Forecast Combination in the Macroeconomy

Christopher G. Gibbs*

University of Oregon

July 30, 2013

Abstract

This paper studies forecast combination from a macroeconomic perspective. We introduce the concept of Forecast Combination Equilibrium to model boundedly rational agents who possess a menu of different models to forecast an endogenous state variable. The agents combine the forecasts using one of three optimal combination strategies or simple averaging, known as equal weights. The different equilibrium outcomes are compared to each other and to rational expectations. We find that optimal combination strategies can produce multiple equilibria and endogenous volatility under broad conditions. The equilibrium outcomes of the strategies are an interesting metric to consider when selecting combination strategies for actual forecasting. The standard metric to assess forecast combination strategies, real-time out-of-sample forecasting, does not capture the self-referential effect of forecasting in the macroeconomy.

JEL Classifications: E17; E31; C52; C53

Keywords: Forecast combination, macroeconomics, econometric learning, time-varying volatility, model selection, Lucas Critique.

*Author contact at cgg1127@gmail.com. I would like to thank George Evans, Bruce McGough, Jeremy Piger, and participants of the University of Oregon Macro Group for valuable comments and suggestions. All errors are my own.

1 Introduction

A popular alternative to rational expectations in dynamic macroeconomics is to model agents as econometricians. The approach, known as econometric learning, is commonly used as a stability criterion for rational expectations equilibria and as a selection mechanism for models with multiple equilibria. It is also used as a rational basis to justify boundedly rational economic behavior. Evans and Honkapohja (2013) call the justification the cognitive consistency principle. The cognitive consistency principle states that economic agents should be modeled to be "as smart as (good) economists."

The standard econometric learning approach assumes that agents possess a single subjective forecast model with initially unknown parameters. The agents estimate the unknown parameters given data and forecast recursively, updating their parameter estimates as new data becomes available. If the subjective forecast model is specified correctly, then the recursively formed parameter estimates typically converge to rational expectations.

In the practice of forecasting, however, econometricians often possess a menu of different subjective forecast models. The menu of different models reflects diverse views on the structure of the economy or different misspecifications that must be made to satisfy degrees of freedom restrictions when there exists limited data. An econometrician can select among the models by assigning a fitness measure to each forecast or she can devise a way to combine them. The forecasting literature has studied both forecast selection and forecast combination and it has found that combination is the more robust and efficient solution. Bates and Granger (1969) first demonstrated the forecasting efficiency gains that are possible with forecast combination. The paper spawned an entire subfield of econometrics dedicated to developing new combination techniques and explaining the origins of the results. Surveys of this literature are found in Clemen (1989), Granger (1989), Timmermann (2006), and Wallis (2011).

Despite the dominance of forecast combination in the forecasting literature, theoretical models that have studied agents with a menu of forecasts, overwhelmingly model agents that select, rather than combine forecasts. A brief list of examples are Brock and Hommes

(1997 and 1998), Branch and Evans (2006, 2007, and 2011), Branch and McGough (2008 and 2010), and Gibbs (2012). The agents in these models select forecasts by a process called dynamic predictor selection. The agents use a fitness measure to select models that evolve with the dynamics of the economy. The evolving fitness measure prompts the agents to switch among the forecast models over time. Dynamic predictor selection is often used to motivate heterogeneous expectations and is shown to produce a number of economic phenomena such as multiple equilibria, time-varying volatility, and exotic dynamics, which can match dynamics observed in actual economic data. It is, however, an open question whether these economic phenomena exist if agents choose the more common and robust solution to the model selection problem of forecast combination.

In addition, combined forecast are also the most common way forecasts are widely distributed to the private sector and to policymakers. Prominent examples of widely distributed combined forecasts include the Survey of Professional Forecasters, the Michigan Inflation Expectations Survey, and the Blue Chip Consensus Forecasts, which are all reported as either mean or median combined forecasts. The wide dissemination and use of these combined forecasts are an example of how combined forecasts may have macroeconomic implications. Given these notable examples and embracing the cognitive consistency principle, it is natural to model agents to form combined forecasts.

1.1 Contribution

I study the forecast combination solution to the model selection problem in a macroeconomy. I introduce a formal equilibrium concept called a *Forecast Combination Equilibrium*. 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), and Branch (2004). The equilibrium concept creates a general framework in which any forecast combination strategy can be studied as an expectation formation strategy of agents in a macroeconomic model.

The concept proposes that a continuum of identical agents considers a menu of underparameterized forecast models.[1] The consideration of underparameterized and parsimonious models follows the recommendations of the forecasting literature.[2] The agents combine the forecasts from the menu models to form a single expectation. The identical agents employ a predetermined strategy to combine the forecasts and the resulting equilibrium beliefs are explored. The forecast combination strategies are judged on their ability to obtain or approximate the Rational Expectations Equilibrium (REE) in keeping with the traditional use of learning techniques as selection and robustness measures.

The agents' combination strategy is to form a weighted sum of the menu of forecasts. I study and compare the outcomes of choosing weights using two different forecast combination strategies proposed in the forecasting literature. The two strategies are optimal weights and simple averaging, known as equal weights. Optimal weights provides estimates of the weights that explicitly minimize the expected squared forecast error, while equal weights mimics the most common way forecasts are combined, as in the widely publicized measures mentioned previously. The two different strategies are used to combine the same menu of forecasts. The existence of Forecast Combination Equilibrium are established for the different strategies and studied using econometric learning.

The optimal weights forecast combination strategy is shown to have different equilibrium outcomes depending on the self-referential feedback of the model. In a model with negative feedback, an optimal weights combined forecast results in a unique equilibrium that is stable under learning and for some parameter assumptions, identical to rational expectations. In a model of positive feedback, however, optimal weights can result in multiple equilibria of which one is identical to rational expectation, but is never stable under learning. The multiple equilibria that are stable under learning create an environment where endogenous volatility and other non-standard dynamics occur under constant gain learning, which is similar to what is observed when agents engage in dynamic predictor selection. The result is significant because unlike the majority of dynamic predictor selection models, the agents under optimal weights take into account all available information

[1] The consequences of combining erroneous information as regressors and the contemplation of sunspot solutions is explored in a follow-up paper.

[2] For example see Hendry and Clements (1998).

and there exists no heterogeneity among agents' expectations. The result suggests that forecast uncertainty may be the key driver of the dynamics observed in this literature.

The equal weights forecast combination strategy bounds agents' expectations away from rational expectations in all relevant parameterizations of the model under study. But unlike optimal weight strategies, the equilibrium under equal weights is unique. The deviations from rationality depend again on the type and magnitude of expectational feedback. Under some parameterizations of the model, the equilibrium under equal weights provides predictions that are very similar to rational expectations predictions and for other parameterizations there are significant discrepancies. Both combination strategies considered in the proposed equilibrium concept provide a consistent modeling framework to think about how agents respond to relative changes in information coming from multiple sources and how their subjective or ad hoc weighting of that information effects economic outcomes.

1.2 Forecast Combination Equilibria and Actual Forecasting

The Forecast Combination Equilibrium concept is presented as an open concept in which numerous combination strategies can be applied because there exists no firm consensus in the forecast combination literature on the best combination strategy. The best forecast combination strategies are actually referred to as a puzzle, called the forecast combination puzzle.[3] The strategies that are consistently found to perform best are simple strategies such as averaging forecasts. This result is obtained despite the fact that there is significant time variation in the relative efficiency of popular forecasting models that should be exploitable by more sophisticated combination routines.

The criteria used to evaluate the efficacy of different forecast combination strategies and which generates the forecast combination puzzle is pseudo out-of-sample forecasting efficiency. Pseudo out-of-sample forecasting is an exercise where an existing data set is partitioned into in-sample and out-of-sample subsets. The in-sample subset is used to estimate a menu of forecast models and initialize the combination strategy to recursively

[3] The puzzle was first called "the forecast combination puzzle" by Stock and Watson (2004), but has been noted in the literature by many authors over the last 40 years.

forecast the out-of-sample subset. If a combination strategy forecasts the out-of-sample subset well versus some benchmark, then the strategy is deemed effective.

This method of evaluation and justification for forecast combination strategies has the potential to suffer from a external validity problem along the lines of the Lucas Critique. A main objective of researchers in this field is to publish and widely distribute the forecasting strategies they develop. If a combination strategy were to show a substantial improvement in forecasting efficacy over existing strategies, and it were widely adopted by firms, used to produce forecasts for policymakers, or used to create widely publicized forecasts such that the forecasts influence decision making on a macroeconomic level, then there is reason to believe that the forecast efficiency of that strategy will not continue. A link between the forecasting strategy and the data generating process is created which may render invalid the demonstrated efficacy of a strategy to predict past data. The same way a macroeconomic policy change based on empirical correlations in past data can often fail to have the intended effect as described by Lucas (1976).

This scenario is described by the Forecast Combination Equilibrium concept and it provides a way to model the general equilibrium effects of a widely used forecast combination strategy. Since these strategies are largely atheoretic with respect to economic theory and because the current evaluation method yields a puzzle, the Forecast Combination Equilibrium concept provides another perspective from which to evaluate forecast combination strategies. Granger (1989) and Wallis (2011) both remark that the forecast combination literature is large and repetitive, but important, and this concept offers a new way to design and evaluate strategies.

The remainder of the paper proceeds as follows. Section 2 introduces a general framework and equilibrium concept in which to study forecast combination. Section 3 proposes forecast combination strategies from the forecasting literature to analyze and characterizes the forecast combination equilibria that exist. Section 4 uses econometric learning to study the stability of the Forecast Combination Equilibria when agents estimate parameters and weights in real time. Section 5 demonstrates the time-varying volatility that optimal forecast combination can generate. Section 6 discusses the relationship between endogenous weight forecast combination strategies and the Lucas Critique. Section 7 concludes.

2 A General Framework

To fix ideas I present a macroeconomy that has a unique Rational Expectations Equilibrium in which to study forecast combination. I then propose a plausible way for boundedly rational agents to possess a menu of different forecasts based on standard practices from the forecasting literature. Finally, I present an equilibrium concept to study the properties of different forecast combination strategies when employed by dynamic optimizing agents in place of rational expectations.

2.1 A Reduced Form Economy

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,

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

where y_t is a scalar state variable, x_{t-1} is a $n \times 1$ vector of exogenous and observable shocks, and v_t is white noise.[4] The model is the reduced form version of two well-known macroeconomic models depending on the value of α. The model is the reduced form version of the Muth (1961) cobweb model for $\alpha < 0$ and the Lucas-type aggregate supply model of Lucas (1973) for $0 < \alpha < 1$.

The standard assumption under rational expectations is that agents form expectations using a linear combination of the exogenous observable shocks, which are model consistent. The agents' expectations or forecasts can be represented as

$$E_{t-1} y_t = \phi' z_{t-1}, \tag{2}$$

[4] The model permits many different shock structures such as VAR(p) or VARMA(p,q) processes.

where ϕ is a $(n+1)\times 1$ vector of coefficients that reflect agents beliefs about the effect of the exogenous observables on y_t and $z_{t-1} = (1\ x'_{t-1})'$. The necessary and sufficient condition for the expectation to be rational is that in equilibrium its forecast errors are orthogonal to the agents' information sets,

$$Ez_{t-1}(y_t - \phi' z_{t-1}) = 0, \tag{3}$$

where 0 is an $(n+1)$ vector of zeros. The unique beliefs that satisfy (3) and constitute a rational expectations equilibrium are $\phi = (1-\alpha)^{-1}(\mu\ \ \zeta')'$.

2.2 Misspecified Models and Forecast Combination Equilibria

To study forecast combination I deviate from rational expectations and assume there exists uncertainty over the correct specification to forecast y_t. I assume that agents consider k different underparameterized versions of equation (2) that each omit one or more of the exogenous shocks in x_{t-1}.[5] The k underparameterized models are denoted as $y_{i,t} = \phi'_i z_{i,t-1}$ for $i = 1, 2, ...k$, where ϕ_i and $z_{i,t-1}$ are $m \times 1$ vectors such that $m \leq n$.

The use of underparamerterized and misspecified models mimics standard practices in the forecasting literature. Macroeconomic forecasters typically possess limited data and must make restrictions on the number of parameters that are estimated in any given model. Also, the use of many predictors is found to create estimation uncertainty in the form of model overfitting that reduces out-of-sample forecast accuracy. Empirical examples of the efficacy of using parsimonious forecasts are Ohanian and Atkeson (2001), Ang, Bekaert, and Wei (2007), and Stock and Watson (2004), who show that simple univariate time series models forecast inflation or output better respectively, than more correctly specified and theoretically grounded models. A survey of the literature on forecasting with many predictors is given by Stock and Watson (2006).

The agents, in accordance with the cognitive consistency principle, choose to combine the k different forecasts to create a single forecast of y_t. The agents combine the fore-

[5] In a companion paper with Bruce McGough, I explore the case where agents consider sunspots representations as possible forecast models.

casts using a weighted sum approach that is standard in the forecasting literature. The weighted sum of the k underparameterized model is given by

$$E_{t-1}y_t = \sum_{i=1}^{k} \gamma_i \phi_i' z_{i,t-1}, \tag{4}$$

where $\gamma_i \in \mathbb{R}$ is the weight given to i^{th} model.

I require that in equilibrium the individual forecasts meet a set of optimality conditions similar to the condition presented for rational expectations. I formalize these conditions in a new equilibrium concept called a *Forecast Combination Equilibrium*.

