HETEROGENEOUS EXPECTATIONS, FORECAST COMBINATION, AND ECONOMIC DYNAMICS

by

CHRISTOPHER G. GIBBS

A DISSERTATION

Presented to the Department of Economics
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2013

DISSERTATION APPROVAL PAGE

Student: Christopher G. Gibbs

Title: Heterogeneous Expectations, Forecast Combination, and Economic Dynamics

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Economics by:

George Evans	Chair
Jeremy Piger	Core Member
Bruce McGough	Core Member
Yue Fang	Institutional Representative

and

Kimberly Andrews Espy	Vice President for Research & Innovation/ Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2013

DISSERTATION ABSTRACT

Christopher G. Gibbs

Doctor of Philosophy

Department of Economics

June 2013

Title: Heterogeneous Expectations, Forecast Combination, and Economic Dynamics

This dissertation examines the forecast model selection problem in economics in both theoretical and empirical settings. The forecast model selection problem is that there often exists a menu of different suitable models to forecast the same economic variable of interest. The theoretical portion of this dissertation considers agents who face this problem in two distinct scenarios. The first scenario considers the case where agents possess a menu of different forecast techniques which includes rational expectations but where the selection of rational expectations is costly. The assumptions that are necessary to include rational expectations as a choice are characterized and the equilibrium dynamics of a model under the appropriate assumptions is studied and shown to exhibit chaotic dynamics. The second scenario considers agents who possess a menu of econometric forecast models and examines the equilibrium outcomes when agents combine the different forecasts using strategies suggested by the forecasting literature. The equilibrium outcomes under these forecasting assumptions are shown to exhibit time-varying volatility and endogenous structural breaks, which are common features of macroeconomic data.

The empirical portion of the dissertation proposes a new dynamic combination strategy for the forecast model selection problem to forecast inflation. The procedure builds on recent research on inflation persistence in the U.S. and on explanations for the efficacy of simple combination strategies, often referred to as the forecast combination puzzle. The new combination strategy is shown to forecast well in real-time out-of-sample forecasting exercises.

CURRICULUM VITAE

NAME OF AUTHOR: Christopher G. Gibbs

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon
Westminster College, Salt Lake City, Utah
University of Utah, Salt Lake City, Utah

DEGREES AWARDED:

Doctor of Philosophy in Economics, 2013, University of Oregon
Master of Science in Economics, 2013, University of Oregon
Bachelor of Science in Economics, 2008, Westminster College
Bachelor of Science in Mathematics, 2008, Westminster College

AREAS OF SPECIAL INTEREST:

Macroeconomics and time series econometrics

GRANTS, AWARDS AND HONORS:

Invited Young Economist, 2011, Nobel Laureates Meetings, Lindau, Germany

Best Ph.D. Field Paper, 2010, University of Oregon

Dale Underwood Scholarship, 2009, University of Oregon

First-year Econometrics Scholarship, 2009, University of Oregon

Economics Faculty Appreciation Award, 2007, Westminster College

ACKNOWLEDGEMENTS

I wish to acknowledge my advisors George Evans and Bruce McGough for all the time and effort they have put into listening to my ideas and encouraging me along the way.

To Lauren for all the support, love, and aid throughout this process.

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

Figure | Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

This dissertation consist of three stand-alone pieces of research that are each variations on a common theme: the selection and combination of forecasts. The research topic covers a broad literature in economics that includes empirical and theoretical research. This dissertation makes contributions to both literatures and in many instances bridges the divide between dynamic macroeconomic theory and the applied practice of forecasting.

The first chapter considers an expectation formation problem in the dynamic macroeconomic theory literature. The literature considers deviations to standard equilibrium outcomes that may arise when individual agents face a cost to forming optimal or rational expectations. The goal of the literature is to endogenize the rationality of the agents by explicitly modeling the fact that information and complex computations are costly. The seminal paper in this literature is Brock and Hommes (1997) who posit a model where agents possess a menu of ways to form expectations that each have different implementation costs. Brock and Hommes show that when heterogeneous agents are allowed to endogenously choose between costly rational expectations or free naive expectations that the market outcomes can exhibit chaotic dynamics.

However, the results obtained by Brock and Hommes raise questions about the validity of assuming rational expectations, even for a cost, in a model where agents can explicitly choose non-rational expectations. The problem is that rational expectations requires agents to coordinate independently on a unique expectation to hold and the traditional assumptions that underlie the coordination may not be satisfied in the presence of non-rational agents. I study the assumptions underpinning rational expectations in the Brock and Hommes model to characterize the conditions under which coordination is a strategic or rationalizable outcome of the model using the eductive approach to studying expectations.

The eductive justification of rational expectations is that rational expectations directly follows from assuming the common knowledge of rationality, meaning that each agent is rational and knows that all other agents are rational and so on. I show that conditions under which coordination is found to not be a consequence of the common knowledge of rationality coincide with the conditions for chaotic dynamics in the Brock and Hommes model. I then propose an extension to their model to explore the impacts of coordination failures on equilibrium dynamics.

The second chapter considers multiple ways to form expectations in an applied forecasting problem. In the practice of forecasting there often exists a plethora of suitable ways to forecast the same macroeconomic variable of interest. It is also common to find that among the suitable methods the efficiency or accuracy of the different forecasts varies greatly over time. The instability and uncertainty that surrounds many forecast methods has prompted practitioners to consider ways to combine the many reasonable forecasts into one. A natural question that arises when combining forecasts is whether or not there exist an optimal way to combine the forecasts.

The pursuit of this question has led to an empirical puzzle. The best forecast combination procedure in practice is often a simple average of forecasts. The result is consistently obtained despite the fact that the theoretically optimal weights for a combination strategy based on the properties of the data under study often deviate substantially from the equal weights implied by averaging. The result is known as the forecast combination puzzle. In this chapter I propose a new dynamic forecasting procedure for inflation that can robustly beat the simple average of forecasts by exploiting a hypothesis for why the forecast combination puzzles exists and new empirical finding from the inflation literature. The procedure is shown to create robust and efficient forecasts on four different measures of inflation in real-time out-of-sample forecasts.

The third chapter returns to dynamic macroeconomic theory and considers the effects of the widespread use of forecast combination strategies, like those discussed in Chapter 2, on the macroeconomy. I proposes an equilibrium concept where homogeneous economic agents possess a menu of ways to forecast an endogenous variable. The agents combine the forecasts using a forecast combination strategy to create a single forecast of the endogenous variable of interest to make dynamic decisions in the model.

