determined by the stability of an associated differential equation in notional time

$$\dot{\Theta} = T(\Theta) - \Theta,$$

where $\dot{\Theta}$ is given by rearranging equations (8) and (11), the equilibrium conditions for

[14] An explicit example of how the regressions are stacked is given in the appendix.

an OW-FCE. The function $T(\Theta)$ is known as the T-map.[15]

The E-stability of a fixed point of the model under study is usually dependent on the value of α, the feedback parameter on expectations. The standard result in the literature is that if agents consider a single correctly specified PLM, then the resulting fixed point is the REE and it is E-stable provided that $\alpha < 1$.[16]

3.2 The Lucas-type Monetary Model

The Lucas-type monetary model I consider follows Branch and Evans (2007). The economy is described by an aggregate supply equation (AS), an aggregate demand equation (AD),

$$\begin{aligned} AS: \quad y_t &= \xi(p_t - E^*_{t-1}p_t) + \rho_1' x_t && (14) \\ AD: \quad y_t &= m_t - p_t + \rho_2' x_t + v_t, && (15) \end{aligned}$$

and a monetary policy rule (MP) given by

$$MP: \; m_t = p_{t-1} + \rho_3' x_t + \eta_t. \tag{16}$$

The variables p_t and m_t are logs of the price level and money supply respectively, y_t is the log deviation of output from a deterministic trend, x_t is a 2×1 vector of serially correlated shocks, and v_t and η_t are i.i.d. white noise. The shocks follow a stationary VAR process

$$x_t = Bx_{t-1} + \epsilon_t. \tag{17}$$

Inflation in the economy is defined as $\pi_t = p_t - p_{t-1}$.

Imposing the inflation definition and combining equations (14), (15), and (16) yields an expectation augmented Phillips curve in the familiar reduced form:

$$\pi_t = \alpha E^*_{t-1}\pi_t + \zeta' x_{t-1} + w_t, \tag{18}$$

where $\alpha = \frac{\xi}{1+\xi}$, $\psi = \frac{(\rho_2' + \rho_3' - \rho_1')}{1+\xi}$, $\zeta' = \psi B$ and $w_t = \psi\epsilon_t + \frac{\eta_t + v_t}{1+\xi}$. The agents consider all non-trivial underparameterizations of the Phillips curve as forecast rules:

[15]The derivation is shown in the appendix. Guse (2008) shows that the technical conditions for the stochastic recursive algorithm theorems that underpin the E-stability principle are satisfied for the block recursive least squares algorithm.

[16]See Evans and Honkapohja (2001) for a complete analysis and discussion of this result.

$$
\begin{aligned}
\hat{\pi}_{1,t} &= a_1 + b_1 x_{1,t-1} \\
\hat{\pi}_{2,t} &= a_2 + b_2 x_{2,t-1}.
\end{aligned}
$$

The agents are assumed to create optimal combined forecasts to minimize the expected squared forecast error of their combined forecast.

3.2.1 Existence and Stability

To show the existence and E-stability of multiple equilibria in the Lucas monetary model, I first consider the case that permits the Fundamental OW-FCE by assuming A1, A2, and A3. The T-map for this case is given by[17]

$$
T\begin{pmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \alpha(a_1\gamma_1 + a_2\gamma_2) \\ \alpha\gamma_1 b_1 + \zeta_1 \\ \alpha(a_1\gamma_1 + a_2\gamma_2) \\ \alpha\gamma_2 b_2 + \zeta_2 \\ \frac{a_1^2\alpha\gamma_1 + a_1 a_2\gamma_2(\alpha-1) + b_1\sigma_1^2(b_1\alpha\gamma_1+\zeta_1)}{a_1^2+b_1^2\sigma_1^2} \\ \frac{a_2^2\alpha\gamma_2 + a_1 a_2\gamma_1(\alpha-1) + b_2\sigma_2^2(b_2\alpha\gamma_2+\zeta_2)}{a_2^2+b_2^2\sigma_2^2} \end{pmatrix}. \tag{19}
$$

The mapping is non-linear and has seven distinct fixed points. The algebraic representations of the solutions, however, are immense and are impractical to list here. Therefore, I use a bifurcation argument to show existence and E-stability for a subset of the equilibria.

To this aim, consider the associated differential equation for least squares learning in this case

$$
\dot{\Theta} = T(\Theta) - \Theta \tag{20}
$$

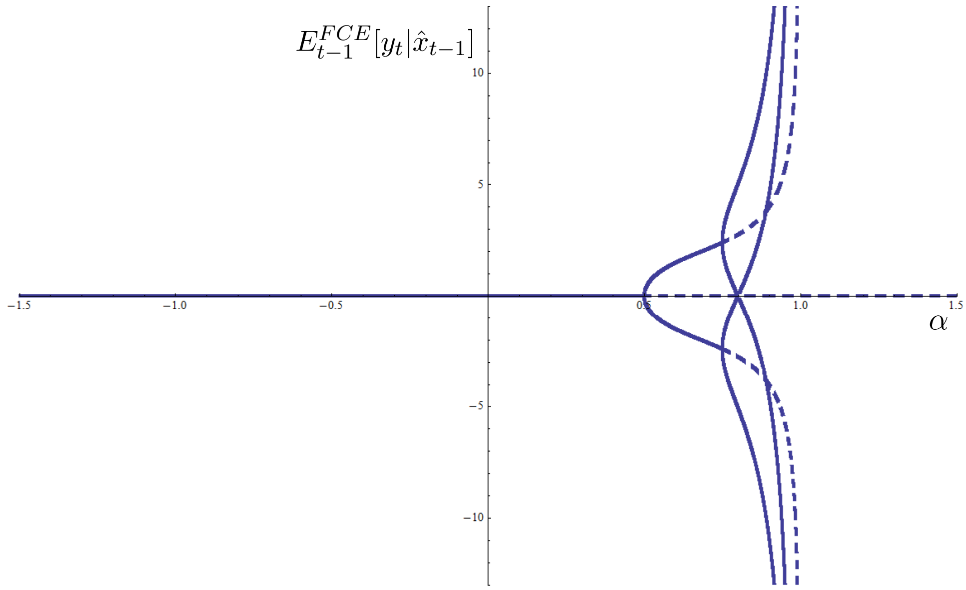
and the Fundamental FCE fixed point. Since $k = 2$, Proposition 3 says that multiple equilibria will arise in this model when $\alpha > 1/2$, which suggests that a bifurcation of the system occurs at $\alpha = 1/2$. By characterizing the type of bifurcation that occurs at this point, the number of equilibria and the E-stability of those equilibria can be determined.

Lemma 1: *The Fundamental OW-FCE steady state of the dynamic system given by (19) and (20) experiences a supercritical pitchfork bifurcation at* $\alpha = 1/2$.

A supercritical pitchfork bifurcation occurs when a fixed point switches stability from

[17] The derivation of the following T-map is given in the appendix.

Figure 1: Multiple equilibria in the Lucas monetary model



Notes: Illustration of the OW-FCEs in the Lucas-type monetary model for different values of α, the feedback parameter on expectations. The vertical axis depicts the equilibrium expectation for a fixed realization of x_t. E-stability is indicated by the solid lines. Parameter values are given in the text.

stable to unstable and two new stable equilibria come into existence.

Proposition 4: *For the economy under study (18) represented by (19)*

1. *There exist at least three OW-FCEs for* $\frac{1}{2} < \alpha < 1$.
2. *The Fundamental OW-FCE is E-stable for* $\alpha < \frac{1}{2}$.
3. *At least two non-fundamental OW-FCEs are E-stable for some* $\frac{1}{2} < \alpha < 1$.

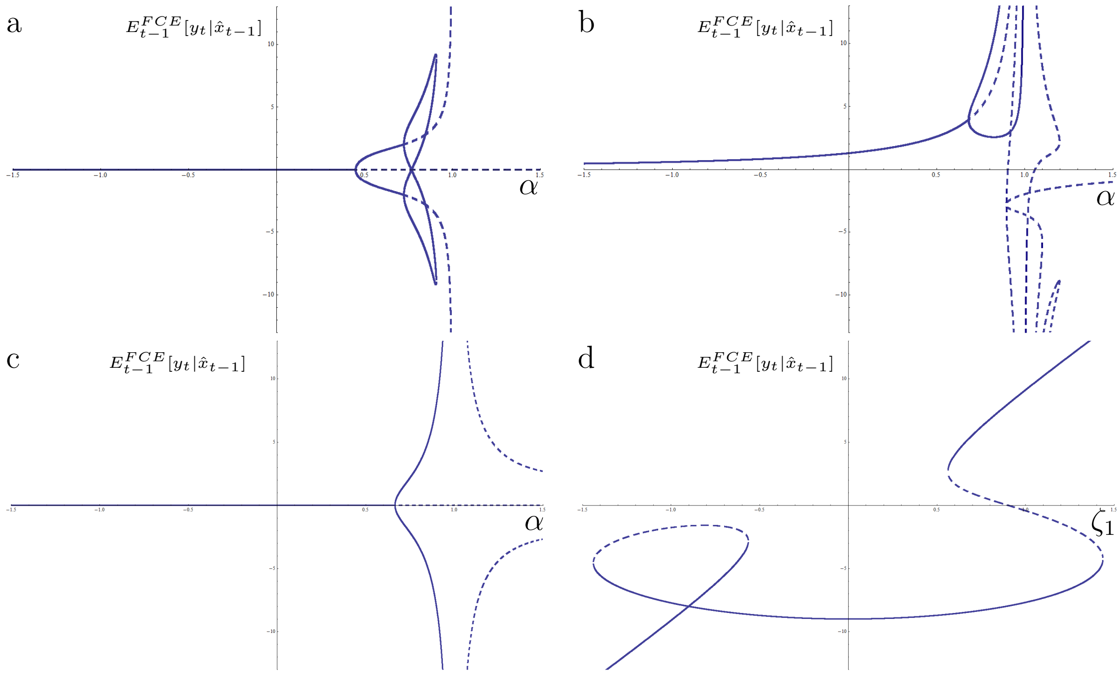
Proposition 4 shows that although the Fundamental OW-FCE is a possible equilibrium, it is not the equilibrium the economy will achieve under learning when $\frac{1}{2} < \alpha < 1$.[18] The economy instead will obtain other non-fundamental equilibria that are distinct from the RE predictions of the model.