Definition 1: A *Forecast Combination Equilibrium* (FCE) is a set of beliefs $\{\phi_1, \phi_2, ..., \phi_k\}$ that describes a vector of forecasts $Y_t = (y_{1,t}\ y_{2,t}...\ y_{k,t})' \in \mathbb{R}^k$, given weights $\Gamma = (\gamma_1\ \gamma_2...\ \gamma_k)' \in \mathbb{R}^k$, such that $E_{t-1}y_t = \sum_{i=1}^{k} \gamma_i y_{i,t}$ and

$$E z_{i,t-1}(y_t - \phi_i' z_{i,t-1}) = 0 \tag{5}$$

for all $i = 1, 2, ..., k$.

The general definition requires agents to consider only plausible models in equilibrium, where plausible is defined as the conditional expected forecast error of each model is zero. The general definition also leave the selection of weights as exogenous. The selection of weights is a non-trivial problem in the empirical practice of forecasting and I am interested in studying modifications from this basic definition to explore how exogenous and endogenous selection of weights alters equilibrium outcomes.

A Forecast Combination Equilibrium is a natural extension of the Restricted Perceptions Equilibrium (RPE) concept. In an RPE the agents are required to have an optimal forecast given their restricted information set. In an FCE the agents have a similar restriction that is model specific. Each individual forecast model is required to be optimal given the information set used to create it. The individual forecasts are, however, not necessarily optimal given the total information set of the agents. The definition reflects the behavior of forecasters in the real world, who optimally fit different misspecified models conditional on their included information and then combine the forecasts.

The equilibrium definition is also related to the Misspecification Equilibrium (ME) concept developed in Branch and Evans (2006). They analyze heterogeneous agents that select forecasts from a list of misspecified models using a fitness measure. The aggregate forecast in the economy is the weighted average of the different forecasts chosen by the agents with the weights equal to the measure of agents that chose each forecast. An ME requires in equilibrium that the individual models satisfy the same orthogonality condition given in Definition 1. The Forecast Combination Equilibrium concept is distinct from ME because there is no heterogeneity in forecast choice. The agents form homogeneous expectations by incorporating the entire menu of forecast to form a single prediction.

2.2.1 Existence of an FCE

I begin my analysis of forecast combination by establishing the conditions that must be met for an FCE to exist given an exogenous vector of weights. Suppose that the agents possess a menu of misspecified forecasts $Y_t = (y_{1,t} \; y_{2,t} \; ... \; y_{k,t})' \in \mathbb{R}^k$ and they choose to combine them to form a single forecast using the weights $\Gamma = (\gamma_1 \; \gamma_2 \; ... \; \gamma_k)' \in \mathbb{R}^k$. The expectation of the agents, $E_{t-1}y_t$, is given by equation (4) and the economy under the combined forecasts can be written as

$$y_t = \mu + \alpha \sum_{i=1}^{k} \gamma_i \phi_i' z_{i,t-1} + \zeta' x_{t-1} + v_t. \tag{6}$$

The economy is said to be in a Forecast Combination Equilibrium if given the weights Γ, the individually misspecified forecasts are individually optimal in accordance with equation (15) given in Definition 1. This implies the beliefs of the agents represented by the ϕ_i's must satisfy the following system of equations

$$\begin{aligned} Ez_{1,t-1}(\mu+\alpha\sum_{i=1}^{k}\gamma_i\phi_i' z_{i,t-1}+\zeta' x_{t-1}+v_t-\phi_1' z_{1,t-1}) &= 0 \\ &... \\ Ez_{k,t-1}(\mu+\alpha\sum_{i=1}^{k}\gamma_i\phi_i' z_{i,t-1}+\zeta' x_{t-1}+v_t-\phi_k' z_{k,t-1}) &= 0. \end{aligned} \tag{7}$$

There exists a unique FCE given the following condition is satisfied:

Existence Condition: Given Y_t and Γ, a unique FCE exists if $det(\Delta) \neq 0$, where

$$\Delta = \begin{bmatrix} (1-\alpha\gamma_1)(u_1\Sigma_z u_1') & -\alpha\gamma_1 u_1\Sigma_z u_2' & ... & -\alpha\gamma_1 u_1\Sigma_z u_k' \\ -\alpha\gamma_2 u_2\Sigma_z u_1' & (1-\alpha\gamma_2)(u_2\Sigma_z u_2') & ... & -\alpha\gamma_2 u_2\Sigma_z u_k' \\ ... & ... & ... & ... \\ -\alpha\gamma_k u_k\Sigma_z u_1' & -\alpha\gamma_k u_k\Sigma_z u_2' & ... & (1-\alpha\gamma_k)(u_k\Sigma_z u_k') \end{bmatrix},$$

$Ez_{t-1}z_{t-1}' = \Sigma_z$, and u_i is an $m \times (n+1)$ sector matrix that selects the appropriate elements out of Σ_z that correspond to the i^{th} underparameterization.[6]

The existence condition is derived in detail in the appendix.

If there exists an FCE for a given Γ and menu of forecast Y_t, then in general there exists and open set U of weight vectors, such that $\Gamma \in U$, which will also satisfy the existence condition for the same Y_t. This follows from the fact that the eigenvalues of Δ are a continuous function of Γ. The existence of multiple weights that constitute and FCE for a given Y_t implies that there is not a unique weight vector for the agents to choose. The next section studies different possible ways agents may choose weights based on recommendations from the forecasting literature.

[6]The condition is a necessary, but not sufficient condition for an ME in Branch and Evans (2006).

3 Exogenous and Endogenous Selection of Weights

In this section I consider the possibility that agents either exogenously impose recommended weights or endogenously choose weights according to an optimality criterion to form combined forecasts. I restrict my analysis to the homogeneous selection of weights by all agents to characterize the possible equilibrium outcomes when agents coordinate on single combination strategy. The analysis of the homogeneous case is sufficiently complicated to leave the questions of heterogeneity to future research.

The goal of forecast combination is to choose weights to create the optimal combined forecast. I employ the standard definition of optimal used in the forecasting literature, which is that an optimal forecast minimizes the expected squared error of a forecast.[7] Given a set of underparmeterized models Y_t, the forecast combination problem is

$$\min_{\{\Gamma\}} E[(y_t - \Gamma' Y_t)^2]. \tag{8}$$

Endogenous solutions to (8) are given by Granger and Ramanathan (1984). They propose three different ordinary least squares regressions that estimate the optimal solution to (8), given past data on y_t and Y_t. The three different specifications are justified by the statistical properties of the forecasts Y_t. I consider all three strategies as ways agents can endogenously select weights.

The three specifications Granger and Ramanathan propose are

$$y_t = \gamma_1 y_{1,t} + \gamma_2 y_{2,t} + ... + \gamma_k y_{k,t} + e_t \tag{9}$$

$$y_t = \gamma_1 y_{1,t} + \gamma_2 y_{2,t} + ... + \gamma_k y_{k,t} + e_t \ : \ s.t. \sum_{i=1}^{k} \gamma_i = 1 \tag{10}$$

$$y_t = \gamma_0 + \gamma_1 y_{1,t} + \gamma_2 y_{2,t} + ... + \gamma_k y_{k,t} + e_t, \tag{11}$$

where e_t represents the error term. The regression coefficients of each model are explicit solutions to (8). The first specification, equation (9), is argued by Granger and

[7] The effect of using other metrics, such as asymmetric loss functions, is an interesting avenue for future research.

Ramanathan to be the optimal solution when the menu of forecasts is believed to be unbiased. I call this specification optimal weights (OW).

The two other specifications are modifications to the OW case. The second specification imposes the restriction that the weights sum to one. I call this case restricted optimal weights (ROW). The restriction is argued to guarantee that the combined forecast of unbiased forecasts is unbiased. The restriction is also argued to ensure that the combined forecast optimally uses the available information. Diebold (1988) shows that combined forecasts that do not impose this restriction can generate serially correlated forecast errors in out-of-sample forecasting exercises.

The third specification adds a constant term to the optimal weights (OWC). The constant term is used to remove any bias that may exist in the elements of Y_t from the combined forecast. The addition of a constant to the weights regression is not a trivial modification to the OW case because in forecasting, parsimony is key. The estimation error introduced by the addition of extra parameters when there exists limited data can reduce forecast efficiency.[8] The modification is also not an obvious addition based on the objective function given by (8). Although, it is a more natural modification when considered in a regression framework.

A natural exogenous weights choice is for agents to consider equal weights. The equal weights solutions is given by $\gamma_i = 1/k$ for $i = 1, 2, ..., k$. This specification is only a solution to (8) under very specific conditions, but as mentioned in the introduction, it is found to work well in practice. Equal weights also serve as a good comparison to the optimal weights because it does not require knowledge of the distribution of y_t or the vector of forecasts Y_t to implement.[9]

[8] A specific example of this appear in Smith and Wallis (2009) who show that estimation uncertainty is one explanation for the forecast combination puzzle.

[9] A popular forecast combination strategy not studied in this paper is to weight forecasts by the inverse of their past mean squared error measured over a rolling window. The weights in this case are not explicit solutions, but, like equal weights, they are found to be effective. For empirical examples see Bates and Granger (1969) or Stock and Watson (2004).

3.0.2 Assessing FCEs

The four different strategies are used to modify the FCE definition to incorporate the selection of weights as an equilibrium condition. Rational expectations is the natural benchmark for our characterization of these different types of FCEs. The resulting equilibria are compared to an REE in four categories:

1. equilibrium differences in beliefs

2. equilibrium differences in forecasts

3. stability under learning

4. and dynamics under real-time learning.

The first two categories address a forecast combination strategy's ability to approximate rational expectations in equilibrium. I capture these categories in a new definition.

Definition 2: An FCE $\{\phi_1, \phi_2, ..., \phi_k\}$ is called a *fundamental FCE* if the individual model beliefs $\phi_i = (a_i \; b_i')'$ are equivalent to the REE beliefs, such that $a_i = (1-\alpha)^{-1}\mu$ and $b_i = (1-\alpha)^{-1}(\zeta_{i,1} \; \zeta_{i,2} \; ... \; \zeta_{i,m-1})'$ for $i = 1, 2, ...k$ and $E_{t-1}^{REE} y_t = E_{t-1}^{FCE} y_t$.

The notation $E_{t-1}^{REE} y_t = E_{t-1}^{FCE} y_t$ denotes equivalence between the equilibrium forecasts. This condition is necessary because equal beliefs in general do not imply equivalent forecasts or vice versa because the combination weights affect the equilibrium expectations.

The third category assesses the likelihood an FCE is an actual outcome when agents must infer their beliefs from past data. The fourth category assesses the dynamics on the off equilibrium paths when agents form forecasts recursively using real-time econometric learning. The off equilibrium paths under learning can sometimes diverge far from the equilibrium dynamics when there exists unobservable stochastic shocks. The comparisons are made with respect to the value of α assumed in the model. The value of α determines the type of economic model represented by equation (1) by determining the amount feedback a forecast has on the actual realization of the data. If $\alpha = 0$, the model has no self-referential component and forecasting is reduced to a purely statistical exercise.

3.1 Characterizing FCEs

This section characterizes the possible FCEs under each of the four proposed combination strategies for an example of the reduced form economy and a specific menu of forecast models. I consider an economy driven by a 2×1 vector x_{t-1} of exogenous and observable shocks. The shocks are assumed i.i.d. with $Ex_{i,t-1} = 0$ for $i = 1, 2$ and

$$Ex'_{t-1}x_{t-1} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}. \tag{12}$$

The results are not dependent on this assumption, but the simple structure simplifies analysis to better illustrate the intuition.

The agents' menu of forecasts consists of all non-trivial underparameterizations of the data generating process. The menu of forecasts is

$$y_{1,t} = a_1 + b_1 x_{1,t-1} \tag{13}$$

$$y_{2,t} = a_2 + b_2 x_{2,t-1}., \tag{14}$$

which can be express as $y_{i,t} = \phi_i' z_{i,t-1}$ with beliefs $\phi_i = (a_i \ b_i)'$ for $i = 1, 2$. The inclusion of all non-trivial underparameterizations make the agents' information sets equivalent to the information set under rational expectations. This assumptions allows for a precise characterization of the difference between an FCE and the REE.

3.1.1 Equal Weights

The equal weights solution requires the least amount of information for the agents to impose. The equilibrium outcomes provide an illustrative example of how forecast combination alters equilibrium beliefs and forecasts.

Definition 3: An Equal Weights Forecast Combination Equilibrium (EWFCE) is a set of belief $\{\phi_1, \phi_2\}$ that describes a vector of forecasts $Y_t = (y_{1,t} \ y_{2,t})'$, given weights

$\Gamma = (\frac{1}{2}\ \frac{1}{2})'$, such that $E_{t-1}y_t = \sum_{i=1}^{k} \gamma_i y_{i,t}$ and

$$Ez_{i,t-1}(y_t - \phi_i' z_{i,t-1}) = 0 \tag{15}$$

for all $i = 1, 2$.

The set of beliefs that constitute an EWFCE for the model under consideration are ϕ_1 and ϕ_2 that satisfy

$$\begin{aligned} Ez_{1,t-1}(y_t - \phi_1' z_{1,t-1}) &= 0 \\ Ez_{2,t-1}(y_t - \phi_2' z_{2,t-1}) &= 0. \end{aligned} \tag{16}$$

These conditions can be represented as a projected T-map. A projected T-map is a mapping from the individual beliefs to the actual outcomes of the economy under forecast combination. The mapping from beliefs to outcomes is a useful representation to calculate equilibrium beliefs and is the key to analyzing the stability of any equilibria under real-time econometric learning. The projected T-map is also equivalent to constructing the Δ matrix to establish existence of an FCE discussed in Section 2. I translate the conditions in each of the four cases into a projected T-map to solve for the equilibrium beliefs.