The combined forecasts or expectations studied in the model are self-referential, which allows for the exploration of possible unanticipated endogenous effects of employing common forecasting strategies that cannot be captured in forecasting exercises like those undertaken in Chapter 2. Forecast combination strategies are largely atheortic with respect to economic theory and their justification lies solely on their ability to perform well in "pseudo" out-of-sample forecasting exercises. Results and conclusions based on these types of studies may lack external validity because they do not capture the self-referentiality that exists in the macroeconomy, where forecasts lead to changes in agent behavior that ultimately dictates the accuracy of the forecast. I show that the consideration of self-referentiality of combined forecasts can have significant implications for macroeconomic dynamics and model predictions.

CHAPTER II

STRATEGIC EXPECTATIONS: FICTITIOUS PLAY IN A MODEL OF RATIONALLY HETEROGENEOUS EXPECTATIONS

Introduction

The dominant paradigm in dynamic macroeconomics is to model homogeneous forward-looking agents who possess rational expectations (RE). The calculation of RE is intuitively a costly process due to the amount of information and number of computations an economic agent must undertake to obtain a proper expectation. This cost implies that under certain circumstances it may be utility or profit maximizing to forgo calculating a rational expectation.

A model of rationally heterogeneous expectations (RHE) explicitly captures this scenario by assuming that agents possess a menu of ways to form expectations about the future value of endogenous state variables. The menu of ways to form expectations, known as predictor rules, includes rational expectations along with other less sophisticated rules. The predictor rules have different costs to implement based on the information and cognitive power necessary to calculate an expectation. The agents evaluate the predictors based on a fitness criterion that incorporates the costs of using each rule. The agents endogenously select among the predictor rules over time as the relative fitness of each rule changes.

The seminal paper in this literature is Brock and Hommes (1997), "A rational route to randomness" (hereafter referred to as BH).[1] BH models heterogeneous agents

[1]The RHE genesis is from the related idea of calculation equilibria proposed by Evans and Ramey (1992 and 1998) who studied agents who pay a cost to incrementally improve an expectations. They found that under certain circumstance it is optimal for the agents to not select fully rational expectations.

in a cobweb model who possess an explicit choice of rational expectations for a cost or naive expectations for free. The agents select among the two predictor rules based on the past observed net profits of each predictor. The model results in an equilibrium price process that BH calls Adaptively Rational Equilibrium Dynamics (ARED), which under certain conditions is shown to converge to a strange attractor and generates chaotic price dynamics.

The necessary condition identified in the BH model for chaotic dynamics is that the relative slopes of supply and demand must be greater than one. This condition exactly coincides with the necessary and sufficient condition identified by Guesnerie (1992) for the rational expectations equilibrium of the cobweb model to lack strong rationality. The lack of strong rationality means that individual agents, who possess only the common knowledge of rationality and full information about the parameters of the model, cannot coordinate on a unique rational expectation to hold.[2] The coinciding conditions raise the question of whether the outcomes of the model are the plausible result of individually rational firms competing in a market.

This is an interesting question because of the volume of research that has adopted the RHE modeling structure. The basic RHE framework has been studied in asset pricing models as in Brock and Hommes (1998), Gaunersdorfer (2000), and Brock, Hommes, and Woo (2009); the New Keynesian model as in Branch and McGough (2010); and in further detail in the cobweb model in works such as Brock, Dindo, and Hommes (2006) and Branch and McGough (2004 and 2008). The common finding is that heterogeneous agents with an explicit choice to hold costly rational expectations or free naive expectations can lead to exotic or chaotic dynamics. In addition, there exists a large literature that considers predictor rule selection without including the

[2]The existence of strong rationality is referred to as the eductive justification for rational expectations.

choice of a rational predictor such as in Chiarella and He (2003) and Branch and Evans (2006 and 2007). The possible implausibility of the rational predictor based on the lack of strong rationality provides justification for studying models that omit rational expectations.

I employ the eductive learning approach of Guesnerie (1992) to study the strategic motivation of predictor selection and of the rational predictor in the BH model. The term strategic is used to connote that the model's predictions are the unique outcomes resulting from individually rational agents competing in a game. I first establish that predictor selection requires a coordination assumption beyond standard RE in order to be well-defined. I then show that the rational expectations predictor, given this coordination assumption for selection, is not a strategic outcome for a broad range of the relevant parameter space. The two results imply that ARED is not justified as an outcome of rationally heterogeneous agents competing in a market, but is the result of strong implicit assumptions that coordinate agents' actions.[3]

The justification that BH is not a strategic outcome is based on analyzing an agent's predictor choice and expectation formation from a game theoretic perspective. Therefore, I propose to use another concept from game theory, fictitious play, to add a minimal amount of plausible strategic interaction into the BH framework to explore the implications of a deviation from the standard RHE assumptions. I call this new equilibrium concept Fictitious Play Equilibrium Dynamics (FPED).

Fictitious play is an adaptive learning approach to playing a game with an unknown and possibly changing opponent. Fictitious play has agents form an empirical distribution of the likely play of their competitors based on the history of

[3]Hommes and Wagener (2010) show that eductive stability of the underlying cobweb model studied in BH does not imply evolutionary stability when heterogeneous agents select among different predictor rules. But, the eductive stability of the BH framework itself and rational expectations is not studied.

past play. The agents then choose their strategy as a best response to the empirical distribution. The concept was originally proposed by Brown (1951) as a plausible way agents could compute and coordinate on a Nash equilibrium. More recent treatments and variations of the approach are found in Fudenberg and Kreps (1993), Young (2001), and Waters (2009).[4]

There are a number of advantages to adopting fictitious play as an expectation and coordination device for heterogeneous agents in the BH framework. The first is that fictitious play makes agents' predictor selection explicitly strategic by requiring agents to best respond to the expected contemporaneous actions of their competition. This feature is inherent to the BH model, but as Branch and McGough (2010) point out, has not been addressed in the literature. The second is that fictitious play is a forward-looking selection mechanism. The agents try to anticipate the other market participants' expectations based on past observations. This builds on the work of Brock, Dindo, and Hommes (2006) (hereafter Brock et al.), who study forward-looking ARED under rational expectations by adding a learning component to their model. The third is that fictitious play provides a plausible way agents could actually form strategically motivated expectations, which may provide insight into explaining the experimental results of "learning to forecast" studies such as Hommes et al. (2005 and 2007) and Hommes (2011). Finally, fictitious play has stable equilibrium dynamics in the cobweb model if agents use the process to pick production quantities rather than predictor rules as demonstrated by Thorlund-Petersen (1990).[5] In addition, Branch (2002) shows that the inclusion of a free learning-based predictor rule results

[4]Waters (2009) analyzes a fictitious play like mechanism for selecting predictor rules in the cobweb model. The paper, however, does not use it to address the strategic issue of predictor rule selection.

[5]The fictitious play stability results demonstrated in Thorlund-Petersen (1990) is in a multi-player Cournot model. The Cournot model is the finite player version of the cobweb model.

in stable equilibrium dynamics. The stability of the dynamics, when learning is present, suggests that the implicit coordination assumptions of the rational predictor may be the main driver of the dynamics. However, I show this is not the case.