All seven distinct equilibria and their E-stability properties are illustrated numerically in Figure 1. The figure plots α on the horizontal axis and the corresponding conditional equilibrium forecasts ($E_{t-1}^{FCE}[y_t|\hat{x}_{t-1}]$) implied by all OW-FCEs that exist for a given α on the vertical axis.[19] The figure is similar to a bifurcation diagram in that it plots an

[18] Evans et al. (2013) also find that $\alpha = \frac{1}{2}$ is the boundary for non-rational equilibria to exist in their simulations.

[19] The remaining parameters are $\zeta_1 = .9$, $\zeta_2 = -.9$, $Ex_tx_t' = I$, and $\hat{x}_{t-1} = (1,\ 1)'$. The equilibrium forecasts are constructed as $E_{t-1}^{FCE}y_t = \gamma_1^*(a_1^* + b_1^* \times 1) + \gamma_2^*(a_2^* + b_2^* \times 1)$, where $*$ denotes the OW-FCE values. The range of the figures includes negative values of α to illustrate the Cobweb case for the

Figure 2: Robustness of the multiple equilibria



Notes: Illustration of the OW-FCEs in the Lucas monetary model for different values of α, the feedback parameter on expectations. The first diagram (a) depicts the equilibria when there is a positive covariance between the exogenous shocks ($\sigma_{12} = .1$). The second (b) diagram depicts the equilibria when there is a positive intercept ($\mu = 1$). The third diagram (c) depicts the equilibria when the combination weights are restricted to sum to one. The fourth digram (d) depicts the equilibria for a fixed α (0.9) with different responses to shocks in the monetary policy rule. The weights in this case are also restricted to sum to one. The remaining parameter values are the same as those used in Figure 1. E-stability is indicated by the solid lines.

equilibrium outcome as a function of a parameter of interest. E-stability of the equilibria are indicated by solid lines and E-instability is indicated by dashed lines.

The figure shows the classic 'pitchfork' shape associated with the bifurcation identified in Lemma 1. The Fundamental FCE is the unique equilibrium for $\alpha < \frac{1}{2}$. At $\alpha = \frac{1}{2}$, the pitchfork bifurcation occurs and two new equilibrium come into existence. A second bifurcation occurs at $\alpha = \frac{3}{4}$ and results in a total of seven OW-FCEs, four of which are E-stable.

3.2.2 Exploration and Robustness

The key to multiple equilibria in OW-FCEs is retaining the non-linearity in the weights. Any modification that retains the non-linearity will produce multiple OW-FCEs that follow a similar pattern to the one depicted in Figure 1, where there exists a single stable equilibrium for small and negative α and multiple, E-stable equilibria for some $0 < \alpha < 1$.

Figure 2 provides four numerical examples to illustrate the claim. The first two plots

reduced form model as well as the case of interest. There is a unique OW-FCE if the model exhibits negative feedback. An example in the literature of the Cobweb case is Branch and Evans (2006).

show that the existence and stability results extend to the case where the Fundamental OW-FCE assumptions are not met. The first plot shows the equilibria and stability results when there is a non-zero covariance between the elements of x_t ($\sigma_{12} = .1$), while the second plot shows the equilibria and stability results when there is a non-zero intercept ($\mu = 1$).[20]

The third plot in Figure 2 depicts the case where the optimal weights are restricted to sum to one. This case is of particular interest because this restriction is often imposed in the forecasting literature. Restricted optimal weights imply the following orthogonality condition for γ_1

$$E(\hat{\pi}_{1,t} - \hat{\pi}_{2,t})[(\pi_t - \hat{\pi}_{2,t}) - \gamma_1(\hat{\pi}_{1,t} - \hat{\pi}_{2,t})] = 0, \tag{21}$$

which follows directly from solving the constrained expected squared forecast error minization problem. The T-map under this assumption and assumptions A1, A2, and A3 is given by

$$T\begin{pmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ \gamma_1 \end{pmatrix} = \begin{pmatrix} \alpha(a_2 + \gamma_1(a_1 - a_2)) \\ \alpha\gamma_1 b_1 + \zeta_1 \\ \alpha(a_2 + \gamma_1(a_1 - a_2)) \\ \alpha(1-\gamma_1)b_2 + \zeta_2 \\ \frac{(a_1-a_2)(a_1\alpha\gamma_1 + a_2(\alpha(1-\gamma_1)-1) + b_1\sigma_1^2(b_1\alpha\gamma_1+\zeta_1) + b_2\sigma_2^2(b_2(1+\alpha(\gamma_1-1)) - \zeta_2)}{(a_1-a_2)^2 + b_1^2\sigma_1^2 + b_2^2\sigma_2^2} \end{pmatrix}. \tag{22}$$

The T-map again defines a system of polynomial equations with multiple solutions.[21] The numerical solutions in Figure 2 shows that the equilibrium and stability results in this case are analogous to those in the unrestricted case. There exists a unique E-stable equilibrium for small or negative α and two E-stable equilibria for some positive $0 < \alpha < 1$.

Finally, the bifurcation of equilibria in the restricted case and the second bifurcation of equilibria in the unrestricted case occur in the b_i beliefs and depend on ζ as well as α. The ζ vector, as shown in equation (18), is a function of the monetary policy rule. Therefore, uniqueness of FCEs depends on monetary policy for some values of α when weights are restricted to sum to one. Figure 2, panel IV illustrates this dependence for a

[20]The two assumptions violate A2 and A1, respectively.

[21]The multiple equilibria in this case are different from the unrestricted case because every equilibrium shares a common intercept belief. This is due to the fact that $\alpha(\gamma_1 + \gamma_2) = 1$ cannot be satisfied, except trivially when $\alpha = 1$. Therefore, bifurcation of the unique equilibrium as α increases in this case occurs in the b_i beliefs.

fixed α as the monetary policy response to the first shock is varied. For some values of ζ_1, there is a single E-stable equilibrium. While for others, there are two E-stable equilibria.

3.3 The Fisherian Model

Following Cochrane (2011), Branch and Evans (2011b), and Anufriev et al. (2013), I consider a frictionless endowment economy populated by a continuum of identical consumers who seek to maximize the present discounted value of utility over consumption:

$$\max E_t \sum_{j=0}^{\infty} \left(\beta e^{x_{1,t+j}}\right)^j u(C_{t+j}),$$

subject to the budget constraint

$$P_t C_t + B_t = (1 + i_{t-1}) B_{t-1} + P_t Y,$$

where P_t is the price of goods, B_t represents one-period nominal bonds, i_t is the nominal interest rate, Y is the endowment, and $x_{1,t}$ is a mean zero AR(1) discount rate shock.[22] The first order conditions of the consumer's problem yield the standard consumption Euler equation and implies a Fisher relationship given by

$$\frac{1}{1+i_t} = \beta e^{x_{1,t}} E_t \left(\frac{P_t}{P_{t+1}} \right).$$

Taking logs of the Fisher relationship yields the familiar linear approximation

$$i_t = r + E_t \pi_{t+1} + \delta_1 x_{1,t}. \tag{23}$$

The model is closed by assuming monetary policy follows a Taylor rule and targets zero percent inflation:

$$i_t = r + \theta_\pi \pi_t + \delta_2 x_{2,t}, \tag{24}$$

where $x_{2,t}$ is AR(1) policy shocks. Combining Equation (23) and (24), inflation in this economy is described by

$$\pi_t = \frac{1}{\theta_\pi} E_t \pi_{t+1} + \frac{1}{\theta_\pi} \delta' x_t, \tag{25}$$

[22] β is the discount factor and it is assumed to be between zero and one.

which in the reduced form notation of Section 2 may be written generally as

$$\begin{aligned} y_t &= \mu + \alpha E_t y_{t+1} + \zeta' x_t && (26) \\ x_t &= \rho x_{t-1} + \omega_t. && (27) \end{aligned}$$

The agents again consider all non-trivial underparameterizations of the inflation process as forecast rules:

$$\begin{aligned} \hat{\pi}_{1,t} &= a_1 + b_1 x_{1,t} \\ \hat{\pi}_{2,t} &= a_2 + b_2 x_{2,t} \end{aligned}$$

and create optimal combined forecasts to minimize the expected squared forecast error of their combined forecast.[23]