A projected T-map is constructed by specifying the agents' perceived law of motion (PLM) for the economy. The PLM under forecast combination is the combined forecast given the appropriate weights,

$$E_{t-1}y_t = \frac{1}{2}\phi_1' z_{1,t-1} + \frac{1}{2}\phi_2' z_{2,t-1}. \tag{17}$$

The PLM represents how agents form $E_{t-1}y_t$ in equation (1). The PLM can be substituted in for $E_{t-1}y_t$ to produce the actual law of motion (ALM) of the economy,

$$y_t = \mu + \alpha(\frac{1}{2}\phi_1' z_{1,t-1} + \frac{1}{2}\phi_2' z_{2,t-1}) + \zeta' x_{t-1} + v_t. \tag{18}$$

The ALM describes how y_t evolves given agents beliefs and their forecast combination strategy. The ALM is then substituted into the orthogonality conditions, the expectation is taken, and the conditions are simplified so that ϕ_1 and ϕ_2 appear on the right-hand side of the equations and a function of ϕ_1 and ϕ_2 are on the left-hand side. The left-hand side of the equations is the projected T-map. The T-map under equal weights is

$$T\begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \mu + \frac{\alpha}{2}(a_1 + a_2) \\ \mu + \frac{\alpha}{2}(a_1 + a_2) \\ \frac{\alpha}{2}(b_1 + b_2\frac{\sigma_{12}}{\sigma_1^2}) + \zeta_1 \\ \frac{\alpha}{2}(b_2 + b_1\frac{\sigma_{12}}{\sigma_2^2}) + \zeta_2 \end{pmatrix}, \tag{19}$$

where $\zeta = (\zeta_1\ \zeta_2)'$. The fixed points of the projected T-map correspond to EWFCEs of the economy.

Lemma 1: There exists a unique fixed point given by

$$\begin{aligned} a_1 = a_2 &= \frac{\mu}{1-\alpha} \\ b_1 &= \frac{(\frac{\alpha}{2}-1)\zeta_1\sigma_1^2\sigma_2^2 - \frac{\alpha}{2}\zeta_1\sigma_{12}^2 - \zeta_2\sigma_{12}\sigma_2^2}{\frac{1}{4}\sigma_{12}^2\alpha^2 - (1-\frac{\alpha}{2})^2\sigma_1^2\sigma_2^2} \\ b_2 &= \frac{(\frac{\alpha}{2}-1)\zeta_2\sigma_1^2\sigma_2^2 - \frac{\alpha}{2}\zeta_2\sigma_{12}^2 - \zeta_1\sigma_{12}\sigma_1^2}{\frac{1}{4}\sigma_{12}^2\alpha^2 - (1-\frac{\alpha}{2})^2\sigma_1^2\sigma_2^2}, \end{aligned} \tag{20}$$

which is not a fundamental FCE.

The lemma is obtained by solving for a fixed point of the T-map and by comparing the resulting beliefs to rational expectations.

The rational expectations beliefs for the given model are $a = \frac{\mu}{1-\alpha}$, $b_1 = \frac{\zeta_1}{1-\alpha}$, and $b_2 = \frac{\zeta_2}{1-\alpha}$. The difference in the beliefs between the EWFCE and REE are from the misspecification of the two underparameterized models and the interaction of the forecast combination strategy in a self-referential environment. The misspecification error is an omitted variable. The bias is captured by the σ_{12} terms in the b_i beliefs. If $\sigma_{12} = 0$, then the omitted variable bias is removed and the EWFCE beliefs collapse to $b_i = \frac{\zeta_i}{1-\alpha/2}$

for $i = 1, 2$, where the remaining difference is due to the combination strategy and the feedback from expectations. The use of equal weights prevents the agents from fully responding to a predicted change from one of the individual models. The attenuated response to the prediction impacts the actual realization of y_t when $\alpha \neq 0$. This alters the actual relationship between x_{t-1} and y_t in equilibrium, which is reflected by the agents beliefs in the EWFCE.

In addition, the equilibrium forecasts provided by the FCE do not equal the REE forecasts. A way to characterize this difference is to calculate the expected squared forecast error in the REE compared to the EWFCE. The expected squared forecast error in the REE is $Ev_t^2 = \sigma_v^2$ and the expected squared in the EWFCE is

$$E(y_t - \frac{1}{2}\sum_{i=1}^{2} y_{i,t})^2 = \sigma_v^2 + \xi_1\sigma_1^2 + \xi_2\sigma_2^2 + \xi_1\xi_2\sigma_{12}, \tag{21}$$

where $\xi_i = (\frac{1}{2}(\alpha - 1)b_i + \zeta_i)$ and b_i is the EWFCE belief given previously for $i = 1, 2$. The equal weights forecast has a higher expected squared forecast error than under rational expectations in equilibrium.

3.1.2 Optimal Weights

The optimal weights case uses the regression specification (9) to form optimal weights for the menu of forecasts. The regression specification can be translated into an extra orthogonality condition that must be satisfied in equilibrium. I formalize these conditions into a new definition.

Definition 4: An Optimal Weights Forecast Combination Equilibrium (OWFCE) is a set of beliefs and weights $\{\phi_1, \phi_2, \Gamma\}$ such that $E_{t-1}y_t = \sum_{i=1}^{2} \gamma_1 y_{i,t}$and

$$\begin{aligned} EY_t(y_t - \Gamma' Y_t) &= 0 \\ Ez_{1,t-1}(y_t - \phi_1' z_{1,t-1}) &= 0 \\ Ez_{2,t-1}(y_t - \phi_2' z_{2,t-1}) &= 0, \end{aligned} \tag{22}$$

where the 0's are 2×1 vectors of zeros.

To study OWFCEs I again translate the equilibrium conditions into a projected T-map. The PLM under optimal weights is

$$E_{t-1}y_t = \gamma_1 \phi_1' z_{1,t-1} + \gamma_2 \phi_1' z_{2,t-1}, \tag{23}$$

and the corresponding ALM is

$$y_t = \mu + \alpha(\gamma_1 \phi_1' z_{1,t-1} + \gamma_2 \phi_1' z_{2,t-1}) + \zeta' x_{t-1} + v_t. \tag{24}$$

Substituting the ALM into the conditions of Definition 4, simplifying the system with respect to each coefficient in $\{\phi_1, \phi_2, \Gamma\}$, and taking expectations gives the following projected T-map

$$T\begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \mu + \alpha(a_1\gamma_1 + a_2\gamma_2) \\ \mu + \alpha(a_1\gamma_1 + a_2\gamma_2) \\ \alpha\gamma_1 b_1 + (\alpha\gamma_2 + \zeta_2)\frac{\sigma_{12}}{\sigma_1^2} + \zeta_1 \\ \alpha\gamma_2 b_2 + (\alpha\gamma_1 + \zeta_1)\frac{\sigma_{12}}{\sigma_2^2} + \zeta_2 \\ \frac{\alpha\gamma_1 a_1^2 + b_1(\alpha\gamma_1\sigma_1^2 b_1 + (\alpha-1)\gamma_2\sigma_{12}b_2 + \sigma_1\zeta_1 + \sigma_{12}\zeta_2) + a_1((\alpha-1)\gamma_2 a_2 + \mu)}{(a_1^2 + \sigma_1^2 b_1^2)} \\ \frac{(\alpha-1)(\gamma_1 a_1 a_2 + \gamma_1\sigma_{12}b_1 b_2) + \alpha\gamma_2 a_2^2 + \alpha\gamma_2\sigma_2 b_2^2 + \zeta_1\sigma_{12}b_2 + \zeta_2\sigma_2^2 b_2 + \mu a_2}{(a_2^2 + \sigma_2^2 b_2^2)} \end{pmatrix}. \tag{25}$$

The fixed points of the T-map are OWFCEs of the economy.

The T-map is a system of polynomial equations which suggest there may exist multiple OWFCEs. Obtaining analytic solutions to systems of polynomial equations is notoriously difficult (see Sturnfels (2002)). It is possible to solve this system explicitly with the aid of computers, but the solutions are large and impractical to study. Instead of solving for the entire family of solutions, I analyze the T-map using bifurcation theory to characterize OWFCEs in the neighborhood of rational expectations. The possible equilibria that exist away from the rational expectations are explored numerically.

Lemma 2: If $\sigma_{12} = 0$ and $\mu = 0$, then there exists an OWFCE that is a fundamental FCE with optimal weight $\Gamma = (1\ 1)'$.

The lemma can be established by substituting in the appropriate values into the T-map and checking that it is a fixed point. The requirement that $\sigma_{12} = 0$ is necessary to prevent omitted variable bias in the individual agents beliefs and $\mu = 0$ is required so that optimal weights are correctly specified. The condition that σ_{12} and μ equal zero does not remove the non-linearity from the T-map, so there may exist non-fundamental OWFCEs as well as the fundamental OWFCE in some cases. The existence of non-fundamental OWFCEs implies that rational expectations may only be one of many equilibrium outcomes under optimal weights.

The existence of non-fundamental OWFCEs can be established by monitoring the properties of the fundamental OWFCE as a bifurcation parameter is varied. The existence of a bifurcation can precisely characterize the existence of non-fundamental OWFCEs without having to explicitly solve for them. The natural parameter to study is α, the feedback parameter on expectations, which captures the self-referential element of the model.

To apply bifurcation theory, consider the T-map as a differential equation given by

$$\dot{\Theta} = T(\Theta) - \Theta. \tag{26}$$

The differential equation governs the dynamics of $\Theta = (\phi_1'\ \phi_2'\ \Gamma')'$ in notional time.[10] Bifurcation theory characterizes the existence of OWFCEs by monitoring the stability of a fixed point. If the fixed point of the system has eigenvalues that are equal to zero for some value of α, it may indicate that new fixed points have come into existence by way of a bifurcation.

The fundamental FCE experiences a bifurcation at $\alpha = \frac{1}{2}$. The type of bifurcation and its effect on the fundamental FCE is analyzed by using the center manifold reduction

[10] Notional time is used to distinguish the treatment of the T-map as a differential equation from the actual timing of outcomes in the model.

technique described in Wiggins (1990). The center manifold reduction creates a one-dimensional projection of the bifurcation that fully characterizes the existence of the fixed points in the larger system.

Theorem 1: Given $\mu = 0$ and $\sigma_{12} = 0$, there coexists non-fundamental and fundamental OWFCEs for some $\alpha > \frac{1}{2}$.

The theorem is proved by showing the existence of a pitchfork bifurcation. A pitch fork bifurcation is where a single fixed point destabilizes and creates two new stable OWFCEs. The approximate center manifold is given in Figure (1).

Theorem 1 is a surprising result. It says that optimal weights lives up to its moniker for negative or small positive values of α and can provide equilibrium outcomes that are equivalent to rational expectations. But, as α becomes large, optimal weights can also provide equilibrium outcomes that diverge from rational expectations.[11]

There is an economic explanation for the existence of the non-fundamental OWFCEs when the feedback parameter on expectations is positive. A positive feedback parameter creates a self-fulfilling quality to expectations. The degree to which a forecast has an affect on y_t is determined by the sign of α. In the case where α is negative any beliefs that deviate from rational expectations will result in poor forecasts because y_t will move in the opposite direction of the forecast. In the case where α is positive any beliefs that deviate from rational expectations will be partly confirmed because y_t will move in the same direction as the forecast. The self-fulfilling quality allows the non-fundamental beliefs to interact with the weights to create combinations that constitute an OWFCE.

To illustrate the multiple OWFCEs that exist and their dependence on α, I create a type of pseudo-bifurcation diagram. The diagram is constructed by numerically solving for the entire set of OWFCE that exists for a fixed parameterization of the model and

[11]non-rational or non-fundamental beliefs are associated with stable dynamics in notional time. As indicated when the T-map was introduced, the mapping is key for analyzing the stability of equilibira under real-time econometric learning and the theorem suggests that although the fundamental FCE always exists, it may not always be stable under learning. Unfortunately Theorem 1 does not constitute a proof because the specific technique used is not equivalent to studying the equilibrium under econometric learning, but the intuition is shown to be correct in Section 4.

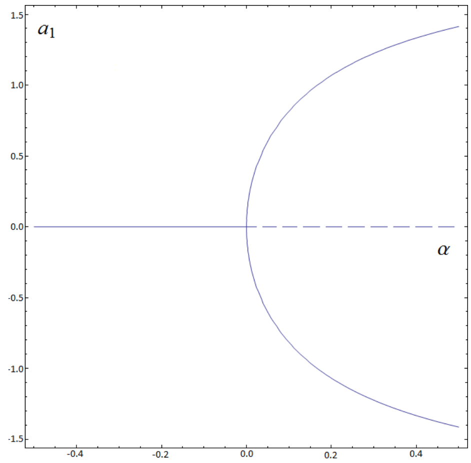

Figure 1: A pitchfork bifurcation on the approximate center manifold of the T-map (26). The bifurcation has been normalized to occur at $(0, 0)$. The solid line indicate stable fixed points and the dashed line indicates unstable fixed points.

a given value of α. The value of α is then incrementally changed and another set of OWFCE is calculated. The resulting beliefs and weights that characterize the OWFCEs for each α are used to construct forecasts for a given specific realization of x_{t-1}. The forecasts are plotted against the corresponding α to produce a diagram that illustrates the dependence of equilibrium beliefs on α.

Figure 2 is the pseudo-bifurcation diagram for α between -1.5 and 1.5 with parameters $\zeta_1 = .9$, $\zeta_2 = -.9$, $\sigma_1^2 = \sigma_2^2 = 1$, $\sigma_{12} = 0$, $\mu = 0$, and $\hat{x}_{t-1} = (1\ 1)'$ as the specific realization of x_{t-1}. The range of α is chosen to cover the relevant regions of the parameter space that are typically explored in the literature.[12] The REE forecast under the given parameters for the fixed $\hat{x}_{t-1}$ is $E_{t-1}^{REE}[y_t|\hat{x}_{t-1}] = 0$ for all α. The simulation shows that the fundamental FCE is the unique OWFCE before the bifurcation and is one of many after the bifurcation. The system also bifurcates a second time at $\alpha = \frac{3}{4}$, which results in six OWFCEs existing simultaneously in addition to the OWFCE that is the fundamental OWFCE.