The equilibrium dynamics under fictitious play replicate many of the key features of ARED in a forward-looking model with either heterogeneous or homogeneous agents. The addition of fictitious play makes the firms' beliefs a state variable of the economy. The market structure is such that there is an incentive to deviate from the predictor rule when a firm believes that the majority of market participants will select it. These incentives result in a perpetual miscoordination between beliefs and outcomes that helps to drive the equilibrium price dynamics.

The remainder of the paper proceeds as follows. Section 2 introduces the BH model and discusses the strategic decision the agents face. Section 3 explores the strategic justification of the BH model. Section 4 introduces Fictitious Play Equilibrium Dynamics and demonstrates its properties. Section 5 concludes.

The BH Model

The model of BH is a linear cobweb model that describes the price dynamics of a competitive market with a non-storable good that has a one-period production lag.[6] The model features a continuum of identical firms who select among a fixed menu of predictor rules to forecast price in order to choose a profit-maximizing quantity of goods to supply to the market. The production decision is made at time t and the goods are produced and sold at time $t+1$. The market demand at time $t+1$ is

$$D(p_{t+1}) = A - Bp_{t+1};\ \ A > 0,\ \ B > 0. \tag{2.1}$$

[6] The notation adopted follows Brock et al. (2006)

The firms supply to the market based on their expectation of price at time $t+1$, which is denoted as $p^e_{j,t+1}$, where j indicates the j^{th} firm's expectation. The firm's expectation is used to produce a profit-maximizing quantity according to

$$S(p^e_{j,t+1}) = argmax_{q_{j,t}}\{p^e_{j,t+1}q_{j,t} - c(q_{j,t})\}. \tag{2.2}$$

The firms each face an identical cost function,

$$c(q_{j,t}) = \frac{q^2_{j,t}}{2b},\ b > 0, \tag{2.3}$$

which results in a unique profit-maximizing quantity of goods to supply to the market, given an expectation of price:

$$S(p^e_{j,t+1}) = bp^e_{j,t+1}. \tag{2.4}$$

The firms choose between two predictor rules to form a forecast of price in each period t. The firms can choose the rational predictor, denoted $p^{1,e}_{t+1} = p_{t+1}$, for an explicit cost $C \geq 0$ or a naive predictor, denoted $p^{2,e}_{t+1} = p_t$, for free. The cost is imposed to represent the extra computational power and information that is necessary to calculate a rational expectation. The fraction of firms that choose the rational predictor is denoted n^1_t and the fraction of that choose the naive predictor is denoted n^2_t. Given the fractions of firms, the market clearing condition is

$$D(p_{t+1}) = n^1_t S(p^{1,e}_{t+1}) + n^2_t S(p^{2,e}_{t+1}). \tag{2.5}$$

Strategic Decisions

There are potentially two strategic decisions that firms face each period. The firms must select a predictor rule, and contingent on the selection of the rational predictor, they must select a rational expectation. The choice of a predictor is strategic because of the explicit cost of choosing the rational predictor. The cost creates a trade-off between investing resources to calculate a correct forecast and lost profits that may occur due to forecast errors. The size of the forecast errors depend on market volatility, which depends on agents' predictor choices. The rational predictor is always a strategic expectation in that it must take into account both the naive and rational forecasts of the other firms in order to determine the correct expectation.

The BH framework introduces a behavioral assumption that removes the strategic element from the predictor selection decision. The firms ignore the strategic aspect of selecting a predictor rule by being backward-looking. The firms choose predictor rules by comparing the past realized profits of each rule, net the cost of implementation. The past profit of each predictor is given by

$$\pi_t^i = \pi(p_t, p_t^{i,e}) = \frac{b}{2} p_t^{i,e}(2p_t - p_t^{i,e}) - C_i \tag{2.6}$$

for $i = 1, 2$ and is considered public information. The past profits are called the fitness measure.

The fraction of firms that choose each predictor is determined by the fitness measure and a discrete-choice econometric model:

$$n_t^i = \frac{e^{\beta U_t^i}}{\sum_{i=1}^{2} e^{\beta U_t^i}}, \tag{2.7}$$

where U_t^i, the fitness measure, is $U_t^1 = \pi_t^1 - C$ or $U_t^2 = \pi_t^2$. The parameter β is called the intensity of choice parameter and it governs the relative size of the fraction of firms that choose each rule based on the realization of the fitness measure. The intensity of choice parameter is a proxy for the firms' uncertainty over their choices. If β is large, then more firms use the rule with highest past profit. BH closes the model by defining $m_t = n_t^1 - n_t^2$ and substituting in the predictor rules into the market clearing condition (2.5) to yield the following system of equations

$$A - Bp_{t+1} = \frac{b}{2}(p_{t+1}(1+m_t) + p_t(1-m_t) \tag{2.8}$$

$$m_{t+1} = Tanh[\frac{\beta}{2}(\frac{b}{2}(p_{t+1}-p_t)^2 - C)]. \tag{2.9}$$

BH defines equations (2.8) and (2.9) as Adaptively Rational Equilibrium Dynamics (ARED).

The ARED system has a unique steady-state

$$E^* = (\frac{A}{B+b}, Tanh[-\frac{\beta C}{2}]) \tag{2.10}$$

and possesses the following stability properties:

Result 1: (*BH Theorem 3.4*) Assume that the slopes of supply and demand satisfy $b/B > 1$

i. When the cost of the rational predictor is $C = 0$, the steady-state $E = (\bar{p}, 0)$ is globally stable.

ii. When the cost of the rational predictor is $C > 0$, then there exists a critical value β_1 such that for $0 \leq \beta < \beta_1$ the steady state is globally stable, while for

$\beta > \beta_1$ the steady state is an unstable saddle path with eigenvalues 0 and

$$\lambda(\beta) = -\frac{b(1 - m^*(\beta))}{2B + b(1 + m^*(\beta))}.$$

At the critical value β_1 the steady state value $m^*(\beta_1) = -B/b$

iii. When the steady state is unstable, there exists a locally unique period 2 orbit $\{(\hat{p}, \hat{m}), (-\hat{p}, \hat{m})\}$ with $\hat{m} = -B/b$. There exists a $\beta_2 > \beta_1$ such that the period cycle is stable for $\beta_1 < \beta < \beta_2$.

Result 2: (*BH Theorem 3.2*) For $\beta = +\infty$, even when the market is locally unstable (i.e. $b/B > 1$) and when the cost of the rational predictor is $C > 0$, the system always converges to the saddle point equilibrium steady state E^*.

Furthermore, if β is set to a significantly large value, $C > 0$, and $b/B > 1$, then ARED may exhibit chaotic price dynamics. Examples of the dynamics are given in Section 4.