3.3.1 Existence and Stability

To show the existence and E-stability of multiple equilibria in the Fisherian model, I again consider the case that permits the Fundamental OW-FCE by assuming A1, A2, and A3. The T-map is given by

$$T\begin{pmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \alpha(a_1\gamma_1 + a_2\gamma_2) \\ \alpha\gamma_1\rho_1 b_1 + \zeta_1 \\ \alpha(a_1\gamma_1 + a_2\gamma_2) \\ \alpha\gamma_2\rho_2 b_2 + \zeta_2 \\ \frac{a_1 a_2(\alpha-1)\gamma_2 + b_1\sigma_1^2\zeta_1\rho_1^2}{a_1^2(1-\alpha) + b_1^2\sigma_1^2\rho_1^2(1-\alpha\rho_1)} \\ \frac{a_1 a_2(\alpha-1)\gamma_1 + b_2\sigma_2^2\zeta_2\rho_2^2}{a_2^2(1-\alpha) + b_2^2\sigma_2^2\rho_2^2(1-\alpha\rho_2)} \end{pmatrix}, \quad (28)$$

which again has seven fixed points. Treating α as the bifurcation parameter, the system experiences a bifurcation at $\alpha = 1/2$.[24]

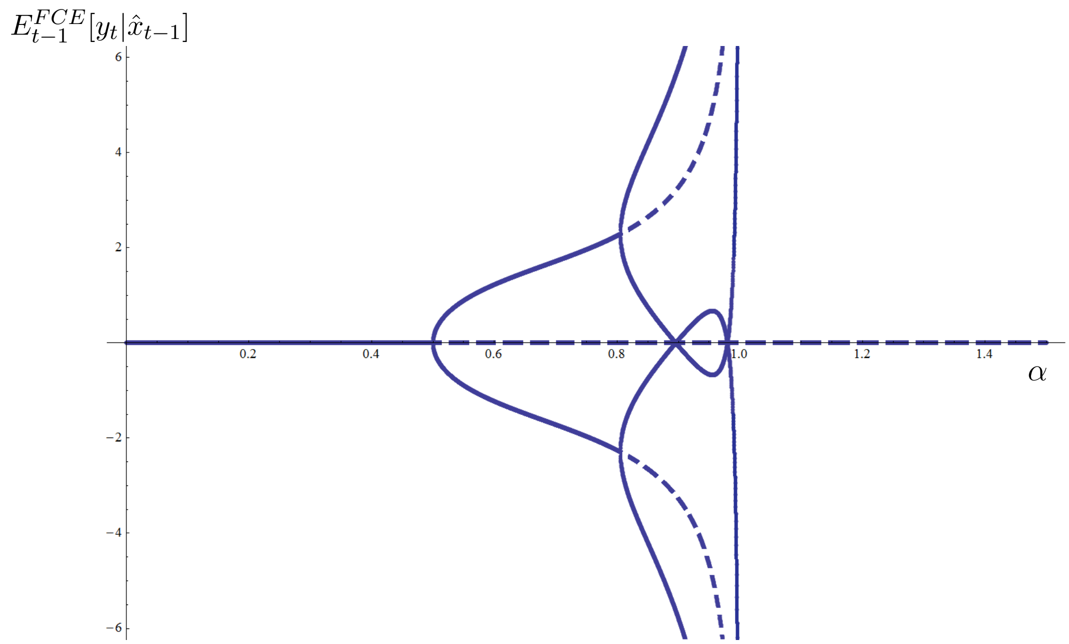
Lemma 2: *The Fundamental OW-FCE steady state of the dynamic system given by (20) and (28) experiences a supercritical pitchfork bifurcation at $\alpha = 1/2$.*

---

[23] As is standard in the adaptive learning literature, it is assumed that the agents know ρ since it is an exogenous observable process, which may be learned with probability one using least squares.

[24] The relationship between α and ζ implied by the underlying model is ignored for the analysis in this section. The relationship does not affect the characteristics or the position with respect to α of the first bifurcation. Numerical simulations with the relationship imposed show that all equilibrium and E-stability results are qualitatively unchanged, although, the timing of the second round of bifurcations and the onset of multiple equilibria in the constrained weights case is altered relative to the fixed ζ case.

Figure 3: Multiple equilibria in the Fisherian model



Notes: Illustration of the OW-FCEs in the Fisherian model for different values of α, the feedback parameter on expectations. The vertical axis depicts the equilibrium expectation for a fixed realization of x_t. E-stability is indicated by the solid lines. Parameter values are given in the text.

Proposition 5: *For the economy under study (25) represented by (28)*

1. *There exist at least three OW-FCEs for $\frac{1}{2} < \alpha < 1$.*

2. *The Fundamental OW-FCE is E-stable for $\alpha < \frac{1}{2}$.*

3. *At least two non-fundamental OW-FCEs are E-stable for some $\frac{1}{2} < \alpha < 1$.*

The full set of equilibria are illustrated in Figure 3 using the same type of diagram employed in the previous section.[25] The number and stability of equilibria are the same as in the previous case. Multiple equilibria, though, is particularly interesting in this case because it is always a function of policy since α is a policy choice ($\alpha = \theta_\pi^{-1}$).

The standard analysis of this model, under RE, requires $\theta_\pi > 1$ in order for a unique REE to exist. This of course coincides with the well-known Taylor principle in this model.[26] Under FCE, however, this policy prescription is altered. Here multiple, learnable equilibria exist if $1 < \theta_\pi < 2$, which have non-zero inflation steady states. For

[25]Figure 3 parameters are $\zeta_1 = .9$, $\zeta_2 = -.9$, $Ex_t x_t' = I$, $\rho_1 = 0.9$, $\rho_2 = 0.9$, and $\hat{x}_{t-1} = (1,\ 1)'$.

[26]See Taylor (1993, 1999) for discussion and background on the Taylor rule. Many other factors can also affect determinacy as shown by Bénassy (2006) or Gliksberg (2009). Cochrane (2011) disputes whether the Taylor principle is sufficient and argues that explosive inflation paths cannot be ruled out. However, Evans and McGough (2015) shows that the explosive paths are in general not learnable.

example, at the parameter values used in Figure 3 and assuming $\theta_\pi = 1.5$, the two learnable FCEs have steady state inflation rates of 1.04% and -1.04%, respectively.[27] Whereas for $\theta_\pi = 3$, there is a unique Fundamental FCE with a steady state inflation rate of 0%. Therefore, more aggressive monetary policy is required in the presence of model uncertainty to ensure a unique and learnable equilibrium with inflation at target.

3.4 Discussion

The economic explanation for the pattern of a unique equilibria bifurcating into multiple stable equilibria as α increases lies in the positive feedback of expectations. When α is close to one, all beliefs have a self-fulfilling quality. The data always moves in the same direction as beliefs. The movement of the data towards beliefs is reinforced by the optimal weights, which allows for some beliefs to become self-fulfilling.

The agents in this process are in some sense subject to a perverse form of the Lucas Critique. This is because the reduced form correlations in the data that are used to calculate the optimal weights are a function of the forecast combination strategy. However, unlike the Lucas Critique, where changes in policy alter the reduced form correlations and invalidate policy choices, the choice of optimal weights may reinforce these correlations. The reinforcement prevents agents from detecting that their forecasts deviate from the fundamentals because their beliefs are consistent with past data.

There is nothing unique to the models studied in this paper with respect to the positive feedback necessary to generate the multiple equilibria and stability results. In fact, it is commonly found that the equilibria and E-stability results of learning models in simple settings carry over to richer, more realistic settings. Therefore, it is likely that any linear rational expectations model with a learnable REE will also have multiple, E-stable equilibria under forecast combination when there exists significant positive feedback.

4 Real-time Application and Comparison to DPS

One of the primary contributions of DPS and the learning and expectations literature is to show how expectation driven fluctuations can capture features of actual laboratory and real world economic data that are not explained by standard RE models. This section illustrates that optimal forecast combination with constant gain learning continues this tradition.

[27] $\theta_\pi = 1.5$ corresponds to the value Taylor (1993) argues best describes U.S. monetary policy.

4.1 Constant Gain Learning

Optimal weight forecast combination is implemented by assuming that the agents estimate the parameters of their forecast rules and the optimal combination weights in real time using constant gain learning. Constant gain learning replaces the usual $\kappa_t = t^{-1}$ in the recursive least squares algorithm (13) with a constant $0 < \bar{\kappa} < 1$. A constant gain places more weight on the most recent observations. The stability of an equilibrium under constant gain learning for a sufficiently small gain is still governed by E-stability.[28] The main difference in this case is that the convergence is now in terms of a distribution of beliefs centered at the equilibrium. Small shocks in the model, therefore, may occasionally push expectations from the basin of attraction of one stable OW-FCE to another.[29]

Constant gain learning is argued by Sargent (2001), Orphanides and Williams (2005, 2006), and Branch and Evans (2006b, 2007), to be an appropriate learning strategy when agents are concerned about structural breaks. In addition, the use of constant gain learning is found to be a route to interesting dynamics in Lucas-type monetary models in the case of model misspecification (Cho et al. (2002), McGough (2006)) and Dynamic Predictor Selection (Branch and Evans (2007)) and is thus a natural assumption to employ.

4.2 Comparison to Dynamic Predictor Selection

Branch and Evans (2007) consider the Lucas-type monetary model discussed in Section 3 and model uncertainty as an explanation for the time-varying volatility of inflation observed in U.S. data. In their setup, agents consider a menu of forecast rules, $\Pi_t = (\hat{\pi}_{1,t},\ \hat{\pi}_{2,t})'$, identical to the menu considered in Section 3. The agents estimate the parameter beliefs of the rules using recursive least squares and select a single forecast in each period by comparing past mean squared forecast errors (MSFE). The MSFE of each model is tracked recursively using the following equation:

$$MSFE_{i,t} = MSFE_{i,t-1} + \lambda((\pi_t - \hat{\pi}_{i,t})^2 - MSFE_{i,t-1}) \tag{29}$$

for $i = 1, 2$, where λ is the weight placed on the most recent observation. The proportion of agents who choose each model based on this fitness criterion is determined by

[28] See Evans and Honkapohja (2001) for a thorough treatment of this result.