[12]The majority of papers look at α between -3 and 1. The action in FCEs is in the positive feedback case, so I restrict the size of α on the left. Some examples from the literature are Brock and Hommes (1997) who use $\alpha = -2.7$, Branch and Evans (2006) who use $\alpha = -2$, Gibbs (2012) who uses $\alpha = -1.5$, and Branch and Evans (2007) who use $\alpha = 0.6$.

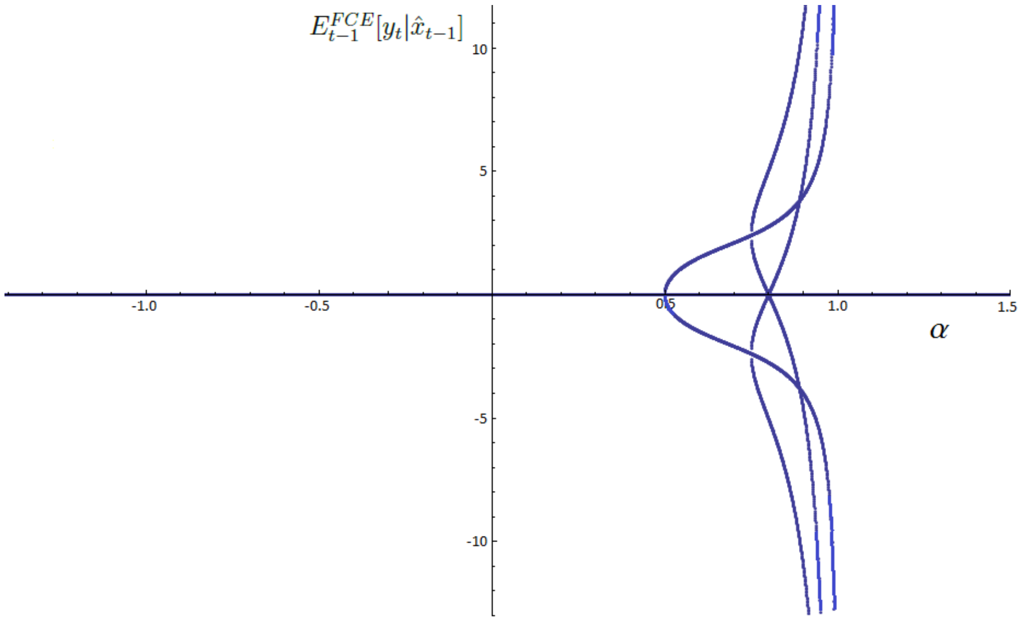

Figure 2: A plot of the OWFCE forecasts for a specific realization of x_{t-1} and for different values of α.

3.1.3 Further Exploration

No fundamental OWFCE exists when either $\sigma_{12} \neq 0$ or $\mu \neq 0$. Figure 3 illustrates the deviations in the forecasts from rational expectations by plotting the equilibrium forecasts for OWFCEs with $\mu = 1$ and the remaining parameters the same as Figure 2. The figure shows the OWFCEs forecast in black and the REE forecast in gray. The deviation between the OWFCE nearest the REE is driven by the weights. This OWFCE has equilibrium beliefs equivalent to rational expectations, but optimal weights $\Gamma = (\frac{\zeta_1^2+\alpha\mu^2}{\zeta_1^2+\mu^2}\ \frac{\zeta_2^2+\alpha\mu^2}{\zeta_2^2+\mu^2})'$. The weights provide a combined forecast that is not equivalent to the rational expectations forecast.

The genesis of the deviation is due to the forecast combination strategy. The strategy is misspecified along the line considered by Granger and Ramanathan (1984). The individual forecasts have positive intercepts that are not correctly accounted for by the optimal weights specification. The specification error is also affected by the value of α, but it is not the main cause of the deviation. Note that if $\alpha = 0$, the equilibrium weights will still not equal $\Gamma = (1\ 1)'$, which is required for the OWFCE forecast to equal the REE forecast.

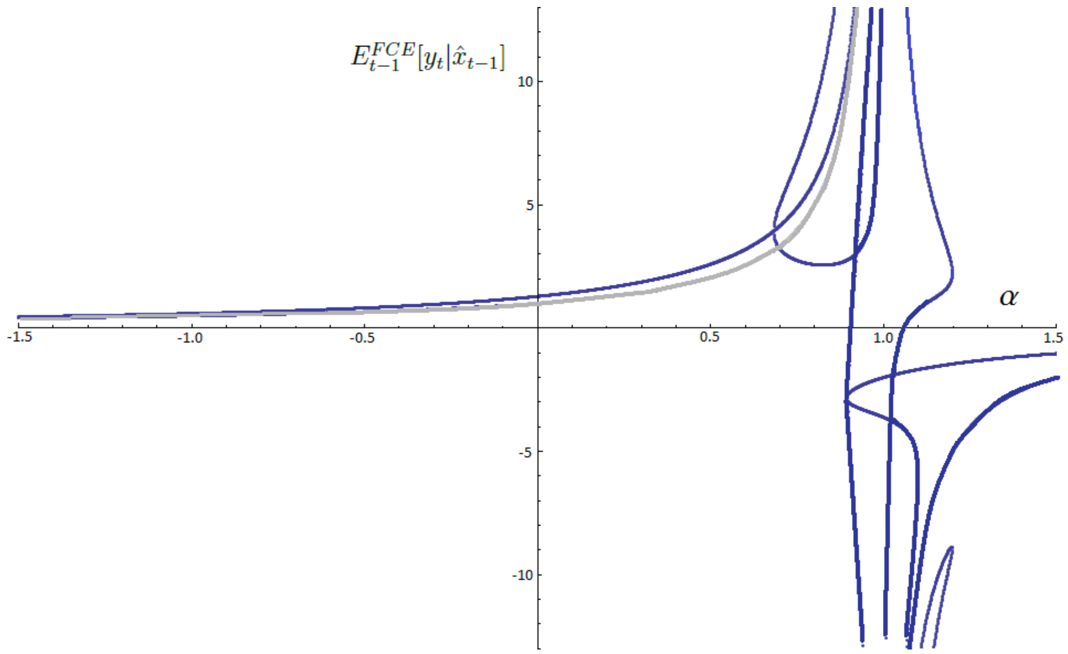

Figure 3: A plot of the OWFCE forecasts (black) and the REE forecast (gray) for a specific realization of x_{t-1} and for different values of α.

The addition of a positive intercept also alters the second bifurcation of the system. The equilibrium forecasts for a relatively large positive α change compared to Figure 2. There now exists multiple OWFCEs for $\alpha > 1$. The maximum number of OWFCEs remains at seven.

3.1.4 Restricted Optimal Weights

The second specification proposed by Granger and Ramanathan (1984) imposes the restriction that the weights sum to one. The purpose of the restriction is to ensure that the combined forecast of unbiased forecasts is unbiased. An FCE with restricted optimal weights can fit into Definition 4 by imposing the restriction that $\gamma_2 = 1 - \gamma_1$ on the first orthogonality condition to yield

$$E(y_{1,t} - y_{2,t})[(y_t - y_{2,t}) - \gamma_1(y_{1,t} - y_{2,t})]. \tag{27}$$

The FCE under restricted regression weights will be referred to as ROWFCE. The system can be represented by a T-map following the same procedure executed under optimal weights. The T-map for ROW is

$$T\begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \\ \gamma_1 \end{pmatrix} = \begin{pmatrix} \mu + \alpha\gamma_1 a_1 + \alpha(1-\gamma_1)a_2 \\ \mu + \alpha\gamma_1 a_1 + \alpha(1-\gamma_1)a_2 \\ \alpha\gamma_1 b_1 + (\zeta_2 + \alpha(1-\gamma_1)b_2)\frac{\sigma_{12}}{\sigma_1^2} + \zeta_1 \\ \alpha(1-\gamma_1)b_2 + (\zeta_1 + \alpha\gamma_1 b_1)\frac{\sigma_{12}}{\sigma_2^2} + \zeta_2 \\ \frac{b_1\Omega_1 - b_2\Omega_2 + (a_1-a_2)\Omega_3}{a_1^2 - 2a_1a_2 + a_2^2 + \sigma_1^2 b_1^2 - 2b_1b_2 + \sigma_2^2 b_2^2} \end{pmatrix} \tag{28}$$

where Ω_1, Ω_2, and Ω_3 are expanded in the footnote.[13]

The possible FCEs under restricted optimal weights are similar to the OW case with a unique FCE that experience a bifurcation to create multiple FCEs. Figure 4 plots the ROWFCE forecasts given by $E_{t-1}^{FCE}[y_t|\hat{x}_{t-1}]$ for $\alpha \in (-1.5, 1.5)$ using identical parameters to Figure 2, but with $\mu = 1$, and $\sigma_{12} = .1$. The ROWFCEs are dissimilar to the OW case because the restriction prevents the existence of a fundamental FCE.

Lemma 3: There does not exist a fundamental FCE in the set of ROWFCEs.

The lemma is obtained by substituting in the REE beliefs into the T-map to verify that they are not a fixed point for any γ_1.

The combined forecast under restricted optimal weights also does not provide a forecast equivalent to rational expectations. Figure 4 in the bottom panel shows the ROWFCE forecasts and the REE forecasts for $\zeta = (.9 \;\; -.65)'$, $\sigma_1^2 = \sigma_2^2 = 1$, $\sigma_{12} = 0$, and $\mu = 0$. In this case the forecast diverges from the rational expectations forecast even when the intercept term and the covariance of the shocks are zero. The restriction that the weights sum to one restricts the possible beliefs and forecasts from ever being equivalent to rational expectations.

3.1.5 Optimal Weights with a Constant

The third solution to the forecast combination problem offered by Granger and Ramanathan (1984) is to add a constant parameter to the weights. The constant is added

[13] $\Omega_1 = \sigma_1^2(\alpha\gamma_1 b_1 + \zeta_1) + \sigma_{12}(b_2(\alpha - \alpha\gamma_1 - 1) + \zeta_2)$, $\Omega_2 = \sigma_{12}(\alpha\gamma_1 b_1 + \zeta_1) + \sigma_2^2(b_2(\alpha - \alpha\gamma_1 - 1) + \zeta_2)$, and $\Omega_3 = \mu + \alpha\gamma_1 a_1 + a_2(\alpha - \alpha\gamma_1 - 1)$.

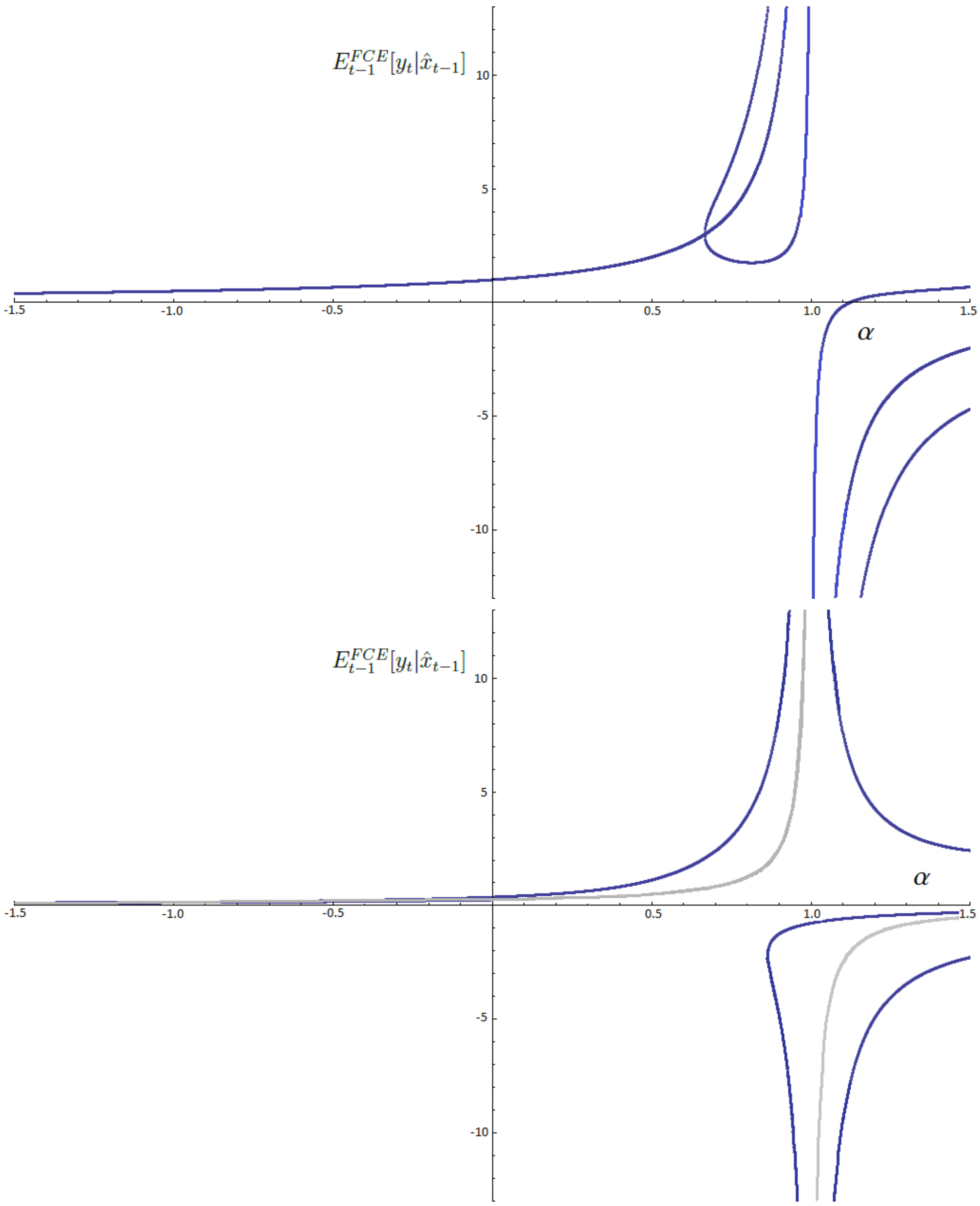

Figure 4: Plots of the ROWFCE forecasts. The top plot demonstrates multiple ROWFCEs for the same parameter values used in Figure (2), but with $\mu = 1$ and $\sigma_{12} = 0.1$. The bottom plot compares ROWFCE forecasts (black) to the REE forecast (gray) for different values of α and $\zeta = (.9 \;\; -.65)'$, $\sigma_1^2 = \sigma_2^2 = 1$, $\sigma_{12} = 0$, and $\mu = 0$.