The Eductive Justification of the Rational Predictor

The rational predictor is inherently a strategic calculation where rational firms must anticipate the combined effect of all agents' expectations on the market price to determine a correct expectation to hold. The strategic justification of this calculation is to determine whether the agents are capable of independently coordinating on the RE price possessing only the *common knowledge of rationality* (CK):

Common Knowledge of Rationality: The individual knowledge by a rational firm of the structure of the economy and that all other firms are rational and know the structure of the economy and that all firms know that the other firms know and so on.

A rational expectation is strategically or eductively justified if the agents, given the stated assumptions, can coordinate on a unique expectation to hold through iterated deletion of strictly dominated strategies. In what follows, I present the basic results from Guesnerie's (1992) analysis of the cobweb model without predictor rule selection to motivate and demonstrate the tools necessary to analyze the more complicated case.

The basic cobweb model is composed of a linear demand and supply functions given by equations (2.1) and (2.4). The unique rational expectations equilibrium (REE) of the model is

$$\bar{p} = \frac{A}{B+b}. \tag{2.11}$$

The REE is strategically or eductively justified if the following definition holds:

Strongly Rational: A rational expectation is strongly rational if it is the unique rationalizable expectation for firms to hold assuming the common knowledge of rationality.

The strong rationality of an expectation is determined by analyzing the strategic choices of agents as a one period normal-form game. The cobweb model can be formulated as a game with a continuum of players on the unit interval, $J = [0, 1]$. The strategy set of the players is all feasible prices $P = [0, A/B]$ such that an individual's price expectation is $p^e \in P$. The aggregate expected price in the market is $\int p_j^e dj$, which given equation (2.4) results in a market supply of $b \int p_j^e dj$ and an equilibrium price $p = D^{-1}(b \int p_j^e dj)$. The j^{th} firm's best response in this game is to pick p^e such that $p^e = p$.

The agents choose a best response using iterated deletion of strictly dominated strategies. Given the common knowledge of rationality, it follows that each firm only chooses strategies that are best responses to the possible profile of strategies other rational firms would actually play. Therefore, there may exist a subset of strategies that are never played because they are never best responses. If the firms eliminate these strategies, then they have a smaller feasible strategy set $P^1 \subset P$ from which to choose. The smaller strategy set P^1, however, can be put through the same deduction and the firms may again be able to discard strategies to arrive at $P^2 \subset P^1 \subset P$. If firms continue this line of reasoning, it may be possible to reduce the strategy set to $P^\infty = p^e$, which implies that each firm will conclude there is only one price expectation a rational firm would choose. For the cobweb model, if $P^\infty = \bar{p}$, it is said that $\bar{p}$ is strongly rational.

Result 3: *Guesnerie (1992): Proposition 1* - Given equations (2.1) and (2.4), i) if $b/B < 1$, then REE of the cobweb model is strongly rational. ii) If $b/B \geq 1$, then the REE of the cobweb model is not strongly rational and the set of rationalizable prices is P.

Therefore, in the basic cobweb, the REE is not a consequence of strategic interaction when $b/B \geq 1$.[7]

[7] The term rationalizable means that a strategy is not strictly dominated by another feasible strategy.

The Common Prior

The assumption necessary for coordination of rational expectation in the absence of strong rationality is that firms hold model consistent beliefs or beliefs that are in line with "the relevant economic theory," as stated in Muth (1961). This belief is called by Aumann and Dreze (2008) the *common prior belief* (CP):

Common Prior Belief: The strategy choices of different types of agents are common knowledge.

The CP is a coordinating assumption akin to the law iterated expectations for heterogeneous agents, where firm j's expectation of firm k's expectation of price is $E^j E^k[p] = E^j[p]$ for $j \neq k$.[8] If agents possess the common knowledge of rationality and the common prior belief, they can select the REE price from the set of rationalizable prices with the expectation that the other firms will make the same choice.

Strategic Interaction in ARED

The cobweb game with predictor selection is more complicated. The firms in the expanded game have two choices: they must select a predictor and then choose an expectation of price. The predictor choice can be thought of as the agents choosing a type. If agents choose the rational predictor, then they play the cobweb game as described in the previous section. If they choose the naive predictor, then they abandon strategy and choose last period's price. The CP assumption in this context

[8] The idea of an analog to the law of iterated expectations for heterogeneous agents is also explored in Branch and McGough (2009) in the context of the New Keynesian model.

is not sufficient to motivate predictor rule choices or the rational predictor because the agents cannot always deduce the distribution of player types.

To illustrate the issue, suppose that the firms who select the rational predictor are endowed with CK and CP, where CP imparts the knowledge of each type's strategy, but does not include information about n_t^1 or n_t^2, the fraction of types. Knowledge of the fraction of firms that choose each rule is necessary to calculate a rational expectation and to choose a predictor rule. Under certain conditions, without knowledge of the fractions, the firms cannot deduce a unique predictor to select. Each predictor or firm type is rationalizable given the other agents' potential actions. To formalize the argument, let the fraction of firms who choose the rational predictor be n, the fraction of naive firms be $1-n$, and let the naive predictor forecast be denoted p_{-1}.

Theorem 1: Suppose that the firms are endowed with CK and CP. The firms can deduce that

i. if $|\bar{p}-p_{-1}| \leq \frac{B}{B+b}\sqrt{\frac{2C}{b}}$, then $n=0$ and $p^{2,e}=p_{-1}$ is the unique rationalizable predictor choice.

ii. if $|\bar{p}-p_{-1}| \geq \sqrt{\frac{2C}{b}}$, then $n=1$ and $p^{1,e}=p$ is the unique rationalizable predictor choice.

iii. if $|\bar{p}-p_{-1}| \in (\frac{B}{B+b}\sqrt{\frac{2C}{b}}, \sqrt{\frac{2C}{b}})$, then both $p^{2,e}=p_{-1}$ and $p^{1,e}=p$ are rationalizable.

Proof: The firms are profit maximizers and will choose the predictor rule with the highest profit. The rational predictor under ARED assuming CK and CP is

$$p = \frac{A - (1-n)bp_{-1}}{B + bn}. \tag{2.12}$$

Equating profits of the predictors for an unspecified n results in

$$\begin{aligned} \pi_1 &= \pi_2 \\ \frac{b}{2}p^2 - C &= \frac{b}{2}p_{-1}(2p - p_{-1}). \end{aligned}$$

Solving the equation the for p_{-1} and substituting in (2.12) yields

$$p_1 - \bar{p} = \pm \frac{B + bn}{B + b}\sqrt{\frac{2C}{b}},$$

where $\bar{p} = A/(B+b)$. (i) If $|\bar{p} - p_{-1}| \leq \frac{B}{B+b}\sqrt{\frac{2C}{b}}$, then the rational predictor is never more profitable than being naive, therefore all agents choose to be naive and $n = 0$. (ii) If $|\bar{p} - p_{-1}| \geq \sqrt{\frac{2C}{b}}$, then the naive predictor is never more profitable than the rational predictor and all agents will choose $n = 1$. (iii) If $|\bar{p} - p_{-1}| \in (\frac{B}{B+b}\sqrt{\frac{2C}{b}}, \sqrt{\frac{2C}{b}})$, then there exists $n \in (0,1)$ such that $\pi_1 = \pi_2$. Therefore, the rational and naive predictors are each rationalizable. Q.E.D.