[29] I do not simulate the decreasing gain case because the E-stability analysis fully characterizes the limiting dynamics. Given a set of initial conditions in a neighborhood of any of the E-stable OW-FCEs, the learning algorithm will converge asymptotically with probability one to the OW-FCE in that neighborhood.

$$n_{i,t} = \frac{exp(-\beta MSFE_{i,t-1})}{\sum_{i=1}^{k} exp(-\beta MSFE_{i,t-1})}, \tag{30}$$

where $n_{i,t}$ is the proportion of agents who choose the i^{th} model and β is a parameter that governs the relative speed at which agents abandon an underperforming forecast rule. The aggregate expectation in the economy is a linear combination of the $\hat{\pi}_{i,t}$'s with the $n_{i,t}$'s as the weights

$$E^*_{t-1}\pi_t = \sum_{i=1}^{k} n_{i,t}\hat{\pi}_{i,t}. \tag{31}$$

The equilibrium concept in this framework is a Misspecification Equilibrium (ME), which was first introduced by Branch and Evans (2006). An ME is characterized by the same condition given in Remark 1, plus a fixed point in the proportion of agents that choose each rule given by equation (30). Branch and Evans (2007) show that there are three possible MEs in this model: two MEs where all agents choose the same forecast rules and a third where some fraction of agents chooses each rule. Analytic E-stability results are not available for these equilibria, but they report that, in simulations, the two homogeneous equilibria are stable under learning.

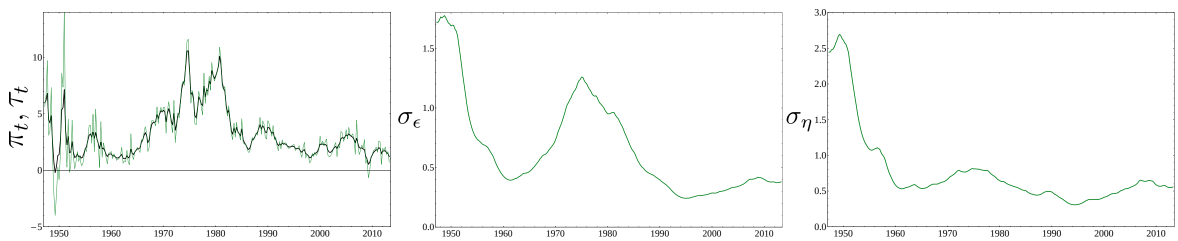
4.2.1 Theoretical Comparison

The key theoretical difference between DPS and FCE is the assumption of how agents use available information. The DPS selection assumption requires agents to disregard potentially useful information from the menu of forecasts based off small differences in MSFE. Models which have proven useful in the past under this assumption are completely ignored when another model performs well over the most recent period. A forecast combination strategy on the other hand allows agents to use all the information they possess to form forecasts, including using past forecast performance to inform combination weights.

In addition, the model selection by the agents in DPS is only partially data driven. The intensity of choice parameter β plays a large role in determining the weights and their transitions over time. This parameter is exogenous and not data dependent. There is no equivalent parameter with forecast combination. The relative weights are chosen endogenously in a way that is consistent with practices prescribed by the forecasting literature. Therefore, in some sense, the forecast combination assumption is a more individually rational response to model uncertainty than DPS.

A drawback to FCE though is the homogeneous expectations assumption. Heterogeneity of expectations is of course a documented feature in surveys of expectations and

Figure 4: UC-SV estimates of U.S. inflation



Notes: UC-SV estimates for the GDP Deflator measure of U.S. inflation. The estimation sample is 1947q1 through 2013q2.

in expectations observed in laboratory experiments. However, Hommes (2013) notes that in many laboratory experiments there is often convergence among participants to near identical expectations that in aggregate exhibit complicated dynamics. The explanation provided by DPS in these cases relies primarily on heterogeneity of forecast rule choices over time rather than cross sectionally. Therefore, there is empirical support for considering homogeneous expectations as an explanation for observed laboratory behavior.[30]

4.2.2 Real-time Learning Comparison

Again following Branch and Evans (2007), I compare and contrast FCE and DPS's ability to qualitatively match non-linear time series properties of U.S. inflation. Stock and Watson (2007) argue that the inflation process in the US is described well by parsimonious unobserved components stochastic volatility model (UC-SV) that allows for changing volatility in both permanent and transient shocks. They model the inflation process as

$$
\begin{aligned}
\pi_t &= \tau_t + \eta_t, \text{ where } \eta_t = \sigma_{\eta,t}\zeta_{\eta,t} \\
\tau_t &= \tau_{t-1} + \epsilon_t, \text{ where } \epsilon_t = \sigma_{\epsilon,t}\zeta_{\epsilon,t} \\
\ln\sigma^2_{\eta,t} &= \ln\sigma^2_{\eta,t-1} + \psi_{\eta,t} \\
\ln\sigma^2_{\epsilon,t} &= \ln\sigma^2_{\epsilon,t-1} + \psi_{\epsilon,t},
\end{aligned}
$$

where $\zeta_t = (\zeta_{\eta,t}, \zeta_{\epsilon,t})$ is i.i.d. $N(0, I_2)$, $\psi_t = (\psi_{\eta,t}, \psi_{\epsilon,t})$ is i.i.d. $N(0, 0.2I_2)$, and ζ_t and ψ_t are independently distributed. Figure 4 illustrates the UC-SV estimates of the time-varying trend and the standard deviations of the two shocks for the GDP Deflator measure of U.S. inflation. The inflation series exhibits non-linear features with changes in volatility to both transient and permanent shocks over time.

Figure 5 illustrates a typical DPS and FCE simulation with constant gain learning

[30]Homogeneity of expectations also implies that no asset trading would take place in a equilibrium in an asset trading model. Extending FCE to the heterogeneous agent case is an interesting topic for future research.

in the Lucas monetary model. The left panel of the figures shows π_t and the intercept belief $a_{1,t}$. The right panel shows the estimated combination weights for FCE and the proportion of agents who choose the first forecast rule under DPS. The same sequence of i.i.d. shocks is used in both simulation. FCE and DPS both generate time-varying volatility, however, only FCE generates sharp movements in the trend of inflation over time.[31]

The sharp movements in trend inflation in the FCE case occur because the optimal weights are estimated to be near one (shown by the shaded areas). Optimal weights equal to one in this case correspond to the unstable Fundamental FCE. Both optimal weight beliefs near one are essentially beliefs that the forecast rules are biased towards zero. The combination weights compensate for this perceived bias by scaling the forecasts away from zero.[32] Because of positive feedback, this belief is initially self-confirming and is incorporated into the estimate of each forecast rule's intercept, which drives the trend of inflation far away from its steady state value.

Similar behavior also occurs in the FCE case where weights are restricted to sum to one. Figure 6 shows $a_{1,t}$ for DPS, unrestricted FCE, and restricted FCE, where identical shocks generate three distinct predictions. The intercept beliefs track each other closely for the first 600 periods before significant breaks occur in the FCE beliefs. The break in the restricted weights case occurs for a similar reason as the break in the unrestricted case. One of the weights is estimated to be larger than one, which implies a belief that the corresponding forecast rule is biased towards zero. Again the positive feedback of expectations makes this belief initially self-confirming, which drives beliefs and trend

[31] Constant gain learning on its own generates some variation in the second movements of inflation even in the case of a unique equilibrium. It can also generate escape dynamics, where trend inflation moves abruptly away from steady state values as explored by Cho et al. (2002) and McGough (2006). Therefore, FCE is not a necessary condition to generate this type of behavior.

[32] Optimal weights larger than one are a known feature in the actual practice of forecasting and will arise under quite general conditions including when weights are restricted to sum to one. To formally illustrate this argument, I present the derivation of optimal weights given by Timmermann (2006) who considers forecasting a mean zero process with two unbiased forecast rules. Ignoring time, the two forecast rules' errors can be written as $fe_1 = \pi - \hat{\pi}_1$ and $fe_2 = \pi - \hat{\pi}_2$, where $fe_1 \sim (0, \sigma_1^2)$, $fe_2 \sim (0, \sigma_2^2)$, $\sigma_{12} = \rho_{12}\sigma_1\sigma_2$, and ρ_{12} is the correlation of the forecast errors. The combined forecast error is given by $fe_c = \gamma_1 fe_1 + (1-\gamma_1) fe_2$, which implies