to offset biases that may exist in the forecasts contained in Y_t. A forecast combination equilibrium under optimal weights with a constant can be fit into Definition 4 by redefining Γ to include an intercept such that $\Gamma = (\gamma_0\ \gamma_1\ \gamma_2)'$. The FCE under optimal weights will be referred to as OWCFCE. The transformed equilibrium conditions can then be represented as a projected T-map following the same procedure executed for optimal weights. The T-map for OWC is

$$T\begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \\ \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \mu + \alpha(\gamma_0 + \gamma_1 a_1 + \gamma_2 a_2) \\ \mu + \alpha(\gamma_0 + \gamma_1 a_1 + \gamma_2 a_2) \\ \alpha\gamma_1 b_1 + (\alpha\gamma_2 b_2 + \zeta_2)\frac{\sigma_{12}}{\sigma_1^2} + \zeta_1 \\ \alpha\gamma_2 b_2 + (\alpha\gamma_1 b_1 + \zeta_1)\frac{\sigma_{12}}{\sigma_2^2} + \zeta_2 \\ \mu - \gamma_1 a_1 - \gamma_2 a_2 + \alpha(\gamma_0 + \gamma_1 a_1 + \gamma_2 a_2) \\ \frac{\alpha\gamma_1 a_1^2 + b_1(\alpha\gamma_1\sigma_1^2 b_1 + (\alpha-1)\gamma_2\sigma_{12}b_2 + \sigma_1^2\zeta_1 + \sigma_{12}\zeta_2) + a_1((\alpha-1)(\gamma_0+\gamma_2 a_2)+\mu)}{a_1^2 + b_1\sigma_1^2} \\ \frac{\alpha\gamma_2 a_2^2 + b_2((\alpha-1)\sigma_{12}b_1 + \alpha\gamma_2\sigma_2 b_2 + \sigma_{12}\zeta_1 + \sigma_2\zeta_2) + a_2((\alpha-1)(\gamma_0+\gamma_1 a_1)+\mu)}{a_2^2 + \sigma_2^2 b_2^2} \end{pmatrix}. \tag{29}$$

Lemma 4: There exists a unique OWCFCE given by

$$\begin{aligned} a_1 = a_2 &= \frac{\mu}{1-\alpha} & \gamma_0 &= \frac{(\sigma_1^2\sigma_2^2 - (\sigma_{12})^2)\zeta_1\zeta_2\mu}{(\alpha-1)(\sigma_1^2\zeta_1 + \sigma_{12}\zeta_2)(\sigma_{12}\zeta_1 + \sigma_2^2\zeta_2)} \\ b_1 &= \frac{\sigma_1^2\zeta_1 + \sigma_{12}\zeta_2}{\sigma_1^2(1-\alpha)} & \gamma_1 &= \frac{\sigma_1^2\zeta_1}{\sigma_1^2\zeta_1 + \sigma_{12}\zeta_2} \\ b_2 &= \frac{\sigma_{12}\zeta_1 + \sigma_2^2\zeta_2}{\sigma_2^2(1-\alpha)} & \gamma_2 &= \frac{\sigma_2^2\zeta_2}{\sigma_2^2\zeta_2 + \sigma_{12}\zeta_1} \end{aligned}$$

and if $\sigma_{12} = 0$, then the OWCFCE is the fundamental FCE with weights $\Gamma = (\frac{\mu}{\alpha-1}\ 1\ 1)'$.

The result provides intuition for why there exist multiple FCEs under OW or ROW. The addition of the intercept term correctly specifies the weights for all possible parameterizations of the economy. The correct specification eliminates the possibility of a self-reinforcing bias originating in the misspecified models. For example, note that the non-fundamental forecasts in Figure 2 are non-zero. The forecasts deviate from zero because one or both of the individual models posits, incorrectly, a positive (negative) value for the intercept. The incorrect belief of a positive (negative) value for the intercept exists because of the interaction between the weights and the expectational feedback, which biases the agents' beliefs compared to rational expectations. The addition of the intercept, γ_0, eliminates the possibility of a sustained bias in the combined forecast from this interaction and results in a unique equilibrium.

3.2 Results Summary

The findings of the section are summarized in Table 1. The table is organized by forecast combination strategies and by the feedback parameter to illustrate the different outcomes that occur for positive, negative, and zero feedback. The table denotes the number of possible equilibria, whether one of the forecast combination equilibrium is equivalent to REE as indicated by the existence of a fundamental FCE, and the conditions under which the fundamental FCE occurs.

There are two important results shown in this section. The first is that optimal weighting strategies can provide equilibrium outcomes equivalent to rational expectations under certain conditions. This shows that optimally combined underparameterized models can be a route to rational expectations. The second result is that optimal combination strategies can deviate from rational expectations in surprising ways, in that, there coexists multiple FCEs under OW and ROW in the positive feedback case. The existence of the multiple FCEs suggest that although rational expectations outcomes are possible, they may not be likely.

The deviations from rational expectations also show that the recommendations from the out-of-sample forecasting literature do not always carry over into the self-referential environment. The recommendations of Granger and Ramanathan are consistent with the multiple FCE case of OW, when σ_{12} and μ equal zero, and consistent for many of the ROWFCEs, which prove never to be fundamental. The next section determines whether the FCEs found under the four solutions to the forecast combination problem are learnable if agents estimate the weights and beliefs in real time.

		Number of FCE	Fundamental FCE	Fundamental Condition
EW	$\alpha = 0$	1	no	-
	$\alpha < 0$	1	no	-
	$\alpha > 0$	1	no	-
OW	$\alpha = 0$	1	yes	$\sigma_{12} = 0$, & $\mu = 0$
	$\alpha < 0$	1	yes	$\sigma_{12} = 0$, & $\mu = 0$
	$\alpha > 0$	7	yes (1/7)	$\sigma_{12} = 0$, & $\mu = 0$
ROW	$\alpha = 0$	1	no	-
	$\alpha < 0$	1	no	-
	$\alpha > 0$	3	no	-
OWC	$\alpha = 0$	1	yes	$\sigma_{12} = 0$
	$\alpha < 0$	1	yes	$\sigma_{12} = 0$
	$\alpha > 0$	1	yes	$\sigma_{12} = 0$

FCE Representative Results Summary

Table 1: Tabulated representative results for the FCEs under equal weights (EW), optimal weights (OW), restricted optimal weights (ROW), and optimal weights with a constant (OWC). The Fundamental FCE column denotes existence. The notation (1/7) indicates that only one of the 7 OWFCEs is a fundamental FCE. The Condition column gives the necessary condition for the existence of fundamental FCE result to be obtained. A dash indicates that there is no broad or economically significant restriction.

4 Learning FCEs

This section assesses the stability of the FCEs identified in Section 3 under recursive least squares learning following the method of Evans and Honkapohja (2001). The FCE concept follows the cognitive consistency principle which makes econometric estimation the natural way agents would form forecasts and combination weights. In real time, the agents are executing the econometric procedures recursively by estimating models on existing data, forming an expectation, and then interacting in the economy to form a new data point. The econometric learning analysis acts an equilibrium selection mechanism by characterizing the likelihood of convergence to a given FCE from nearby initial beliefs.

4.1 E-stability

The agents form their estimates of belief and weights using recursive least squares learning. The estimation of multiple individual models and combination weights requires the use of the Seemingly Unrelated Regression (SUR) method of estimation. The SUR method allows the agents' estimation strategy to be written in a way that standard learning results can be applied. The ability of the agents' estimation strategy to be written in this form is an advantage of the optimal weights strategies studied in this paper. In related work by Evans et al. (2012), the agents use Bayesian model averaging, which is found to not be emendable to standard learning analysis.

The SUR is written recursively as

$$\begin{aligned} \Theta_t &= \Theta_{t-1} + \kappa_t R_t^{-1} \mathbf{z_{t-1}}(\mathbf{y_t} - \mathbf{z'_{t-1}}\Theta_{t-1}) \\ R_t &= R_{t-1} + \kappa_t(\mathbf{z_{t-1}}\mathbf{z'_{t-1}} - R_{t-1}), \end{aligned} \tag{30}$$

where the first equations governs the evolution of the belief and weight coefficients, the second equation is the estimated second moments matrix, and κ_t is the gain sequence that governs the weight given to new observations. To estimate the SUR under optimal weights (OW), the agents stack three copies of y_t into the vector $\mathbf{y_t} = (y_t \; y_t \; y_t)'$ and stack the regressors into the matrix

$$\mathbf{z_t} = \begin{pmatrix} z_{1,t} & 0 & 0 \\ 0 & z_{2,t} & 0 \\ 0 & 0 & Y_{t+1} \end{pmatrix}, \tag{31}$$

where the zeros are 2×1 vectors of zero so that $\mathbf{z_t}$ is a 6×3 matrix. The other forecast combination techniques can fit into this form by making the appropriate changes to y_t and z_t.

The possible rest points of (30) are equivalent to the FCEs determined in Section 3. The stability of these FCEs are determined by appealing to the E-stability principle.[14] The E-stability principle states that the stability of a rest point of (30) is governed by the stability of an associated differential equation. The associated differential equation is determined by fixing the parameter Θ and taking the limit of the expected values of

[14]Guse (2008) explores using SUR to analyze RPE under E-stability. Guse shows that the E-stability results can be applied directly to SUR.

(30) as t goes to infinity.[15] The resulting system is

$$\begin{aligned} \frac{d\Theta}{d\tau} &= R^{-1}E\mathbf{z}\mathbf{z}'(T(\Theta) - \Theta) \\ \frac{dR}{d\tau} &= E\mathbf{z}\mathbf{z}' - R. \end{aligned} \tag{32}$$

The stability of an FCE under the econometric learning process is determined by

$$\frac{d\Theta}{d\tau} = T(\Theta) - \Theta, \tag{33}$$

which is the same differential equation studied in Section 3, where $T(\Theta)$ is the appropriate T-map derived previously. The stability of this equation evaluated at fixed point governs the stability of (30). The condition for stability is that the Jacobian of the T-map, evaluated at an FCE, has eigenvalues with real parts less than one. The condition for the stability of the REE when agents learn using a single, correctly specified model given by equation (2) is $\alpha < 1$.

4.2 Equal Weights

The Equal Weights FCE can easily be characterized analytically. Stability under learning of EWFCE requires the same condition as stability of the REE when agent learn using a correctly specified model in the standard learning analysis.

Theorem 2: The EWFCE is E-stable if $\alpha < 1$, $\alpha < \frac{2}{\rho+1}$, and $\frac{2}{1-\rho}$ where ρ is the correlation between $x_{1,t-1}$ and $x_{2,t-1}$.

The binding condition of Theorem 2 is $\alpha < 1$. Although, it is worth noting that if the agents did not include intercepts in their misspecified forecast models, then the conditions for E-stability would be relaxed. This reflects the dampening effect the weights have on agents' beliefs that was discussed in Section 3.

[15]See Evans and Honkapohja (2001) for a more detailed explanation.

4.3 Optimal Weighting Strategies

Analytic characterizations of the E-stability conditions for the optimal weights cases is difficult because the eigenvalues that indicate stability are large polynomials. I study E-stability in these cases by providing analytic results for special cases that are tractable and use numerical simulation to show evidence that the results generalize.

In the OW case, Theorem 1 strongly suggests that after the bifurcation the two non-fundamental OWFCEs that come into existence are stable under learning and that the fundamental OWFCE is unstable. Unfortunately, Theorem 1 does not provide a proof because it required the system to be reduced to a smaller dimension than needed to determine stability. However, if the bifurcation in the larger system shares the same stability properties as the smaller system, then the fundamental OWFCE should be stable under learning before the bifurcation and unstable after.

Theorem 3: The fundamental OWFCE is E-stable if $\alpha < 1/2$.

Theorem 3 shows that the fundamental steady state behaves as expected, which suggests that the two non-fundamental equilibria are stable under least squares learning. This result is confirmed numerically.

For simplicity I illustrate the E-stability of the different OWFCEs and ROWFCEs by modifying the pseudo-bifurcation diagrams. Figures 2, 3, and 4 are replicated in Figure 5 with solid lines corresponding to FCEs that are stable under learning and dashed lines corresponding to FCEs that are not stable. The upper right plot of Figure 5 shows the result predicted by Theorem 3 with the fundamental FCE destabilizing at $\alpha = \frac{1}{2}$. The intuition from the bifurcation analysis is also seen in the ROW case. The unique FCE destabilizes and the system bifurcates to produce two new stable equilibria.