The agents are unable to determine the contemporaneous strategy choices of the other firms for a range of feasible prices because if the strategies are chosen in exactly the right proportions, then a firm is indifferent between the rational or naive predictors. Any $n \in [0,1]$ could result in the model because there is no way to coordinate which firms choose the rational predictor and which firms chooses the naive predictor to enforce the correct fraction given a realization of p_{-1} in the relevant range.[9] Therefore, the standard assumptions invoked when using rational expectations are not sufficient to justify the firms' deduction of n or the realization of n as a market outcome when there exists rationally heterogeneous expectations.[10]

Rationally heterogeneous expectations requires an additional assumption to justify the knowledge of n for the range of feasible prices where there exists heterogeneous predictor choices. We define this assumption as the *non-rational agent common prior belief* (NACP)

Non-rational Agent Common Prior: The fraction of firms who choose the rational predictor is n and the selection of the rational predictor imparts common knowledge among the rational firms of n.

The NACP assumption is similar to the notion of a Correlated equilibrium, where firms are told a strategy to choose and it is rational for a firm to choose that strategy given all players follow their instructions.[11] The NACP assumption is invoked in the BH model by assuming it is common knowledge that predictor selection is based on past profit and knowledge of the discrete choice rule given by equation (2.7) that governs the precise fraction of firms that choose each rule.

The knowledge of n is necessary to form a rational expectation, which implies that NACP is a necessary condition to form a rational expectation in ARED. Thus,

[9]The third case is an example of a mixed strategy equilibrium. However, the playing of mixed strategies is not permitted.

[10]Note that this is a stronger result than that of strong rationality because it holds for all b/B.

[11]See Aumann and Dreze (2008) for more information on the relationship between Correlated equilibria and RE.

the next question is to determine whether CK and NACP are the only necessary conditions to make the rational predictor a strategic outcome of the model.

Theorem 2: Given CK and NACP, the rational predictor of ARED is i) strongly rational if $nb/B < 1$. ii) The rational predictor is not strongly rational if $nb/B \geq 1$ and the set of rationalizable prices is $P_0 = [0, (A - b(1-n)p_{-1})/B]$.

Proof: The agents who choose the RE predictor possess CK and NACP, which implies they know market clearing is given by

$$A - Bp = nbp^e + (1-n)bp_{-1}$$

Market clearing can be rewritten as

$$p = \mu_0 + \mu_1 p^e,$$

where $\mu_0 = (A - b(1-n)p_{-1})/B$ and $\mu_1 = nb/B$. The agents possess CK, so the result is obtained by applying Result 3. Q.E.D.

The intuition for theorem 2 is that the rational agents know at a minimum $(1-n)bp_{-1}$ of goods will be supplied to the market. This implies that the highest price the market will obtain is $p^H = D^{-1}((1-n)bp_{-1})$, which results in an initial feasible strategy set of $P_0 = [0, p^H]$. The agents can perform iterated deletion of strictly dominated strategies by exploring whether p^H is ever a best response. If the agents that choose the rational predictor expect p^H, then the price next period is $p_1 = D^{-1}(nbp^H + (1-n)bp_{-1})$. If $nb/B < 1$, then $0 < p_1 < p^H$ and all $p^e < p_1$, such that $p^e \in P_0$, are revealed to never be best responses. This results in a smaller feasible set $P_1 = [p_1, p^H]$ of possible best responses. The agents can continue the deduction by evaluating whether p_1 is a best response. The firms will find that this again leads to the elimination of more strategies. If the deduction is continued it will eventually converge to the rational expectation price forecast. If however $nb/B \geq 1$, then $p_1 \leq 0$ and no prices are eliminated for being a best response. The entire set P_0 is rationalizable.

The conditions of Theorem 2 for strong rationality are less restrictive than in the standard case presented in Result 3. For $n < 1$, the firms can coordinate on a rational expectation for values of b/B where coordination is not justified in the basic cobweb framework. This has a nice economic intuition because as n increases, the number of agents whose forecasts are known with certainty decreases. The more uncertainty that enters the market, the harder the agents find it to coordinate on a single rational expectation. However, this result does not necessarily imply that the rational predictor is strategically justified for a range of b/B that greatly exceeds Result 3 because n is dynamic in ARED. If $b/B > 1$, then the instability of the price dynamics will cause n to fluctuate between zero and one. The fluctuations may lead to periods where coordination is not strategically justified. The likelihood of unjustified coordination rises with the instability of the dynamics and with the likelihood of chaotic dynamics.

Theorems 1 and 2 show that CK, CP, and NACP are the implicit coordination assumptions that are necessary to coordinate actions in ARED. The ARED framework is not well defined under only the CK assumption by Theorem 1, and the rational predictor is in general not strategically justified under CK and NACP by Theorem 2 in the relevant parameter space. Hence, there is in general no strategic justification for the interesting ARED outcomes in the current modeling framework. The dynamics are not the result of rationally heterogeneous firms competing in a market, but they are the manifestation of strong coordinating assumptions. This implies that the inclusion a rational predictor rule is no more justified than the inclusion of a heuristic, econometric, or boundedly rational rules on the menu of predictor choices to coordinate agent behavior.

Strategic Expectations

The inclusion of the rational predictor, however, is a natural modeling choice in dynamic macroeconomics. Therefore, in this section I propose an extension to ARED that adds a strategic element to the model while retaining as much of the original structure as possible. The dynamic properties of the extension are characterized and compared to ARED.

The CK and CP assumptions are standard in macroeconomics and are retained in what follows. The NACP is non-standard and I focus my attention on adding strategic interactions to model this assumption. I modify the NACP assumption employed in BH by assuming that firms engage in fictitious play. The firms construct an empirical estimate of the distribution of predictor choices in an average period to create an expectation of the contemporaneous predictor rule choices that will take place in the current period. The firms then choose a predictor rules as a best response to this expectation.

Fictitious play is added to the existing framework by assuming firms estimate the difference in the fraction of firms who choose the rational and naive predictors. The firms use the estimate to pick a best response and if they choose the rational predictor, they use the estimate of m_t along with CK and CP to form an expectation.

The firms estimate m_t using a simple learning rule

$$\zeta_{t+1} = \zeta_t + \gamma(m_t - \zeta_t), \tag{2.13}$$

where ζ_{t+1} is the estimated difference in the fraction of firms that choose each predictor and $0 < \gamma \leq 1$ is the gain parameter that governs the weight placed on new observations.[12] The firms use the estimate to calculate the expected profitability of the two predictor rules and the predictor with the highest expected profit is chosen.