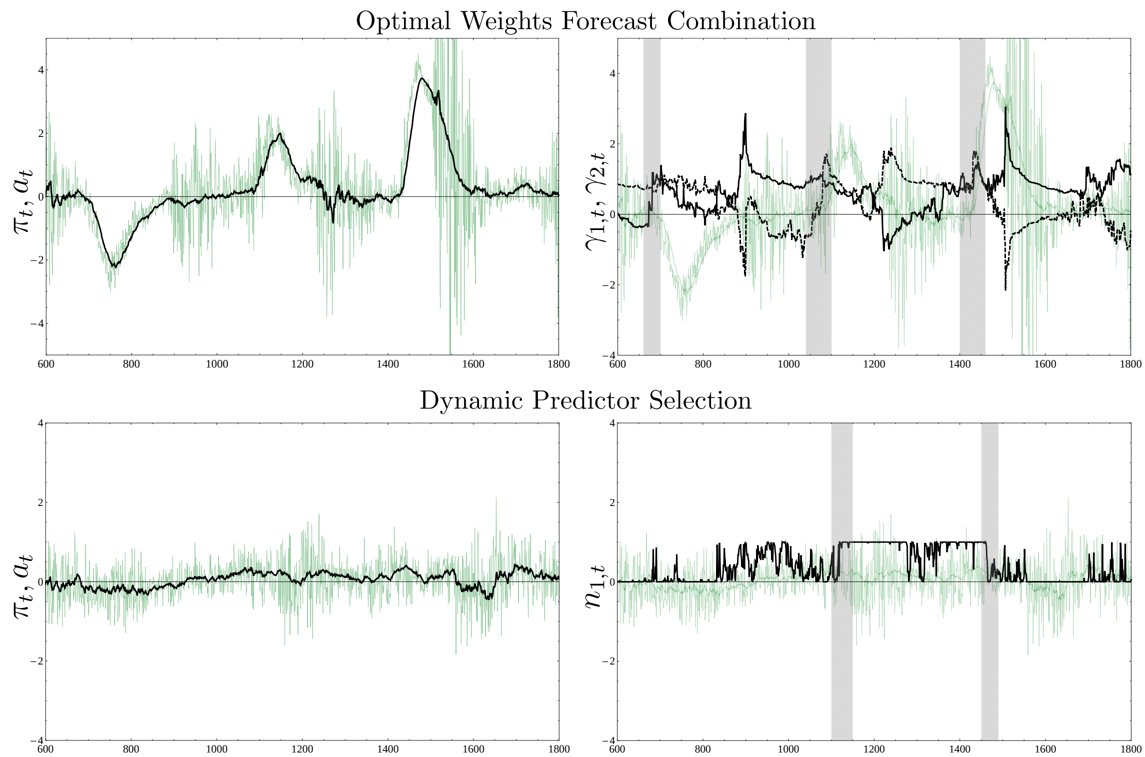
$$\sigma_c^2(\gamma_1) = \gamma_1^2\sigma_1^2 + (1-\gamma_1)^2\sigma_2^2 + 2\gamma_1(1-\gamma_1)\sigma_{12}. \tag{32}$$

Minimizing equation (32) with respect to γ_1 yields the optimal weight

$$\gamma_1^* = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}. \tag{33}$$

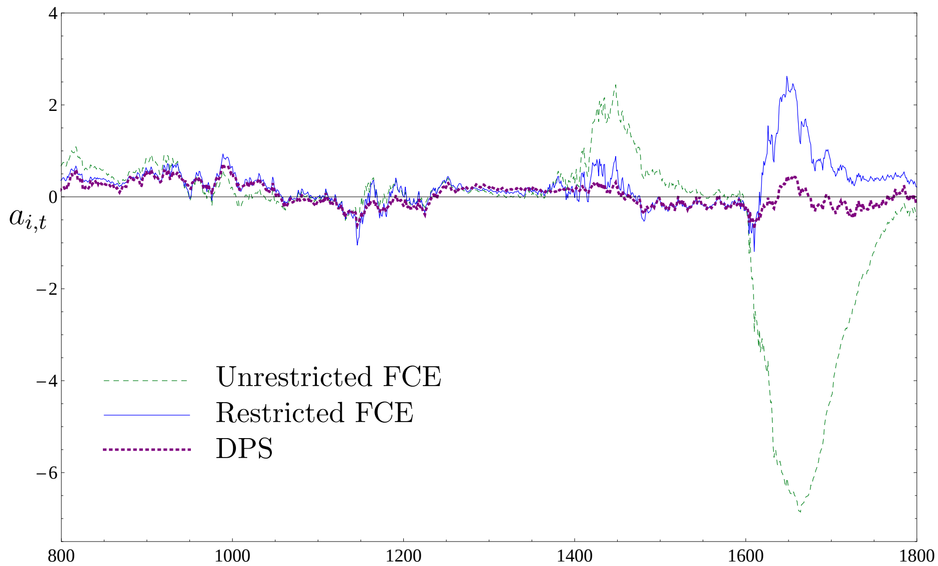
It is simple to show that $\sigma_c^2(\gamma_1^*) \leq \min(\sigma_1^2, \sigma_2^2)$ and that in general the optimal weights need not be in the unit simplex if $\rho_{12} > \sigma_2/\sigma_1$.

Figure 5: FCE and DPS simulations



Notes: Simulated paths of inflation for FCE and DPS. The shared parameter values for the simulations are $\alpha = 0.9$, $\zeta = (.5, .75)'$, $E[x_t x_t'] \sim N(0, 0.1I)$, $w_t \sim N(0, 0.25)$, and a gain of $\kappa = 0.08$. The remaining DPS parameters are $\beta = 50$ and $\lambda = 0.35$. The first column shows the path of inflation and the intercept beliefs $(a_{1,t})$. The second column shows the optimal combination weights and the DPS population weights. The shaded areas indicate periods where beliefs are in the neighborhood of the unstable Fundamental FCE in the forecast combination case and indicate transitions between the two stable equilibria in the DPS case.

Figure 6: Intercepts beliefs



Notes: The intercept beliefs ($a_{i,t}$) from a simulation of inflation in the Lucas monetary model. Parameter values for the simulation are the same as those used in Figure 5.

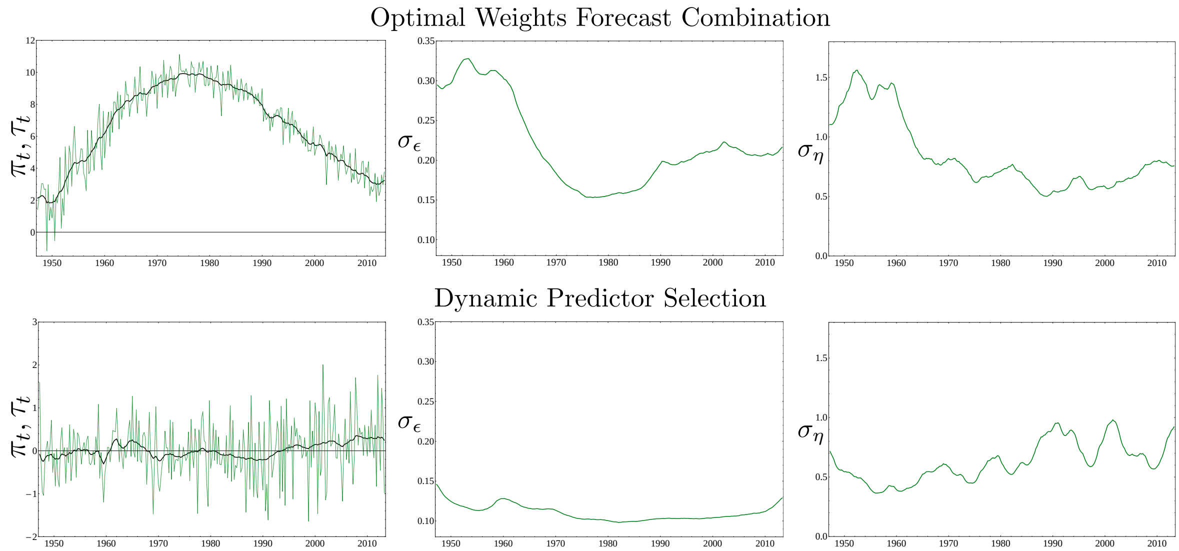
inflation away from steady state. The DPS weights by contrast are constrained to the unit simplex, which rules out these kinds of self-fulfilling beliefs with respect to trend inflation.

Finally, Figure 7 shows the UC-SV estimates for simulated inflation under FCE and DPS. The simulations uses the same parameter values as Figure 5, except for the gain parameter. I follow Branch and Evans (2006b) and set the gain to $\bar{\kappa} = 0.0115$ to calibrate the rate at which agents respond to new information to be consistent with actual forecasting behavior at a monthly frequency.[33] This allows the speed at which agents respond to shocks to be plausible for a comparison to actual data. The economy is simulated for 795 periods and then aggregated by simple averaging to 265 periods to mimic the quarterly observations of the GDP Deflator. The FCE simulation is chosen to include a period with a break in trend inflation to match the break in inflation observed in the actual data.[34] The FCE simulation as expected implies time-varying volatility in both the permanent and transient shocks to inflation. DPS by contrast only replicates the time-varying volatility of the transient shocks to inflation.

[33]Branch and Evans (2006b) find that $\bar{\kappa} = 0.0345$ explains the quarterly forecasts observed in the SPF. I have scaled the number to make it consistent with monthly observations.

[34]The simulated path is a 795 period subsample taken from a larger simulation following a 15,000 period burn-in to illustrate common possible dynamics.

Figure 7: UC-SV estimates of simulated inflation



Notes: Simulated inflation from a Lucas monetary model with forecast combination and Dynamic Predictor Selection. The stochastic trend τ_t, the standard deviation of the permanent shock σ_ϵ, and the standard deviation of the transient shock σ_η are estimated using the UC-SV model described in the text. Parameter values for the simulation are the same as those used in Figure 5 with the exception of the gain. The gain is set to $\kappa = 0.015$. See the text for further explanation.

5 Conclusion

This paper fills a gap in the literature on expectations and bounded rationality by exploring the equilibrium and dynamic consequences of homogeneous agents who employ forecast combination techniques, rather than model selection, to form expectations because of model uncertainty. The consequence of heterogeneous agents who face model uncertainty, but engage in model selection, is well explored in the literature. This paper establishes that despite the similarities of the two approaches to capturing model uncertainty that forecast combination yields distinct equilibrium and dynamic predictions.