The FCE under OWC is also not tractable analytically and is analyzed using numeric simulation. A numerical investigation of the parameter space shows that the E-stability condition can vary from $0 < \alpha < 1$, depending on the value of the intercept, the correlation between the exogenous shocks, and ζ. Figure 6 plots a grid of the parameter space

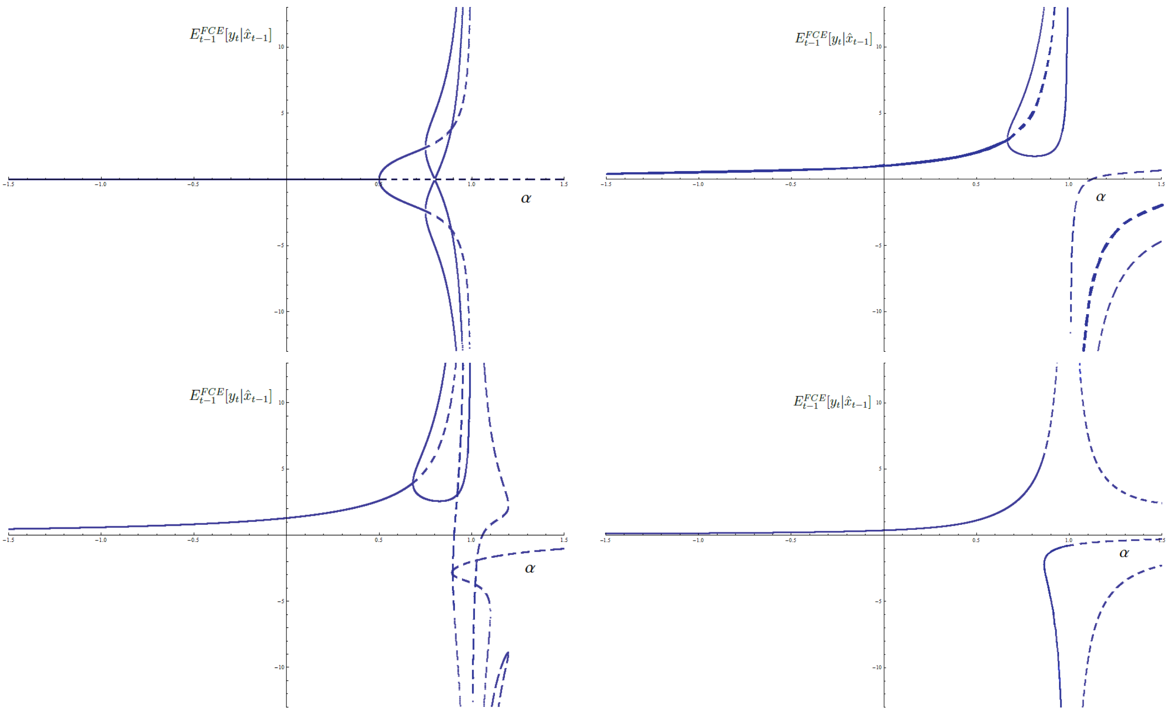

Figure 5: The $E_{t-1}^{FCE}[y_t|\hat{x}_{t-1}]$ under OW (left) and ROW (right) for different values of α. E-stability is indicated by solid lines and E-instability is indicated by dashed lines.

for α, σ_{12}, and μ with the remaining parameters $\zeta = (.9 \ \ -.9)'$, $\sigma_1^2 = \sigma_2^2 = 0$. The light regions of the figure indicates E-stability and the dark regions indicates E-instability. The figures demonstrates a non-linear relationship between E-stability and the parameters of the model.

4.4 Discussion

The E-stability results are summarized in Table 2. The E-stability analysis reveals that the fundamental FCEs under OW have stricter conditions for E-stability than rational expectations. However, many of the non-fundamental OWFCEs that exist are learnable for the same parameters as rational expectations. The non-fundamental OWFCEs are the equilibria the economy obtains when α is large and positive, despite the existence of a fundamental OWFCE.

The OWC case provides a unique equilibrium, but the equilibrium is not stable over the same parameter space as rational expectations. The E-stability of the OWCFCE varies non-linearly with parameters of the model and in many cases the E-stability con-

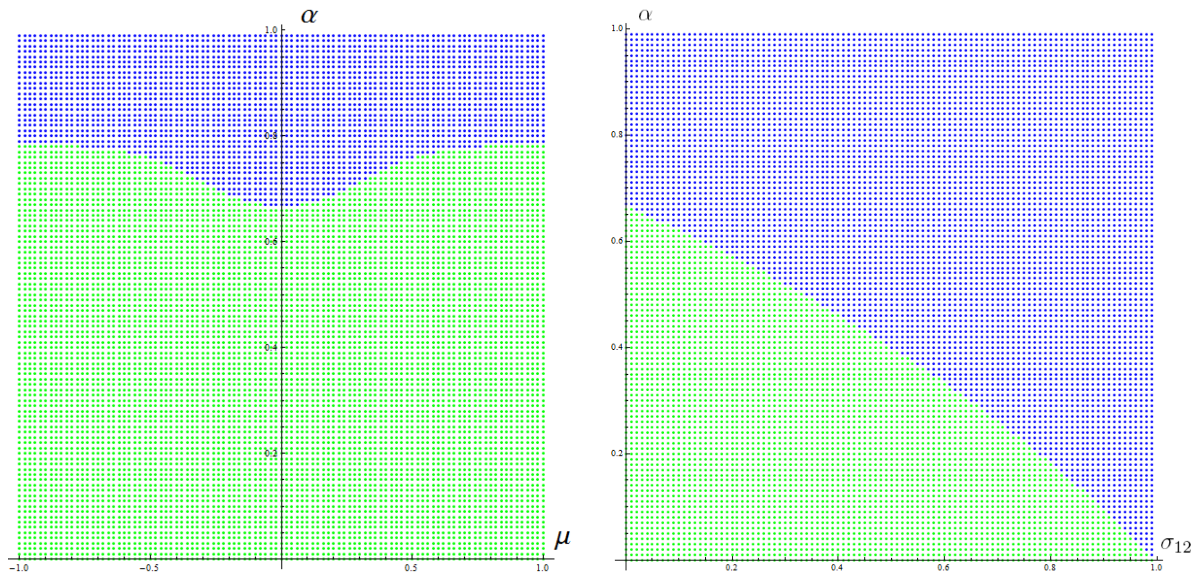

Figure 6: E-stability of the fundamental FCE under OWC for pairs of parameters (α, σ_{12}), and (α, μ). The light portion corresponds to the parameter space that is E-stable.

dition for α is far lower than is traditionally found in the literature. The result implies that even a correct optimal forecast combination specification does not guarantee rational expectations under learning.

The best combination strategy compared to rational expectations is equal weights. Equal weights always results in a unique and learnable equilibrium over the same parameter space as rational expectations. This is in contrast to OW, which has as many as four non-fundamental coexisting and stable equilibria, or ROW, which has two stable and non-fundamental coexisting equilibria. It is also in contrast to the OWCFCE, which is not learnable at all for portions of the parameter space.

5 Learning in Real Time

The last metric to assess the different forecast combination strategies is to analyze the dynamics generated by the strategies under econometric learning. The result from the dynamic predictor selection literature is that when agents use constant gain learning the economy can experience time-varying volatility as the economy transitions endogenously between equilibria as shown in Branch and Evans (2007). I demonstrate that this behavior also occurs under certain conditions when agents use optimal weights to combine forecasts.

E-stability Results

	Fundamental FCE	Non-Fundamental	Coexisting	Condition
EW	-	$\alpha < 1$	-	-
OW	$\alpha < \frac{1}{2}$	$\alpha < 1$	4	$\mu = 0$
	-	$\alpha < 1$	4	$\mu \neq 0$
ROW	-	$\alpha < 1$	2	-
OWC	$\alpha < \frac{2}{3}$	$\alpha \lesssim 0.8^{*}$	-	$\mu = 0$
	$\alpha \lesssim 1^{**}$	$\alpha \lesssim 1^{**}$	-	$\mu \neq 0$

Table 2: Tabulated E-stability results for the FCEs under equal weights (EW), optimal weights (OW), restricted optimal weights (ROW), and optimal weights with a constant (OWC). The Coexisting column indicates the maximum number of stable FCEs that exist for the stated conditions in the Non-Fundamental column. *The E-stability condition on α is a non-linear function of σ_{12} and ζ that ranges between $(0, 0.8)$ (see Figure (6)). **The E-stability condition is a non-linear function of σ_{12}, ζ, and μ that ranges between $(0, 1)$.

Constant gain learning is used to model agents that are concerned about structural breaks as argued in Orphanides and Williams (2006) and Branch and Evans (2006b and 2007). Constant gain learning assumes that agents place more weight on new information when forming their parameter estimates. The placement of higher weight on new observation can cause agents' expectations to drift in response to the random shocks in the economy. When there exists multiple FCEs the drifting causes the economy to transition from one stable FCE to another.

5.1 Endogenous Volatility

The endogenous volatility is driven by the existence of multiple equilibria when agents use OW or ROW with constant gain learning. The following simulation assumes agent use OW to form combined forecasts. Similar results are obtained using the ROW strategy. The time-varying volatility that may occur alters both the mean and variance of y_t after an endogenous break. The agents estimate beliefs and weights in real time using the SUR recursive formula given by (30).

The simulation is conducted with parameters $\alpha = .9$, $\mu = 0$, $\zeta = (.9\ .9)'$, $v_t \sim N(0, 1)$,

$\sigma_1^2 = \sigma_2^2 = 1$, $\sigma_{12} = 0$, and a gain parameter of $\kappa = 0.05$.[16] There are four E-stable OWFCEs under this parameterization of the model. The OWFCEs are

$$\Theta_1 = \begin{pmatrix} 2.846 \\ 1.014 \\ 2.846 \\ 7.985 \\ 0.125 \\ 0.985 \end{pmatrix}, \Theta_2 = \begin{pmatrix} -2.846 \\ 1.014 \\ 2.846 \\ 7.985 \\ 0.125 \\ 0.985 \end{pmatrix}, \Theta_3 = \begin{pmatrix} -2.846 \\ 7.985 \\ -2.846 \\ 1.014 \\ 0.985 \\ 0.125 \end{pmatrix}, \Theta_4 = \begin{pmatrix} 2.846 \\ 7.985 \\ -2.846 \\ 1.014 \\ 0.985 \\ 0.125 \end{pmatrix}.$$

The white noise shocks, v_t, cause the agents to occasionally move away from the neighborhood of one stable equilibrium into the attractor of another stable equilibrium, which results in a time series that exhibits endogenous volatility. Figure (7) shows a time series of y_t generated under constant gain learning with a 100 period moving average of the variance shown below. The simulation shows the economy endogenously transitioning in response to white noise shocks from Θ_1 to Θ_2.

The transition between Θ_1 and Θ_2 results in a small change in the mean of y_t. The change in the mean can temporarily become very large if the transition is between distant FCEs such as between Θ_1 and Θ_3. To illustrate this change, I increase the variance of the white noise shocks to increase the likelihood that the agents' beliefs move far away from an initial FCE. Figure (8) shows the dynamics of the system when $v_t \sim N(0, 2)$. The time paths of y_t, b_1, b_2 and Γ show that the system transitions from Θ_1 to Θ_3 and results in a large temporary deviation in y_t.

5.1.1 VAR Shocks

Next, I simulate the model assuming a VAR(1) shock structure for x_{t-1}. The simulation demonstrates that the main equilibrium results of the paper carry over to more compli-

[16] There is a debate over the plausibility of large gain parameters. I do not address this debate here, but note that the gain I selected is typical for the literature. For example, Orphanides and Williams (2006) and McGough (2006) use gains between 0.01 and 0.03, while Branch and Evans (2007) uses gains that vary from 0.01 to 0.15.

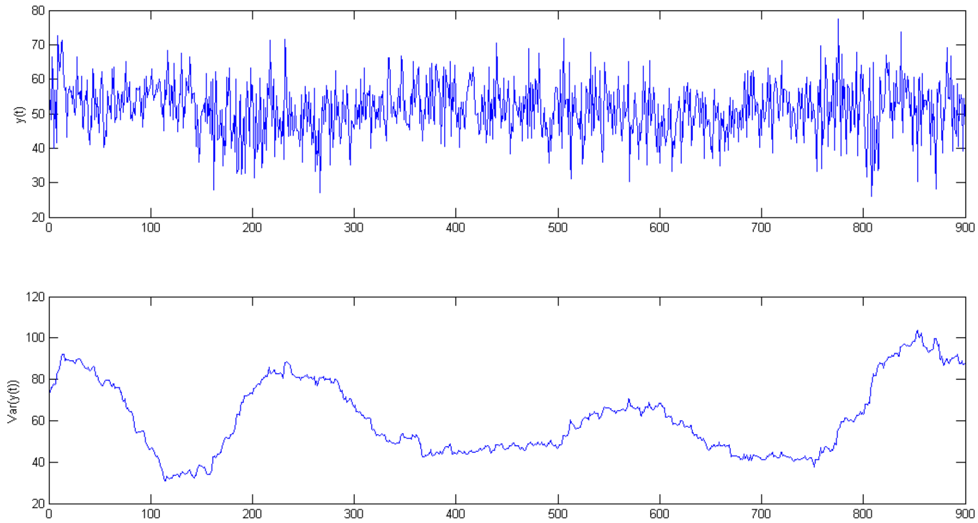

Figure 7: Time-varying volatility generated by OW forecast combination under constant gain learning.