This best-response assumption requires the firms to be forward-looking, which introduces a simultaneity issue into the BH framework when the cost of the rational predictor is positive. The firm must know the rational forecast in order to know whether it is profitable to pay the cost and choose the rational predictor. This is a known issue of dynamic predictor selection models and is addressed in Brock et al. (2006), who investigate forward-looking ARED under rational expectations. I employ a similar framework to the one they developed to overcome the simultaneity issue.

The Expert Manager

Brock et al. (2006) posits the existence of a single expert manager who has the ability to form rational expectations and charges the firms a fee for her service. I modify this concept and assume that there exists a continuum of expert managers on the unit interval that have structural knowledge about the economy equivalent to the CK and CP assumptions. An expert manager can be hired by the firm for the fee $C \geq 0$ in order to calculate an expert prediction. An expert manager can only provide services to one firm each period. There is perfect competition among

[12]The gain parameter is a common feature of learning models in macroeconomics. For a more in depth discussions see Evans and Honkapohja (2001).

the managers and any firm can hire any manager, which results in a uniform fee that is equal to the expert managers' marginal cost of producing a prediction $C \geq 0$. The expert managers possess CK and CP which implies that their price expectations are coordinated given a shared belief of ζ_t, the fraction of firms believed to hire an expert manager in each period. The fraction of firms that hire an expert manager and the expected net profits of doing so are public information.[13] The information can be thought of as provided by a government agency that monitors the market. The fraction of firms that hire an expert manager in a given period is determined by the discrete-choice random-utility framework of the original BH model given by equation (2.7).

The expert managers use the belief ζ_t, CK, and CP to produce the following price prediction:

$$p^{1,e}_{M,t+1} = \frac{2A - bp_t(1-\zeta_t)}{2B + b(1+\zeta_t)}. \tag{2.14}$$

An expert manager's prediction is identical to the rational predictor of ARED, except that m_t is replaced with ζ_t. An expert manager is acting strategically in using ζ_t to form a forecast by attempting to predict the contemporaneous strategy choices of the other firms and the effect those choices will have on realized profits. The expert managers' prediction is used to calculate the expected profits of paying for their

[13] An additional assumption is made in Brock et al. (2006) that the fee is unknown to firms to prevent the firms from reverse engineering the price forecast from the advertised expected net profit. I omit this assumption because the reverse engineering violates the spirit of the cost associated with rational expectations. A firm that reverse engineers the price would have to pay the same cost as hiring a manger to do so.

service or remaining naive using the following equations:

$$\pi^{1,e}_{t+1} = \frac{b}{2}(\frac{2A - bp_t(1-\zeta_t)}{2B + b(1+\zeta_t)})^2 - C \tag{2.15}$$

$$\pi^{2,e}_{t+1} = \frac{b}{2}p_t(\frac{2A - bp_t(1-\zeta_t)}{2B + b(1+\zeta_t)} - p_t). \tag{2.16}$$

The resulting expected net profits are advertised to the firms and used as fitness measure to make a hiring decision.

As acknowledged in Brock et al. (2006), the expert manager is a limited description of market behavior. But, given the difficulties that exist with modeling heterogeneous agents, it is an interesting theoretical benchmark to consider and a necessary assumption in order to make direct comparisons to past work.

Fictitious Play Equilibrium Dynamics

The expert manager's profit advertisements, the predictor $P^{1,e}_{M,t+1}$, the discrete choice model, and the learning rule together form Fictitious Play Equilibrium Dynamics (FPED):

$$A - Bp_{t+1} = \frac{b}{2}(\frac{2A - bp_t(1-\zeta_t)}{2B + b(1+\zeta_t)}(1+m_t) + p_t(1-m_t)) \tag{2.17}$$

$$\zeta_{t+1} = \zeta_t + \gamma(m_t - \zeta_t) \tag{2.18}$$

$$m_t = Tanh[\frac{\beta}{2}(\frac{b}{2}(\frac{2A - bp_t(1-\zeta_t)}{2B + b(1+\zeta_t)} - p_t)^2 - C)]. \tag{2.19}$$

FPED has two state variables p_t and ζ_t. The introduction of fictitious play makes market beliefs, as opposed to market outcomes, a state variable of the economy.

The switch from outcomes to beliefs only introduces subtle changes to the dynamics properties of the market for small values of the intensity of choice parameter. Fictitious Play Equilibrium Dynamics retains the same unique fixed point as ARED

$$E^* = (p^*, \zeta^*) = (\bar{p}, Tanh[-\frac{\beta C}{2}]) \tag{2.20}$$

and it possesses the following stability properties.

Theorem 3: Assume that the slopes of supply and demand satisfy $b/B > 1$ and $0 < \gamma \leq 1$

i. When the cost of the rational predictor is $C = 0$, the steady-state $E = (\bar{p}, 0)$ is locally stable.

ii. When the cost of the rational predictor is $C > 0$, then there exists a critical value β_1 such that for $0 \leq \beta < \beta_1$ the steady state is locally stable, while for $\beta > \beta_1$ the steady state is an unstable saddle path with eigenvalues $1 - \gamma$ and

$$\lambda(\beta) = -\frac{b(B + b\zeta^*(\beta) + (b + B)\,Tanh(\beta C/2))}{B(b + 2B + b\zeta^*(\beta))}.$$

At the critical value β_1 the steady-state value $\zeta^*(\beta_1) = -B/b$

iii. When the steady state is unstable, there exists a locally unique period 2 orbit $\{(\hat{p}, \hat{\zeta}), (-\hat{p}, \hat{\zeta})\}$ with $\hat{\zeta} = -B/b$. There exists a $\beta_2 > \beta_1$ such that the period cycle is stable for $\beta_1 < \beta < \beta_2$.

The proof appears in the appendix. Result 1 and Theorem 3 reveal that the FPED has similar dynamics to ARED locally and for small values of β. However, The two systems diverge globally and for large values of β, when β approaches positive infinity.

Theorem 4: Assume that the slopes of supply and demand satisfy $b/B > 1$, $C > 0$ and $\beta = +\infty$, then $E = (\bar{p}, -1)$ is an unstable saddle path if $0 < \gamma < 1$ and globally stable if $\gamma = 1$.

Theorem 4 illustrates the effect that strategic interaction has on coordination. When $0 < \gamma < 1$, the firms have uncertainty over the contemporaneous choices of the other firms in the market, which is grounded in past observed behavior. The firms observe heterogeneous predictor rule selection between periods and they posit that there may exist heterogeneous predictor rule selection within each period. The best response of the firms is to incorporate this uncertainty into their expectations. The uncertainty drives a wedge between beliefs and actual actions that can never be overcome. It is only when the firms discard information (i.e. $\gamma = 1$) that the uncertainty is removed and coordination is achieved.[14] The same intuition of perpetual miscoordination of belief and actual actions taken in the market holds in the case with finite-β.