The Forecast Combination Equilibrium concept permits multiple equilibria when agents employ a forecast combination strategy to minimize the expected squared forecast error of the combined forecast. The number and stability of the equilibria depend on the menu of forecast rules considered as well as the aggressiveness of monetary policy. In particular, multiple learnable FCEs are shown to exist for common values of the inflation reaction coefficient in the Taylor rule that typically would induce determinacy under rational expectations and expectational stability in wide class of learning models. More aggressive monetary policy is required when forecasts are formed using optimal weights forecast combination to guarantee a unique and expectationally stable equilibrium.

Simulations of inflation dynamics using the equilibrium concept with constant gain learning of parameter beliefs are compared to simulations using Dynamic Predictor Selection. Model uncertainty in both cases endogenously generates time-varying volatility of inflation in a model that only experiences i.i.d. random shocks. However, only forecast combination also generates periodic endogenous movements in trend inflation. The dynamics generated by the forecast combination assumption replicate non-linear features of U.S. inflation dynamics, which illustrates the potential of the concept to explain a range of economic phenomena.

6 Appendix

Selector Matrix Example: The selector matrices given by the u_i's are $m_i \times n+1$ matrices that can be thought of as identity matrices, which have the rows that do not correspond to included exogenous variables deleted. As an example, consider the case where $z_{t-1} = (1, x_{1,t-1}, x_{2,t-1}, x_{3,t-1})'$ and the misspecified model is given by

$$\hat{y}_{1,t} = a_1 + b_1 x_{1,t-1} + b_3 x_{3,t-1}.$$

The model can be written as $\hat{y}_{1,t} = \phi_1' u_1 z_{t-1}$, where $\phi_i = (a_1, b_1, b_3)'$ and

$$u_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Definitions for the Alternative Timing of Expectations: Many DSGE models and asset pricing models have the following reduced form

$$\begin{aligned} y_t &= \mu + \alpha E_t y_{t+1} + \zeta' x_t \\ x_t &= \rho x_{t-1} + \omega_t, \end{aligned}$$

where x_t is vector of mean zero autoregressive processes. The unique REE of this model is

$$y_t = \Omega'(I - \alpha\rho_z)^{-1} z_t,$$

where $\Omega = (\mu, \zeta')'$, ρ_z is square diagonal matrix with the elements of $(1, \rho')$ on the diagonal, and $z_t = (1, x_t')'$. The misspecified forecast rule in this case are given by

$$\hat{y}_{i,t} = \phi_i' u_i z_t$$

which implies

$$E_t^* y_{t+1} = \sum_{i=1}^{k} \phi_i' u_i \rho_z z_t.$$

The corresponding condition for an FCE parameter belief given by Equation (8) is

$$\phi_i = \left[(1-\alpha\gamma_i)\, u_i \Sigma_z \rho_z' u_i'\right]^{-1} (u_i \Sigma_z \Omega + \alpha \sum_{j\neq i} \gamma_j u_i \Sigma_z \rho_z' u_j' \phi_j).$$

The corresponding condition for an optimal weight belief given by Equation (11) is

$$\gamma_i = [\phi_i' u_i \rho_z \Sigma_z \rho_z' (-\alpha\rho_z + I) u_i' \phi_i]^{-1} (\phi_i' u_i \rho_z \Sigma_z \rho_z' (\Omega + \sum_{j\neq i} \gamma_j (\alpha\rho_z - I) u_j' \phi_j)).$$

Recursive Least Squares Learning Example and the Derivation of T-map: For the Lucas-type monetary model example, the agents' recursively estimates three regressions: two regressions to estimate the coefficients of $\hat{\pi}_{1,t}$ and $\hat{\pi}_{2,t}$ and a third to estimate Γ. The estimation can be written jointly and recursively as

$$\begin{aligned} \Theta_t &= \Theta_{t-1} + \kappa_t R_{t-1}^{-1} \mathbf{z}_{t-1} (\mathbf{y}_t - \mathbf{z}_{t-1}' \Theta_{t-1}) \\ R_t &= R_{t-1} + \kappa_t (\mathbf{z}_{t-1} \mathbf{z}_{t-1}' - R_{t-1}), \end{aligned} \tag{34}$$

where the first equations governs the evolution of the belief and weight coefficients, $\Theta_t = (\hat{\phi}_{1,t}', \hat{\phi}_{2,t}', \hat{\Gamma}_t')'$, and the second equation is the estimated second moments matrix. The recursive form uses a block matrix structure to estimate all coefficients simultaneously, where $\mathbf{y}_t = (\pi_t\ \pi_t\ \pi_t)'$, the regressors are stacked into the matrix

$$\mathbf{z}_t = \begin{pmatrix} u_1 z_t & 0 & 0 \\ 0 & u_2 z_t & 0 \\ 0 & 0 & \Pi_{t+1} \end{pmatrix}, \tag{35}$$

and $\Pi_t = (\hat{\pi}_{1,t}, \hat{\pi}_{2,t})'$.

The derivation of the T-map follows Chapter 13 of Evans and Honkapohja (2001) and can be computed by calculating

$$ER^{-1}\mathbf{z}_{t-1}(\mathbf{y}_t - \mathbf{z}_{t-1}\Theta_{t-1}).$$

The nesting of the equilibrium conditions (8) and (11) can be seen by multiplying the $\mathbf{z}_{t-1}$ through the brackets to yield

$$ER^{-1}\begin{pmatrix} u_1 z_{t-1}z'_{t-1}\Omega + \alpha\sum_{i=1}^{2}\gamma_i u_1 z_{t-1}z'_{t-1}u'_i\phi_i - u_1 z_{t-1}z'_{t-1}u'_1\phi_1 \\ u_2 z_{t-1}z'_{t-1}\Omega + \alpha\sum_{i=1}^{2}\gamma_i u_2 z_{t-1}z'_{t-1}u'_i\phi_i - u_2 z_{t-1}z'_{t-1}u'_2\phi_2 \\ Y_t z'_{t-1}\Omega + \alpha Y_t\Gamma' Y_t - Y_t\Gamma' Y_t \end{pmatrix}$$

and then pushing through the expectations operator, rearranging terms, and factoring out $\mathbf{z}_{t-1}\mathbf{z}'_{t-1}$ to arrive at

$$R^{-1}E\mathbf{z}_{t-1}\mathbf{z}'_{t-1}\begin{pmatrix} (u_1\Sigma_z u'_1)^{-1}(u_1\Sigma_z\Omega + \alpha\gamma_2 u_1\Sigma_z u'_2\phi_2) + (\alpha\gamma_1 - 1)\phi_1 \\ (u_2\Sigma_z u'_2)^{-1}(u_2\Sigma_z\Omega + \alpha\gamma_1 u_2\Sigma_z u'_1\phi_1) + (\alpha\gamma_2 - 1)\phi_2 \\ (\phi'_1 u_1\Sigma_z u'_1\phi_1)^{-1}(\phi'_1 u_1\Sigma_z(\Omega + (\alpha-1)\gamma_2 u_2\phi_2)) + (\alpha-1)\gamma_1 \\ (\phi'_2 u_2\Sigma_z u'_2\phi_2)^{-1}(\phi'_2 u_2\Sigma_z(\Omega + (\alpha-1)\gamma_1 u_1\phi_1)) + (\alpha-1)\gamma_2 \end{pmatrix}.$$

The associated differential equation of the system can then be written as

$$\begin{aligned} \dot{\Theta} &= R^{-1}E\mathbf{z}\mathbf{z}'(T(\Theta) - \Theta) \\ \dot{R} &= E\mathbf{z}\mathbf{z}' - R, \end{aligned}$$

where since $R^{-1}E\mathbf{z}\mathbf{z}' = I$ at a fixed point, the appropriate equation to analyze a given equilibrium is

$$\dot{\Theta} = T(\Theta) - \Theta.$$

Proposition 1: To prove existence first note that by construction $u_i u'_i = I$, where I is an $m_i \times m_i$ identity matrix, and consider the following two lemmas:

Lemma 3: If Σ_z is diagonal (A2), then $u_i\Sigma_z^{-1}u'_i u_i\Sigma_z u'_i = I$

Proof: By assumption Σ_z is a diagonal matrix and by construction $u_i'u_i = \Psi$ is a square diagonal matrix with only ones and zeros on the diagonal. Matrix multiplication of diagonal matrices implies that $\Sigma_z^{-1}u_i'u_i\Sigma_z$ can be computed as

$$\begin{aligned} diag(\sigma_{z,1}^{-1}, \sigma_{z,2}^{-1}, ..., \sigma_{z,k}^{-1}) * diag(\psi_1, \psi_2, ..., \psi_k) * diag(\sigma_{z,1}, \sigma_{z,2}, ..., \sigma_{z,k}) \\ = diag(\sigma_{z,1}^{-1}\psi_1\sigma_{z,1}, \sigma_{z,2}^{-1}\psi_2\sigma_{z,2}, ..., \sigma_{z,k}^{-1}\psi_k\sigma_{z,k}) \\ = diag(\psi_1, \psi_2, ..., \psi_k). \end{aligned}$$

Therefore, $u_i\Sigma_z^{-1}u_i'u_i\Sigma_z u_i' = u_i\Psi u_i' = u_iu_i'u_iu_i' = I$. $\square$

Lemma 4: If Σ_z is diagonal (A2), then $(u_i\Sigma_z u_i')^{-1}u_i\Sigma_z = u_i$