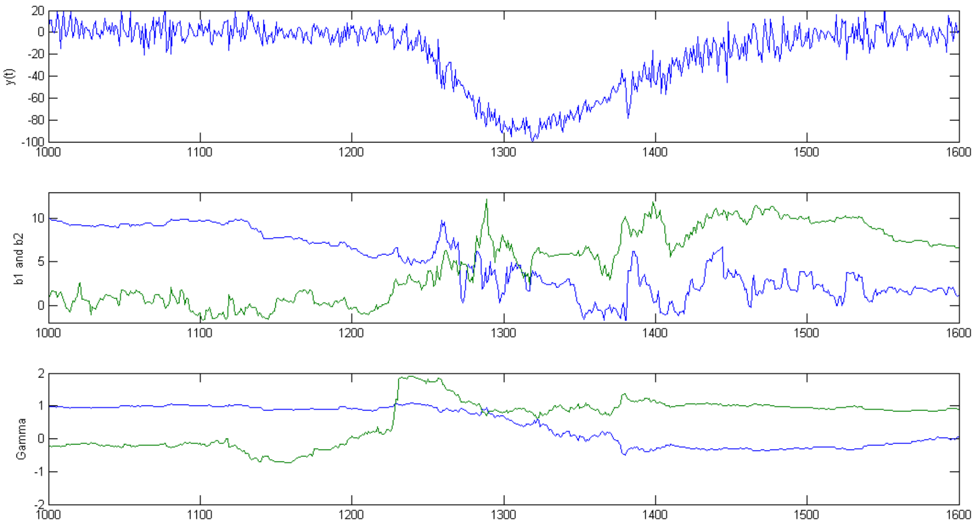

Figure 8: Time-varying volatility generated by OW forecast combination under constant gain learning.

cated shock structures. The model does not need to be altered to accommodate this new shock structure. The list of models assumed for the agents still represents all non-trivial underparameterizations of the VAR(1) process. The only change to agent behavior is in the estimated beliefs, which are altered due to different misspecification errors when compared to the case with i.i.d shocks.

The exogenous shocks of the economy are

$$x_t = Ax_{t-1} + \epsilon_t, \tag{34}$$

where A is 2×2 and ϵ_t is 2×1. The simulation uses similar parameters to those consider by Branch and Evans (2007). The parameters are $\alpha = .95$, $\zeta = (.5\ .5)'$, $\mu = 0$, and

$$A = \begin{pmatrix} .5 & .001 \\ .001 & .3 \end{pmatrix}, \Sigma_x = \begin{pmatrix} .2668 & .1190 \\ .1190 & 3.5166 \end{pmatrix}, \Sigma_\epsilon = \begin{pmatrix} .2 & .1 \\ .1 & 3.2 \end{pmatrix}, \tag{35}$$

where $\Sigma_x = Ex_{t-1}x'_{t-1}$ and $\Sigma_\epsilon = E\epsilon_t\epsilon'_t$. The agents combined the forecasts using OW. Figure (9) shows the time path of y_t and a 100 period rolling window average of the variance of y_t. The time-varying volatility shown in the figure is similar to the results obtained by Branch and Evans (2007) under dynamic predictor selection.

6 Forecast Combination and the Lucas Critique

Forecast combination techniques are particularly interesting to consider from the macroeconomic perspective because they are largely atheoretic with respect to economic theory. In fact, the very act of combining forecasts is to give up on finding a model of the true data generating process. The justification for a combination strategy is derived from the strategy's ability to forecast an existing data set in a pseudo or real-time forecasting exercise. This justification, however, lacks external validity if the forecasting technique is widely adopted throughout the economy. The mass use of the strategy would have general equilibrium effects such as those demonstrated in this paper that may alter the

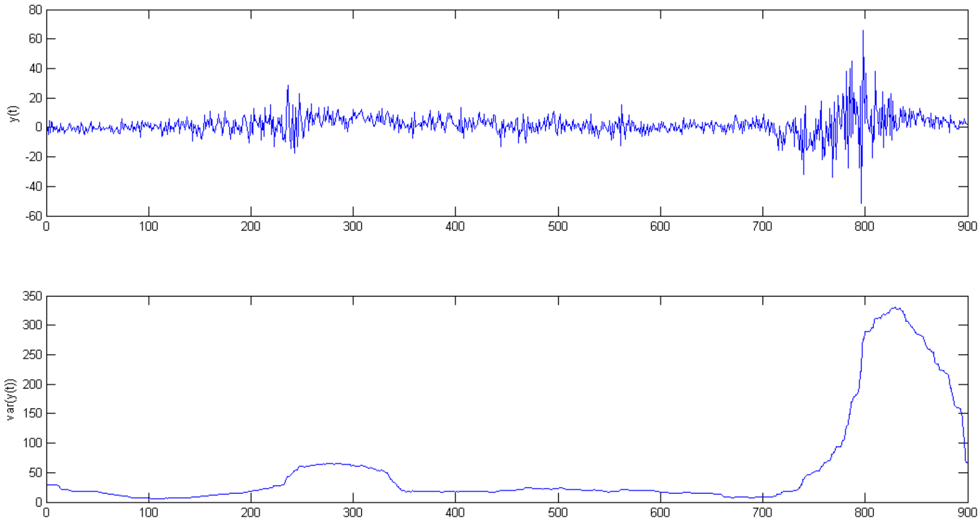

Figure 9: Time-varying volatility generated by OW forecast combination under constant gain learning with x_t following a VAR(1) process.

actual forecast efficiency. The reliance on out-of-sample forecasting exercises for justification and evaluation is subject to the Lucas Critique under the assumption of mass adoption and use. The past forecasting performance established on existing data does not imply continuing future forecasting performance.

In addition, the forecast combination strategies themselves can suffer from a more perverse form of the Lucas Critique. The non-fundamental FCEs that coexist with the fundamental FCE under optimal weights are self-fulfilling equilibria, where past forecasting success that occurred by chance is self-fulfilling as more weight is placed erroneously on the better performing forecast. This form of the Lucas Critique is more perverse because instead of receiving disconfirming information through poor forecasting performance, the agents received confirming information. The confirming information moves agents' beliefs away from the fundamentals of the economy and prevents agents from recognizing their mistakes.

Given that the goal of research in forecasting is to widely disseminate strategies and that forecast combination strategies lack economic justifications, the forecast combination equilibrium framework provides a new way to assess a forecast combination strategy. The ability of a forecast combination strategy to equal or approximate rational expectations, given a menu of forecasts, is a measure of strategy quality. A forecast combination strategy that results in deviations from ration expectations is a strategy that econometricians may not want to promote. When analyzing a forecast strategy from the macroeconomic perspective there is more at stake than mean squared forecast error. Economists should be concerned with putting forward strategies that lead to optimal decisions by individuals, policymakers, and in aggregate.

An example of the usefulness of the Forecast Combination Equilibrium approach to the empirical practice of forecasting is to apply the equilibrium results of this paper to the forecast combination puzzle. I show that the optimal weights forecast combination strategy can result in multiple, non-fundamental, and learnable equilibria, which can move the economy far from the rational predictions. In contrast, I show that equal weights results in a unique, learnable equilibrium that remains in the neighborhood of rational expectations for the majority of the parameter space. These differing results imply that the cause of equal weights dominance in the forecast combination puzzle may be immaterial, because even if optimal weights were shown to be superior in an out-of-sample forecasting exercise, their widespread use may have unintended and possibly undesirable general equilibrium effects.

This paper cannot speak explicitly to the welfare implications of forecast combination because of the reduced form model employed. But, the paper demonstrates the point that different forecast combination strategies will result in different equilibrium outcomes. This difference is a reasonable way to study and think about forecast combination for selection and justification as an addition to the current techniques employed in the forecasting literature.

7 Conclusion

Forecast combination is touted by the forecasting literature as the most robust and efficient way to forecast. In addition, combined forecasts are often the way forecasts are presented to the general public, such as with the Survey of Professional Forecasters. Due to these facts, I adopt the cognitive consistency principle to model boundedly rational agents who combine different forecasts to forecast a single endogenous state variable. The agents follow the actual recommendations of the forecasting literature to combine the forecasts. The concept of Forecast Combination Equilibrium is introduced to describe the equilibrium behavior of the agents.

The equilibrium concept is explored by assuming agents possess a menu of misspecified forecasts that together span the information set needed to form rational expectations. The agents' objective is to combine the menu of misspecified forecasts to create a combined forecast that minimizes expected squared forecast error. The Forecast Combination Equilibria that result are compared to rational expectations.

I find that different types of Forecast Combination Equilibria can both approximate and deviate substantially from rational expectations, depending on how agents combine the forecasts and the assumptions of the model. In a model with negative feedback, the combination of forecasts by optimal weights and equal weights produces unique, learnable equilibria that closely approximate rational expectations. In contrast, a model with positive feedback can have equilibria that diverge from one another and from rational expectations. The Optimal Weights FCE can produce up to six distinct equilibria that each minimize expected squared forecast error, but deviate substantially from rational expectations. These non-fundamental equilibria exist because of the self-referential nature of forecasting in the macroeconomy, where incorrect forecasts can become self-fulfilling. Furthermore, these non-fundamental equilibria are found to be stable under learning.

The use of optimal weights forecast combination strategies by agents, when analyzed under constant gain learning, are shown to exhibit time-varying volatility in the presence of high positive feedback. The dynamics are similar to the those observed in the dynamic predictor selection literature. The results shows that model uncertainty is a key driver in creating these types of outcomes.

Although this paper focuses on the representative agent case, the FCE concept can easily be adapted to accommodate heterogeneous expectations. The heterogeneous expectations case could be used to model specific forecast combination techniques employed by policymakers, such as a central bank, to characterize policy implications of different strategies. The variation in equilibrium outcomes demonstrated in this paper suggests that further study of homogeneous or heterogeneous agents who use forecast combination strategies to form expectations may help explain the stylized facts of macroeconomic and financial data, as well as contribute to the evaluation and design of actual forecasts combination strategies.

Existence Condition: The conditions for existence of an FCE require that beliefs ϕ_i for $i = 1, 2, ...k$ satisfy

$$\begin{aligned} Ez_{1,t-1}(\mu + \alpha \sum_{i=1}^{k} \gamma_i \phi_i' z_{i,t-1} + \zeta' x_{t-1} + v_t - \phi_1' z_{1,t-1}) &= 0 \\ &\cdots \\ Ez_{k,t-1}(\mu + \alpha \sum_{i=1}^{k} \gamma_i \phi_i' z_{i,t-1} + \zeta' x_{t-1} + v_t - \phi_k' z_{k,t-1}) &= 0. \end{aligned}$$

The k underparameterizations can be rewritten as $\phi_i' u_i z_{t-1}$, where z_{t-1} is $(n+1) \times 1$ and u_i is an $m \times (n+1)$ selector matrix that picks the elements out of z_{t-1} that belong in the i^{th} model. Also, the intercept term μ and ζ can be combined in $B = (\mu\ \zeta')'$ to write the system as

$$\begin{aligned} Eu_1 z_{t-1}((B' + \alpha \sum_{i=1}^{k} \gamma_i \phi_i' u_i) z_{t-1} + v_t - \phi_1' u_1 z_{t-1}) &= 0 \\ &\cdots \\ Eu_k z_{t-1}((B' + \alpha \sum_{i=1}^{k} \gamma_i \phi_i' u_i) z_{t-1} + v_t - \phi_k' u_k z_{t-1}) &= 0. \end{aligned}$$

Then simplify

$$
\begin{aligned}
Eu_1 z_{t-1} z'_{t-1}(B + \alpha \sum_{i=1}^{k} \gamma_i u'_i \phi_i) + Eu_1 z_{t-1} v_t - Eu_1 z_{t-1} z'_{t-1} u'_1 \phi_1 &= 0 \\
&\ldots \\
Eu_k z_{t-1} z'_{t-1}(B + \alpha \sum_{i=1}^{k} \gamma_i u'_i \phi_i) + Eu_k z_{t-1} v_t - Eu_k z_{t-1} z'_{t-1} u'_k \phi_k &= 0
\end{aligned}
$$

and take expectations such that $Ez_{t-1}z'_{t-1} = \Sigma_z$, which results in

$$
\begin{aligned}
-\alpha \sum_{i=1}^{k} \gamma_i u_1 \Sigma_z u'_i \phi_i + u_1 \Sigma_z u'_1 \phi_1 &= u_1 \Sigma_z B \\
&\ldots \\
-\alpha \sum_{i=1}^{k} \gamma_i u_k \Sigma_z u'_i \phi_i) + u_k \Sigma_z u'_k \phi_k &= u_k \Sigma_z B.
\end{aligned}
$$

The system of equations has a unique solution given $det(\Delta) \neq 0$

$$
\Delta = \begin{bmatrix}
(1-\alpha\gamma_1)(u_1\Sigma_z u'_1) & -\alpha\gamma_1 u_1 \Sigma_z u'_2 & \ldots & -\alpha\gamma_1 u_1 \Sigma_z u'_k \\
-\alpha\gamma_2 u_2 \Sigma_z u'_1 & (1-\alpha\gamma_2)(u_2\Sigma_z u'_2) & \ldots & -\alpha\gamma_2 u_2 \Sigma_z u'_k \\
\ldots & \ldots & \ldots & \ldots \\
-\alpha\gamma_k u_k \Sigma_z u'_1 & -\alpha\gamma_k u_k \Sigma_z u'_2 & \ldots & (1-\alpha\gamma_k)(u_k\Sigma_z u'_k)
\end{bmatrix}.
$$

Theorem 1: The theorem is proven by establishing the existence of pitchfork bifurcation for the fundamental FCE steady state. The condition for a bifurcation to occur is one of the eigenvalues of the T-map evaluated at the steady state must equal zero. This occurs in the eigenvalue associated with a_1 and a_2 for the fundamental FCE at $\alpha = \frac{1}{2}$. I proceed by describing the basic technique for characterizing a bifurcation following Wiggins (1990) and then show how to apply the technique to the T-map.