Chaotic Dynamics

The hallmark of ARED is the chaotic price dynamics that arise for large but finite values of β. Similar price dynamics are obtained for FPED. The FPED system, however, has two bifurcation parameters that govern dynamics in the model: the gain parameter γ and the intensity of choice parameter β. I demonstrate the effect of both parameter on equilibrium dynamics.

[14]In the limiting case, the miscoordination is perhaps an obvious error on the part of the firms and expert managers, but the solution to the error is not straightforward. An expert manager that recognized the mistake and attempts to make a rational deviation will find themselves in another eductively unjustified decision, where they must consider the possibility that other expert managers have noted the same relationship and are trying to exploit it.

Figure 1 depicts the bifurcation diagrams for β with fixed values of γ, along with the bifurcation diagram for ARED (panel (a)). The figures demonstrate the similar equilibrium dynamics of ARED and FPED around the initial bifurcation, $\beta < \beta_2$, as described in Theorem 3 and Result 1. The differences between plots (b), (c), and (d) are generated by choosing different values for the gain parameter γ. The choice of γ affects dynamics most when β is large. The γ parameter can cause the appearance of stable cycles for ranges of the parameter space that are chaotic for ARED and it causes the variance of the overall price process to increase.[15] The dynamics portrayed in Figure 14 lie in between the results of the rational forward-looking agents of Brock et al. (2006) and the backward-looking agents of ARED depending on the gain. The rational agents of Brock et al. (2006) produce the same initial dynamics for small values of β, but remain in a stable 2-cycle for large values of β. FPED mimics ARED for low values of γ by generating chaotic and complex dynamics.[16] But as γ increases, FPED can transition into stable cycles similar to Brock et al. (2006).

Figure 2 depicts the bifurcation diagrams for γ with fixed values of β. The main diagram of interest is (c), which shows the dynamics of FPED when $\beta = +\infty$. The dynamics of (c) are generated by homogeneous firms who believe that there may exist heterogeneous predictor choices. These dynamics are achieved with no explicit or implicit randomness in the system. The perpetual miscoordination of belief is the primary driver of the outcome. The agents always believe there is a chance of heterogeneity in predictor choices within each period because they observe heterogeneity between periods. The chaotic dynamics witnessed in this case are

[15] I forgo formal demonstrations of sensitive dependence on initial conditions for FPED because the chaotic dynamics that arise in the cobweb model is a well-documented phenomena.

[16] The dynamics in plot (d) appears as if the model has not converged to the attractor. This is not the case. Further analysis reveals there are multiple coexisting attractors that fill out the diagram if plotted.

unique to FPED. No other known version of the BH model generates chaotic dynamics in the $\beta = +\infty$ case without the introduction of explicit random shocks.

The simulations show that rational routes to randomness remain when the market is composed of forward-looking and strategic agents endowed with the usual - CK and CP - rational assumptions. The introduction of a learning rule to motivate the expert manger's or "rational" predictor does not enforce stability in the market. It instead adds a new plausible dimension of miscoordination that was missing from the model that can drive and justify the exotic equilibrium price dynamics.

Forecast Errors

The perpetual miscoordination between beliefs and actions highlighted in the $\beta = +\infty$ case is the key feature of FPED. The price dynamics of FPED are justified as the consequence of miscoordination among strategic firms. The miscoordination, though, does lead to perpetual incorrect predictions of the fraction of firms that choose each predictor and incorrect expert forecasts. This raises the question as to whether the expert manager would want to change the gain parameter on the fictitious play learning rule or the entire rule in response to forecast errors.

A natural reason to desire a change in the learning rule is if the forecast errors are systematic. FPED can produce systematic and non-systematic forecast errors for the learning rule and the expert manager's prediction. Simulations show that whether the forecast errors are systematic is dependent on the parameterization of the model.

Figure 15 shows examples of non-systematic and systematic forecast errors for a single parameterization of FPED with two different gain parameters. The figures in column (a) depict the non-systematic forecast errors from the expert manager's prediction when $\gamma = 0.95$. The top graph (a.1) plots (p_t, m_t) and (p_t, ζ_t) to demonstrate the strange attractor of FPED under this parameterization to illustrate the underlying reason the forecast errors are not systematic, which is mathematical chaos. The plot shows the time path of (p_t, m_t) and (p_t, ζ_t) for 20,000 periods after an initial 10,000 iterations of FPED. The last two graphs of column (a) show the forecast errors of the expert managers' price prediction and the forecast errors of the prediction for the difference in the fraction of predictors chosen over time. The forecast errors for the learning rule appear to almost be white noise (panel (a.3)). The mean forecast error of the learning rule is zero and the first autocorrelation is 0.0776, which suggest there is not a strong empirical reason to alter the rule. The forecast errors of the expert forecast are also mean zero and have a first autocorrelation of 0.2791. Ironically, the higher autocorrelation in the expert managers' price forecasts provides empirical justification for abandoning rational expectations (CK and CP) over fictitious play learning.

The figures of column (b) show the opposite result for the same parameterization of the model with $\gamma = 0.5$. The forecast errors are systematic and the expert manager may indeed wish to alter the forecast rule. This line of reasoning adds a secondary layer of dynamic predictor selection to the model, which further illustrates the complexity of heterogeneous expectations. Further exploration of these issues is interesting, but it is beyond the scope of this paper. My goal is only to provide evidence of the strategic issues that are present in rationally heterogeneous expectations models.

Conclusion

The selection of predictor rules in a model of rationally heterogeneous expectations is inherently motivated by strategic interaction. A firm considers alternatives to costly rational expectations because there may exist a profit opportunity to use a free naive predictor depending on the other firms' predictor choices. However, the introduction of an explicit choice of rational or non-rational expectations requires that additional coordination assumptions be made to coordinate agent behavior. The coordinating assumptions remove the underlying strategic justification for the model's outcomes.

I demonstrate this fact by showing that the standard rational expectations assumptions are insufficient to allow agents to deduce the distributions of predictor rules chosen under certain conditions. The model requires a non-rational agent common prior belief that assigns each agent a predictor rule to choose and which imparts the knowledge of the distribution of predictor rules chosen in the market. Only with this information can a well-defined predictor rule choice be made and a rational expectation calculated. I then show that the rational expectations predictor is not strategically justified in a relevant range of the parameter space, even if agents possess the non-rational agent common prior belief in addition to the common knowledge of rationality. The two results imply that the standard dynamic predictor selection framework that results in chaotic dynamics is not justifiable as the market outcome of competing, strategic, rationally heterogeneous firms.