Proof: Suppose that $(u_i\Sigma_z u_i')^{-1}u_i\Sigma_z = A$, such that $A \neq u_i$, then

$$\begin{aligned} u_i\Sigma_z &= u_i\Sigma_z u_i'A \\ u_i &= u_i\Sigma_z u_i'A\Sigma_z^{-1} \\ u_iu_i' &= u_i\Sigma_z u_i'A\Sigma_z^{-1}u_i' \\ I &= u_i\Sigma_z u_i'A\Sigma_z^{-1}u_i' \text{ by construction} \\ u_i\Sigma_z^{-1}u_i' &= u_i\Sigma_z^{-1}u_i'u_i\Sigma_z u_iA\Sigma_z^{-1}u_i' \\ u_i\Sigma_z^{-1}u_i' &= A\Sigma_z^{-1}u_i' \text{ by Lemma 2} \\ vec(u_i\Sigma_z^{-1}u_i') &= vec(A\Sigma_z^{-1}u_i') \\ vec(\Sigma_z^{-1})(u_i \otimes u_i) &= vec(\Sigma_z^{-1})(u_i \otimes A) \\ vec(\Sigma_z^{-1})(u_i \otimes (u_i - A)) &= 0 \end{aligned}$$

Thus, either $u_i = 0$, or $A = u_i$ and a contradiction is established. $\square$

Now consider the equation for the equilibrium parameter beliefs given by

$$\phi_i = [(1-\alpha\gamma_i)u_i\Sigma_z u_i']^{-1}(u_i\Sigma_z\Omega + \alpha\sum_{j\neq i}\gamma_j u_i\Sigma_z u_j'\phi_j)$$

using Lemmas 3 and 4 this can be simplified to

$$\phi_i = (1-\alpha\gamma_i)^{-1}(u_i\Omega + \alpha\sum_{j\neq i}\gamma_j A_{i,j}\phi_j),$$

where $u_i u_j' = A_{i,j}$. Now assuming A1, A3, and $\alpha \neq \frac{1}{\omega}$, it follows that $A_{i,j}\phi_j = 0$ for all j in the above sum. To see this, recall that u_i and u_j may only share the first row in common by A3. Therefore, everywhere $A_{i,j}$ has a non-zero value corresponds to the intercept parameter belief. However, due to A1 ($\mu = 0$) and $\alpha \neq \frac{1}{\omega}$, the equilibrium parameter belief must also be zero. Thus, $(1-\alpha)^{-1}u_i\Omega$ and $\Gamma = (1, ..., 1)'$ are EW-FCE beliefs.

Finally, to verify that the EW-FCE is equivalent to the REE, substitute in the equilibrium beliefs, weights, and A3 into the ALM

$$\begin{aligned}
y_t &= (\Omega' + \alpha \sum_{i=1}^{k} (1-\alpha)^{-1}\Omega' u_i' u_i) z_{t-1} + w_t \\
y_t &= (\Omega' + \alpha(1-\alpha)^{-1}\Omega' \begin{bmatrix} \omega & \\ & I \end{bmatrix}) z_{t-1} + w_t.
\end{aligned}$$

Now using A1 such that $\Omega = (0, \zeta')'$, it follows that the above expression simplifies to

$$y_t = (1-\alpha)^{-1}\Omega' z_{t-1} + w_t.$$

To show that proposed FCE is unique it must be the case that

$$\Delta(\Gamma^*) = I - \alpha \sum_{i=1}^{k} \Sigma_z u_i' (u_i \Sigma_z u_i')^{-1} u_i'$$

is invertible. Applying Lemma 4 it follows that

$$\begin{aligned}
\Delta(\Gamma^*)\Sigma_z &= (I - \alpha \sum_{i=1}^{k} \Sigma_z u_i' (u_i \Sigma_z u_i')^{-1} u_i')\Sigma_z \\
\Delta(\Gamma^*)\Sigma_z &= \Sigma_z - \alpha \Sigma_z \sum_{i=1}^{k} u_i' u_i \\
\Delta(\Gamma^*)\Sigma_z &= \Sigma_z - \alpha \Sigma_z \begin{bmatrix} \omega & \\ & I_n \end{bmatrix} \\
\Delta(\Gamma^*) &= (I - \alpha A),
\end{aligned}$$

which is invertible as long as $\alpha \neq 1$ and $\alpha\omega \neq 1$. $\square$

Proposition 2: From the proof of Proposition 1 it follows that

$$\phi_i = (1 - \alpha\gamma_i)^{-1} u_i \Omega$$

and that the beliefs are equivalent to the REE in aggregate. Substituting these beliefs into the optimal weights condition yields

$$\gamma_i(1-\alpha)(1-\alpha\gamma_i)^{-1}\Psi = \Xi + (\alpha - 1)\sum_{j\neq i}\Lambda_j,$$

where $\Psi = \Omega' u_i' u_i \Sigma_z u_i' u_i \Omega$, $\Xi = \Omega' u_i' u_i \Sigma_z \Omega$, and $\Lambda_j = (1-\alpha\gamma_j)^{-1}\gamma_j \Omega' u_i' u_i \Sigma_z u_j' u_j \Omega$. Assuming A1, it follows that intercept belief are zero, which implies $\Lambda_j = 0$ for all $j \neq i$. Now noting that $\Xi\Psi^{-1} = 1$, it follows that

$$\gamma_i(1-\alpha)(1-\alpha\gamma_i)^{-1} = 1,$$

which implies the optimal weight is $\gamma_i = 1$. $\square$

Lemma 1: The methods presented follow Wiggins (1990). A bifurcation may be characterized by deriving an approximation to the center manifold of the dynamic system. The dynamic behavior of the system on the center manifold determines the dynamics in the larger system. To demonstrate the derivation of the center manifold, consider the following dynamic system

$$\dot{x} = Dx \qquad x \in \mathbb{R}^n.$$

The system has n eigenvalues such that $s+c+u=n$, where s is the number of eigenvalues with negative real parts, c is the number of eigenvalues with zero real parts, and u is the number eigenvalues with positive real parts. Suppose that $u=0$, then the system can be written as

$$\begin{aligned}
\dot{x} &= Ax + f(x,y,\epsilon), \\
\dot{y} &= By + g(x,y,\epsilon), \qquad (x,y,\epsilon) \in \mathbb{R}^c \times \mathbb{R}^s \times \mathbb{R}, \\
\dot{\epsilon} &= 0,
\end{aligned} \tag{36}$$

where

$$f(0,0,0)=0, \qquad Df(0,0,0)=0,$$
$$g(0,0,0)=0 \qquad Dg(0,0,0)=0,$$

A and B are diagonal matrices with the corresponding eigenvalues on the diagonal, and $\epsilon \in \mathbb{R}$ is the bifurcation parameter. Suppose that the system has a fixed point at $(0,0,0)$. The center manifold is defined locally as

$$W^c_{loc}(0) = \{(x,y,\epsilon) \in \mathbb{R}^c \times \mathbb{R}^s \times \mathbb{R} \mid y = h(x,\epsilon), |x| < \delta, |\epsilon| < \delta, h(0,0)=0, Dh(0,0)=0\}.$$

The graph of $h(x,\epsilon)$ is invariant under the dynamics generated by the system, which gives the following condition:

$$\dot{y} = D_x h(x,\epsilon)\dot{x} + D_\epsilon h(x,\epsilon)\dot{\epsilon} = Bh(x,\epsilon) + g(x,h(x,\epsilon),\epsilon). \tag{37}$$

The equation can be used to approximate $h(x,\epsilon)$ to form $f(x,h(x,\epsilon),\epsilon)$, which gives the dynamics of the system on the center manifold and allows for a bifurcation to be identified. The conditions for the existence of a supercritical pitchfork bifurcation at (0,0,0) are

$$\begin{array}{ccc} f(0,0,0)=0 & \frac{\partial f}{\partial x}(0,0,0)=0 & \frac{\partial f}{\partial \epsilon}(0,0,0)=0 \\ \frac{\partial^2 f}{\partial x^2}(0,0,0)=0 & \frac{\partial^2 f}{\partial x \partial \epsilon}(0,0,0) \neq 0 & \frac{\partial^3 f}{\partial x^3}(0,0,0) < 0. \end{array}$$

To apply the center manifold reduction technique to the T-map given by equation (20) and (19) it must be put into the normal form of (36). This is done by translating the fixed point $a_i = 0, b_i = \zeta_i/(1-\alpha)$, and $\gamma_i = 1$ for $i = 1,2$ to lie at the origin and finding a linear transformation V to put the eigenvalues of the system into the matrices A and B. Let the translated fixed point at the origin be represented by Θ^*. Then calculating $(DT - I)|_{\Theta^*}$, the Jacobian of the T-map at the fixed point, let V be a linear transformation such that

$$V^{-1}(DT - I|_{\Theta^*})V = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

where A is a zero matrix and B is a diagonal matrix with the stable eigenvalues on the diagonal. The T-map in normal form is thus

$$\begin{aligned}\dot{u} &= Au + f(u, v, \epsilon)\\ \dot{v} &= Bv + g(u, v, \epsilon)\\ \dot{\epsilon} &= 0\end{aligned}$$

where f and g are the terms of order two and higher. Now that it is in normal form, the center manifold can be approximated using a Taylor expansion by taking derivatives of equation (37). The second order approximation of the center manifold is

$$f(u, h(u, \epsilon), \epsilon) \approx 2u\epsilon - \frac{\left(\frac{1}{12} - \frac{i}{6}\right) u^3}{\sigma_2^2 \zeta_2^2} - \frac{\left(\frac{1}{6} - \frac{i}{3}\right) u^3 \epsilon}{\sigma_2^2 \zeta_2^2}.$$

The above equation satisfies the conditions for a supercritical pitchfork bifurcation. The explicit derivation of this center manifold approximation is available on the author's website. □

Proposition 4: The result follows directly from Lemma 1 and Proposition 3. The existence of a supercritical pitchfork bifurcation of the Fundamental OW-FCE steady state at $\alpha = \frac{1}{2}$ implies that steady state is stable under learning for $\alpha < \frac{1}{2}$ and that two new E-stable equilibria have come into existence.