A bifurcation is 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

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

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

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

where

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

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

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

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 $f(x, h(x, \epsilon), \epsilon)$. The sufficient conditions for the existence of a bifurcation at (0,0,0) are

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

The T-map: The point of interest is the fundamental FCE, so I set $\mu = 0$ and $\sigma_{12} = 0$. To simplify the analysis, I reduce the dimension of the system by solving a_2, b_1, and b_2 in terms of a_1, γ_1, and γ_2. Let $\eta = (a_1\ \gamma_1\ \gamma_2)'$ and define differential equations as $\dot{\eta} = T(\eta) - \eta$ where

$$\dot{\eta} = \begin{pmatrix} -\frac{a_1(-1+\alpha(\gamma_1+\gamma_2))}{-1+\alpha\gamma_2} \\ -\frac{a_1^2(-1+\alpha)\gamma_1(-1+\alpha\gamma_1)^2+\sigma_1^2(-1+\gamma_1)(-1+\alpha\gamma_2)\zeta_1^2}{(-1+\alpha\gamma_2)(a_1^2(-1+\alpha\gamma_1)^2+\sigma_1^2\zeta_1^2)} \\ \frac{a_1^2(-1+\alpha)\alpha\gamma_1^2-\sigma_2^2(-1+\gamma_2)\zeta_2^2}{a_1^2\alpha^2\gamma_1^2+\sigma_2^2\zeta_2^2} \end{pmatrix}.$$

The fixed point of the system is given by $(0, 1, 1)$, which corresponds to the fundamental FCE. A change of variables is used to put the system in normal form with the fixed point at (0,0,0), and with the bifurcation occurring at 0 as well. Let $u = a_1$, $\gamma_1 = v + 1$, $\gamma_2 = w + 1$, and $\alpha = \epsilon + \frac{1}{2}$. Using the transformation, the system can be written in the form of (36) with $A = 0$, $B = (1\ 1)'$,

$$\begin{aligned} f(u,v,w,\epsilon) &= -\frac{u(v+w+4\epsilon+2v\epsilon+2w\epsilon)}{-1+w+2\epsilon+2w\epsilon} \\ g(u,v,w,\epsilon) &= \begin{pmatrix} -\frac{u^2(1+v)(-\frac{1}{2}+\epsilon)(-1+(1+v)(\frac{1}{2}+\epsilon))^2+v\sigma_1^2(-1+(1+w)(\frac{1}{2}+\epsilon))\zeta_1^2}{(-1+(1+w)(\frac{1}{2}+\epsilon))(u^2(-1+(1+v)(\frac{1}{2}+\epsilon))^2+\sigma_1^2\zeta_1^2)} \\ \frac{u^2(1+v)^2(-\frac{1}{2}+\epsilon)(\frac{1}{2}+\epsilon)-w\sigma_2^2\zeta_2^2}{u^2(1+v)^2(\frac{1}{2}+\epsilon)^2+\sigma_2^2\zeta_2^2} \end{pmatrix}. \end{aligned}$$

Let $v = ha(u,\epsilon)$ and $w = hb(u,\epsilon)$ such that $h(u,\epsilon) = (ha(u,\epsilon)\ hb(u,\epsilon))'$, then using equation (37) the center manifold must satisfy

$$D_u h(u,\epsilon)[Au + f(u, ha(u,\epsilon), hb(u,\epsilon), \epsilon)] - Bh(u,\epsilon) - g(u, ha(u,\epsilon), hb(u,\epsilon), \epsilon) = 0. \quad (38)$$

Equation (38) can be implicitly differentiated to form a second order Taylor approximations of $ha(u,\epsilon)$ and $hb(u,\epsilon)$. The approximations are substituted into $f(u, \hat{h}a(u,\epsilon), \hat{h}b(u,\epsilon), \epsilon)$

to form the center manifold. Figure 1 is a graph of the center manifold with

$$\begin{pmatrix} \hat{h}a(u,\epsilon) \\ \hat{h}b(u,\epsilon) \end{pmatrix} = \begin{pmatrix} \frac{2\sigma_1^2(-1+2\epsilon)^3\zeta_1^2}{\epsilon(u^2(1-2\epsilon)^2+4\sigma_1^2\zeta_1^2)^2} \\ -\frac{2\sigma_2^2(1+2\epsilon)(\zeta_2-2\epsilon\zeta_2)^2}{\epsilon((u+2u\epsilon)^2+4\sigma_2^2\zeta_2^2)^2} \end{pmatrix}$$

and $\sigma_1^2 = \sigma_2^2 = 1$ and $\zeta = (.9\ .9)'$. The partial derivatives of the center manifold meet the specified conditions for the existence of a pitchfork bifurcation.

Lemma 4: To solve for the FCE, first solve for a_1, a_2, and γ_0 using the corresponding equations. The three linear equations yield

$$\begin{aligned} a_1 = a_2 &= \frac{\mu}{1-\alpha} \\ \gamma_0 &= \frac{(\gamma_1+\gamma_2-1)\mu}{\alpha-1} \end{aligned}$$

Then substitute these back into the four remaining equations of the T-map.

$$\begin{aligned} b_1 &= b_1\alpha\gamma_1 + (b_2\alpha\gamma_2+\zeta_2)\frac{\sigma_{12}}{\sigma_1^2} + \zeta_1 \\ b_2 &= b_2\alpha\gamma_2 + (b_1\alpha\gamma_1+\zeta_1)\frac{\sigma_{12}}{\sigma_1^2} + \zeta_2 \\ \gamma_1 &= \frac{(b_1^2\sigma_1^2(\alpha-1)^2\alpha+\mu^2)\gamma_1 + b_1(\alpha-1)^2(b_2\sigma_{12}(\alpha-1)\gamma_2+\sigma_1^2\zeta_1+\sigma_{12}\zeta_2)}{b_1^2\sigma_1^2(\alpha-1)^2+\mu} \\ \gamma_2 &= \frac{(b_2^2\sigma_2^2(\alpha-1)^2\alpha+\mu^2)\gamma_2 + b_2(\alpha-1)^2(b_1\sigma_{12}(\alpha-1)\gamma_1+\sigma_2^2\zeta_2+\sigma_{12}\zeta_1)}{b_2^2\sigma_2^2(\alpha-1)^2+\mu} \end{aligned}$$

The γ_1 and γ_2 equations can be simplified to

$$\begin{aligned} \gamma_1 &= \frac{b_2\sigma_{12}\gamma_2(\alpha-1)+\sigma_1^2\zeta_1+\sigma_{12}\zeta_2}{b_1\sigma_1^2(1-\alpha)} \\ \gamma_2 &= \frac{b_1\sigma_{12}\gamma_1(\alpha-1)+\sigma_2^2\zeta_2+\sigma_{12}\zeta_1}{b_2\sigma_2^2(1-\alpha)}. \end{aligned}$$

Then substituting γ_1 and γ_2 into b_1 and b_2 yields

$$\begin{aligned} b_1 &= \alpha\frac{b_2\sigma_{12}\gamma_2(\alpha-1)+\sigma_1^2\zeta_1+\sigma_{12}\zeta_2}{\sigma_1^2(1-\alpha)}+(\alpha\frac{b_1\sigma_{12}\gamma_1(\alpha-1)+\sigma_2^2\zeta_2+\sigma_{12}\zeta_1}{\sigma_2^2(1-\alpha)}+\zeta_2)\frac{\sigma_{12}}{\sigma_1^2}+\zeta_1 \\ b_2 &= \alpha\frac{b_1\sigma_{12}\gamma_1(\alpha-1)+\sigma_2^2\zeta_2+\sigma_{12}\zeta_1}{\sigma_2^2(1-\alpha)}+(\alpha\frac{b_2\sigma_{12}\gamma_2(\alpha-1)+\sigma_1^2\zeta_1+\sigma_{12}\zeta_2}{\sigma_1^2(1-\alpha)}+\zeta_1)\frac{\sigma_{12}}{\sigma_1^2}+\zeta_2, \end{aligned}$$

which is linear and b_1 and b_2. This shows that the non-linearity cancels out of the system leaving a unique solution.

The second part of the proposition can be verified by substituting $\sigma_{12} = 0$ and $\mu = 0$ into the OWCFCE beliefs to verify that they equal the REE coefficients.

Theorem 2: The Jacobian matrix for the EW T-map (19) evaluated at the EWFCE is

$$\begin{pmatrix} \frac{\alpha}{2} & 0 & \frac{\alpha}{2} & 0 \\ 0 & \frac{\alpha}{2} & 0 & \frac{\alpha\sigma_{12}}{2\sigma_1^2} \\ \frac{\alpha}{2} & 0 & \frac{\alpha}{2} & 0 \\ 0 & \frac{\alpha\sigma_{12}}{2\sigma_2^2} & 0 & \frac{\alpha}{2} \end{pmatrix}.$$

The eigenvalues of the Jacobian are $\lambda_{1,2,3,4} = 0$, α, $\frac{\alpha}{2}(\rho+1)$, and $\frac{\alpha}{2}(1-\rho)$, where ρ is the correlation coefficient between $x_{1,t-1}$ and $x_{2,t-1}$. The E-stability conditions are that α must satisfy $\alpha < 1$ and $\alpha < \frac{2}{1\pm\rho}$ and since $-1 \leq \rho \leq 1$, the binding condition for stability is $\alpha < 1$.

Theorem 3: The Jacobian matrix for the OW T-map (25) evaluated at the fundamental FCE is

$$\begin{pmatrix} \alpha & 0 & \alpha & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 & \frac{\alpha\zeta_1}{1-\alpha} & 0 \\ \alpha & 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 & \frac{\alpha\zeta_2}{1-\alpha} \\ 0 & -\frac{(-1+\alpha)^2}{\zeta_1} & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & -\frac{(-1+\alpha)^2}{\zeta_2} & 0 & \alpha \end{pmatrix}.$$

The eigenvalues of the matrix are $\lambda_{1,2,3,4}$, 2α, $\alpha-\sqrt{-\alpha+\alpha^2}$, $\alpha-\sqrt{-\alpha+\alpha^2}$, $\alpha+\sqrt{-\alpha+\alpha^2}$, $\alpha+\sqrt{-\alpha+\alpha^2}$. The binding condition for E-stability is $\alpha < \frac{1}{2}$.

References

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

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

[3] J.M. Bates and C.W.J. Granger. The combination of forecasts. *Operations Research*, pages 451–468, 1969.

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

[5] W.A. Branch and G.W. Evans. Intrinsic heterogeneity in expectation formation. *Journal of Economic theory*, 127(1):264–295, 2006.

[6] W.A. Branch and G.W. Evans. A simple recursive forecasting model. *Economics Letters*, 91(2):158–166, 2006b.

[7] W.A. Branch and G.W. Evans. Model uncertainty and endogenous volatility. *Review of Economic Dynamics*, 10(2):207–237, 2007.

[8] W.A. Branch and G.W. Evans. Monetary policy and heterogeneous expectations. *Economic Theory*, 47(2):365–393, 2011.

[9] W.A. Branch and B. McGough. Replicator dynamics in a cobweb model with rationally heterogeneous expectations. *Journal of Economic Behavior & Organization*, 65(2):224–244, 2008.

[10] W.A. Branch and B. McGough. Dynamic predictor selection in a new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control*, 34(8):1492–1508, 2010.

[11] W.A. Brock and C.H. Hommes. A rational route to randomness. *Econometrica: Journal of the Econometric Society*, pages 1059–1095, 1997.

[12] W.A. Brock and C.H. Hommes. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic dynamics and Control*, 22(8-9):1235–1274, 1998.

[13] I.K. Cho, N. Williams, and T.J. Sargent. Escaping nash inflation. *Review of Economic Studies*, 69(1):1–40, 2002.

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

[15] M. Clements and D. Hendry. *Forecasting economic time series*. Cambridge University Press, 1998.

[16] M.P. Clements and D.F. Hendry. Forecasting with breaks. *Handbook of Economic Forecasting*, 1:605–657, 2006.

[17] F.X. Diebold. Serial correlation and the combination of forecasts. *Journal of Business & Economic Statistics*, pages 105–111, 1988.

[18] G.W. Evans and S. Honkapohja. *Learning and expectations in macroeconomics*. Princeton Univ Pr, 2001.

[19] G.W. Evans and S. Honkapohja. Learning as a rational foundation for macroeconomics and finance. In Roman Frydman and Edmund S. Phelps, editors, *Rethinking Expectations: The Way Forward for Macroeconomics*. Princeton Univ Pr, 2013.

[20] C.G. Gibbs. Strategic expectations: fictitious play in a model of rationally heterogeneous expectations. *mimeo*, 2012.

[21] C.W.J. Granger. Invited review combining forecasts—twenty years later. *Journal of Forecasting*, 8(3):167–173, 1989.

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

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

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

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

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

[27] R.E. Lucas. Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy*, volume 1, pages 19–46, 1976.

[28] B. McGough. Shocking escapes. *The Economic Journal*, 116(511):507–528, 2006.

[29] J.F. Muth. Rational expectations and the theory of price movements. *Econometrica: Journal of the Econometric Society*, pages 315–335, 1961.

[30] A. Orphanides and J.C. Williams. Monetary policy with imperfect knowledge. *Journal of the European Economic Association*, 4(2-3):366–375, 2006.

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

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

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

[34] J.H. Stock and M.W. Watson. Forecasting with many predictors. *Handbook of economic forecasting*, 1:515–554, 2006.

[35] B. Sturmfels. *Solving systems of polynomial equations*, volume 97. Amer Mathematical Society, 2002.

[36] A. Timmermann. Forecast combinations. *Handbook of economic forecasting*, 1:135–196, 2006.

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

[38] S. Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*, volume 1. Springer, 1990.