Lemma 2: See Lemma 1 for explanation of the center manifold reduction technique. The second order approximation of center manifold in this case is

$$f(u, h(u, \epsilon), \epsilon) \approx \frac{u\left(4\sigma_1^2 \epsilon \zeta_1^2 \sqrt{-8+\rho_1}\rho_1^2(2+3\rho_1) - u^2(1+2\epsilon)(-2+\rho_1)\left(-2\sqrt{-8+\rho_1} - 10\sqrt{\rho_1} + \sqrt{-8+\rho_1}\rho_1 + \rho_1^{3/2}\right)\right)}{2\sigma_1^2 \zeta_1^2 \sqrt{-8+\rho_1}\rho_1^2(2+3\rho_1)}$$

$f(u, h(u, \epsilon), \epsilon)$ satisfies the conditions for a supercritical pitchfork bifurcation. The explicit derivation of this center manifold approximation is available on the author's website.
□

Proposition 5: The result follows directly from Lemma 2 and Proposition 3. The existence of a supercritical pitchfork bifurcation of the Fundamental OW-FCE steady state at $\alpha = \frac{1}{2}$ implies that steady state is stable under learning for $\alpha < \frac{1}{2}$ and that two new E-stable equilibria have come into existence.

References

Ang, A., G. Bekaert, and M. Wei, "Do macro variables, asset markets, or surveys forecast inflation better?," *Journal of Monetary Economics*, 2007, *54* (4), 1163–1212.

Anufriev, M. and C.H. Hommes, "Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments," *American Economic Journal: Microeconomics*, 2012a, *4* (4), 35–64.

_ **and** _ , "Evolution of market heuristics," *The Knowledge Engineering Review*, 2012b, *27* (02), 255–271.

_ **, T. Assenza, C.H. Hommes, and D. Massaro**, "Interest rate rules and macroeconomic stability under heterogeneous expectations," *Macroeconomic Dynamics*, 2013, *17* (08), 1574–1604.

Atkeson, A. and L.E. Ohanian, "Are Phillips curves useful for forecasting inflation?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 2001, *25* (1), 2–11.

Bates, J.M. and C.W.J. Granger, "The combination of forecasts," *Operations Research*, 1969, pp. 451–468.

Bénassy, Jean-Pascal, "Interest rate rules, inflation and the Taylor principle: an analytical exploration," *Economic Theory*, 2006, *27* (1), 143–162.

Branch, W.A., "Restricted perceptions equilibria and learning in macroeconomics," *Post Walrasian macroeconomics: Beyond the dynamic stochastic general equilibrium model. Cambridge University Press, Cambridge*, 2004.

_ **and B. McGough**, "Replicator dynamics in a cobweb model with rationally heterogeneous expectations," *Journal of Economic Behavior & Organization*, 2008, *65* (2), 224–244.

_ **and** _ , "Dynamic predictor selection in a new keynesian model with heterogeneous expectations," *Journal of Economic Dynamics and Control*, 2010, *34* (8), 1492–1508.

_ **and G.W. Evans**, "Intrinsic heterogeneity in expectation formation," *Journal of Economic Theory*, 2006, *127* (1), 264–295.

_ **and** _ , "A simple recursive forecasting model," *Economics Letters*, 2006b, *91* (2), 158–166.

— **and** — , "Model uncertainty and endogenous volatility," *Review of Economic Dynamics*, 2007, *10* (2), 207–237.

— **and** — , "Monetary policy and heterogeneous expectations," *Economic Theory*, 2011, *47* (2-3), 365–393.

Branch, William A and George W Evans, "Unstable inflation targets," Technical Report, mimeo 2011.

Brock, W.A. and C.H. Hommes, "A rational route to randomness," *Econometrica: Journal of the Econometric Society*, 1997, *65* (5), 1059–1095.

— **and** — , "Heterogeneous beliefs and routes to chaos in a simple asset pricing model," *Journal of Economic dynamics and Control*, 1998, *22* (8-9), 1235–1274.

— **,** — **, and F.O.O. Wagener**, "More hedging instruments may destabilize markets," *Journal of Economic Dynamics and Control*, 2009, *33* (11), 1912–1928.

Bullard, James and Kaushik Mitra, "Learning about monetary policy rules," *Journal of Monetary Economics*, 2002, *49* (6), 1105–1129.

Chiarella, C. and X.Z. He, "Heterogeneous beliefs, risk, and learning in a simple asset-pricing model with a market maker," *Macroeconomic Dynamics*, 2003, *7* (04), 503–536.

Cho, I., N. Williams, and T.J. Sargent, "Escaping nash inflation," *The Review of Economic Studies*, 2002, *69* (1), 1–40.

Clemen, R.T., "Combining forecasts: A review and annotated bibliography," *International Journal of Forecasting*, 1989, *5* (4), 559–583.

Cochrane, John H, "Determinacy and Identification with Taylor Rules," *Journal of Political Economy*, 2011, *119* (3), 565–615.

Evans, George W. and Bruce McGough, "Observability and Equilibrium Selection," *mimeo*, 2015.

Evans, G.W. and S. Honkapohja, *Learning and expectations in macroeconomics*, Princeton University Press, 2001.

— **,** — **, T.J. Sargent, and N. Williams**, "Bayesian Model Averaging, Learning, and Model Selection," in T.J. Sargent and J. Vilmunen, eds., *Macroeconomics at the Service of Public Policy*, Oxford University Press, 2013.

Gliksberg, Baruch, "Monetary policy and multiple equilibria with constrained investment and externalities," *Economic Theory*, 2009, *41* (3), 443–463.

Granger, C.W.J. and R. Ramanathan, "Improved methods of combining forecasts," *Journal of Forecasting*, 1984, *3* (2), 197–204.

Grauwe, Paul De, "Animal spirits and monetary policy," *Economic Theory*, 2011, *47* (2-3), 423–457.

Guse, E.A., "Learning in a misspecified multivariate self-referential linear stochastic model," *Journal of Economic Dynamics and Control*, 2008, *32* (5), 1517–1542.

Harvey, D. and P. Newbold, "Tests for multiple forecast encompassing," *Journal of Applied Econometrics*, 2000, *15* (5), 471–482.

Hendry, D.F. and M.P. Clements, "Pooling of forecasts," *The Econometrics Journal*, 2004, *7* (1), 1–31.

Hommes, C.H., "The heterogeneous expectations hypothesis: Some evidence from the lab," *Journal of Economic Dynamics and Control*, 2011, *35* (1), 1–24.

_ , *Behavioral rationality and heterogeneous expectations in complex economic systems*, Cambridge University Press, 2013.

Lucas, R.E., "Some international evidence on output-inflation tradeoffs," *The American Economic Review*, 1973, *63* (3), 326–334.

McGough, B., "Shocking Escapes," *The Economic Journal*, 2006, *116* (511), 507–528.

Muth, J.F., "Rational expectations and the theory of price movements," *Econometrica: Journal of the Econometric Society*, 1961, *29* (3), 315–335.

Orphanides, A. and J.C. Williams, "The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations," *Journal of Economic dynamics and control*, 2005, *29* (11), 1927–1950.

_ **and** _ , "Monetary policy with imperfect knowledge," *Journal of the European Economic Association*, 2006, *4* (2-3), 366–375.

Preston, Bruce, "Learning about monetary policy rules when long-horizon expectations matter," *International Journal of Central Banking*, 2005.

Robertson, J.C., "Central bank forecasting: an international comparison," *Economic Review-Federal Reserve Bank of Atlanta*, 2000, *85* (2), 21–32.

Sargent, T.J., *The conquest of American inflation*, Princeton Univ Pr, 2001.

Smith, J. and K.F. Wallis, "A simple explanation of the forecast combination puzzle*," *Oxford Bulletin of Economics and Statistics*, 2009, *71* (3), 331–355.

Stock, J.H. and M.W. Watson, "Combination forecasts of output growth in a seven-country data set," *Journal of Forecasting*, 2004, *23* (6), 405–430.

_ **and** _ , "Why has U.S. inflation become harder to forecast?," *Journal of Money, Credit and Banking*, 2007, *39*, 3–33.

Sturmfels, B., *Solving systems of polynomial equations* number 97, American Mathematical Soc., 2002.

Taylor, John B, "Discretion versus policy rules in practice," in "Carnegie-Rochester conference series on public policy," Vol. 39 Elsevier 1993, pp. 195–214.

_ , "A historical analysis of monetary policy rules," in "Monetary policy rules," University of Chicago Press, 1999, pp. 319–348.

Timmermann, A., "Forecast combinations," *Handbook of Economic Forecasting*, 2006, *1*, 135–196.

Wallis, K.F., "Combining forecasts–forty years later," *Applied Financial Economics*, 2011, *21* (1-2), 33–41.

Wieland, Volker and Maik H Wolters, "The diversity of forecasts from macroeconomic models of the US economy," *Economic Theory*, 2011, *47* (2-3), 247–292.

Yang, Y., "Combining forecasting procedures: some theoretical results," *Econometric Theory*, 2004, *20* (01), 176–222.