The introduction of a minimal amount of strategic interaction into the BH model through the introduction of fictitious play is shown to not alter the basic conclusions of the BH model. There still exist rational routes to randomness. Fictitious play actually increases the range of the known parameter space for which these price dynamics are possible. The impact of fictitious play is to change the justification for the price dynamics. The model's outcomes under Fictitious Play Equilibrium Dynamics are the result of strategic behavior and perpetual miscoordination between firms' beliefs of the contemporaneous actions of their competitors and the actual actions taken in the market.

Supplementary Material and Proofs

Theorem 3 Proof: The proof follows closely to the proof of Result 1 from Brock and Hommes (1997) as the two systems are very similar.

i) When $C = 0$ the steady state is $E = (\bar{p}, 0)$. The eigenvalues of the system are $-1 < \lambda_1 = -b/(b + 2B) < 0$ and $\lambda_2 = 1 - \gamma < -1$. Therefore it is locally stable. ii) For $C > 0$, the first eigenvalue is given by

$$\lambda_1(\zeta) = -\frac{b(B + b\zeta^*(\beta) + (b + B)Tanh(\beta c/2))}{B(b + 2B + b\zeta^*(\beta))},$$

while λ_2 remains the same. As β increases from 0 to infinity, $\zeta^*(\beta) = Tanh(-\beta c/2)$ decreases from 0 to -1, which implies that $\lambda_1(\zeta)$ decreases from $-b/(b + 2B)$ to $-b/B < 1$. Thus, the steady state becomes an unstable saddle point for some critical value β_1. For $\beta = \beta_1$, the eigenvalue $\lambda_1 = -1$, which implies $\zeta^*(\beta_1) = -B/b$.

iii) For $\beta > \beta_1$ a 2 - cycle is created as in ARED. The symmetry of the system implies that the 2 - cycle is of the form $\{(p_1, \zeta^*), (p_2, \zeta^*)\}$ where $p_2 = -\lambda_1(\zeta)p_1$. Let $\hat{p} = p1 = -p2 > 0$ and it must be the case that $Tanh(\beta/2(2b\hat{p}^2 - c)) = -B/b$. This equation has a positive solution when $\beta > \beta_1$. The local uniqueness and stability of the 2 -cycle are established using the center manifold reduction technique described by Wiggins (1990). The approximate center manifold is

$$\begin{aligned} u_{t+1} &= f(u_t) \\ &= -\frac{b\left(B+(b+B)\,Tanh\left[\frac{c\epsilon}{2}\right]\right)}{B(b+2B)}u_t + \Omega u_t^3 + O[u]^4 \end{aligned}$$

where $\Omega = \frac{b^2(b+B)^2\epsilon Sech\left[\frac{c\epsilon}{2}\right]^3\left(B(2B+b(2+\gamma))\text{Cosh}\left[\frac{c\epsilon}{2}\right]+b(b+B(1-\gamma))Sinh\left[\frac{c\epsilon}{2}\right]\right)}{B(b+2B)^3\left(2B+b\,Tanh\left[\frac{c\epsilon}{2}\right]\right)}$ and ϵ is substituted for β_1. Stability of the cycle is given by checking that $a = \frac{1}{2}f''(0) + \frac{1}{3}f'''(0) > 0$, which holds at β_1 if $b > B$. Q.E.D.

Theorem 4 Proof: There are two cases to consider:

Case 1: Suppose that $\gamma = 1$ and consider an initial point $E_1 = (p_1, \zeta_1) \neq E^*$. If $\gamma = 1$, then $\zeta_{t+1} = m_t$ and if $\beta = +\infty$ then m_t takes only the values -1 and 1. The naive predictor is unstable so for any $E_1 \neq E^*$ the price will eventually satisfy $|\bar{p}-p_t| > (B+b(1+\zeta_t)/2)/(B+b)\sqrt{2C/b}$, at which point $m_t = 1$ because $\pi^{1,e}_{t+1} > \pi^{2,e}_{t+1}$. The next period the price is p_{t+1} and $\zeta_{t+1} = m_t = 1$. If the expert prediction has the highest expected profit and is chosen by all agents, then the system is mapped onto E^* because $p^{1,e}_{M,t+2} = \bar{p}$ and $\zeta_{t+2} = m_{t+1} = 1$. This occurs if $b/2\bar{p}^2 - C > b/2p_{t+1}(2\bar{p}-p_{t+1})$, which holds if $p_{t+1} > (B+b(1+\zeta_t)/2)/(B+b)\sqrt{2C/b}$. Now p_{t+1} is the market outcome of all firm using the expert prediction, so either $\zeta_t = m_{t-1} = 1$ and $p_{t+1} = \bar{p}$ and the system is already at E^* or $\zeta_t = m_{t-1} = -1$, which implies $|\bar{p} - p_{t+1}| > |\bar{p} - p_t|$ by $b/B > 1$ and the system is mapped to E^* in period $t + 2$.

Case 2: By Theorem 2 the naive predictor is always chosen when price is in a neighborhood of the steady state and the naive predictor is unstable when $b/B > 1$. The stable arm of the unstable saddle path is defined by the set $E_{sa} = \{(\bar{p}, \zeta)|\zeta \in$

$[-1,1]\}$. If the price $p_t = \bar{p}$, then the naive predictor is chosen and $m_t = -1$. This follows from the tie-breaking rule assumed in BH, where firms choose to be naive if profits are equal. The naive predictor is chosen every period and the price remains at $\bar{p}$. For any initial ζ, the learning rule converges asymptotically to $\zeta^* = -1$ for $0 < \gamma < 1$.

The steady state is not globally unstable because for any initial $p_0 > \sqrt{2C/b}$ with $\gamma_0 = 1$, the rational predictor will be chosen and the steady state obtained. But for initial points not mapped onto steady state in the next period, there are no paths to steady state. To show that there are no other paths to E^* consider an initial point $E_1 = (p_1, \zeta_1)$ such that $p_1 \neq \bar{p}$ and $p_0 < (B + b(1+\zeta_1)/2)/(B+b)\sqrt{2C/b}$ of FPED. Without loss of generality suppose that E_1 is one iteration from being mapped onto E^* i.e. E_1 is the last iterate on a path to E^*. This implies that $E_1 = (p_1, 1)$ since $p^{1,e}_{M,t+1} = \bar{p}$ if and only if $\zeta_t = 1$. But if $\zeta_1 = 1$, then m_{t-1} and $\zeta_{t-1} = 1$ because $\zeta_t = \zeta_{t-1} + \gamma(m_{t-1} - \zeta_{t-1})$, $0 < \gamma < 1$, and m_{t-1} can only take on the values of 1 and -1. And if $\zeta_{t-1} = 1$ and $m_{t-1} = 1$, then $p_1 = \bar{p}$ which is a contradiction. Therefore, there does not exist a path from some $E \neq E^*$ to E^*. Q.E.D.

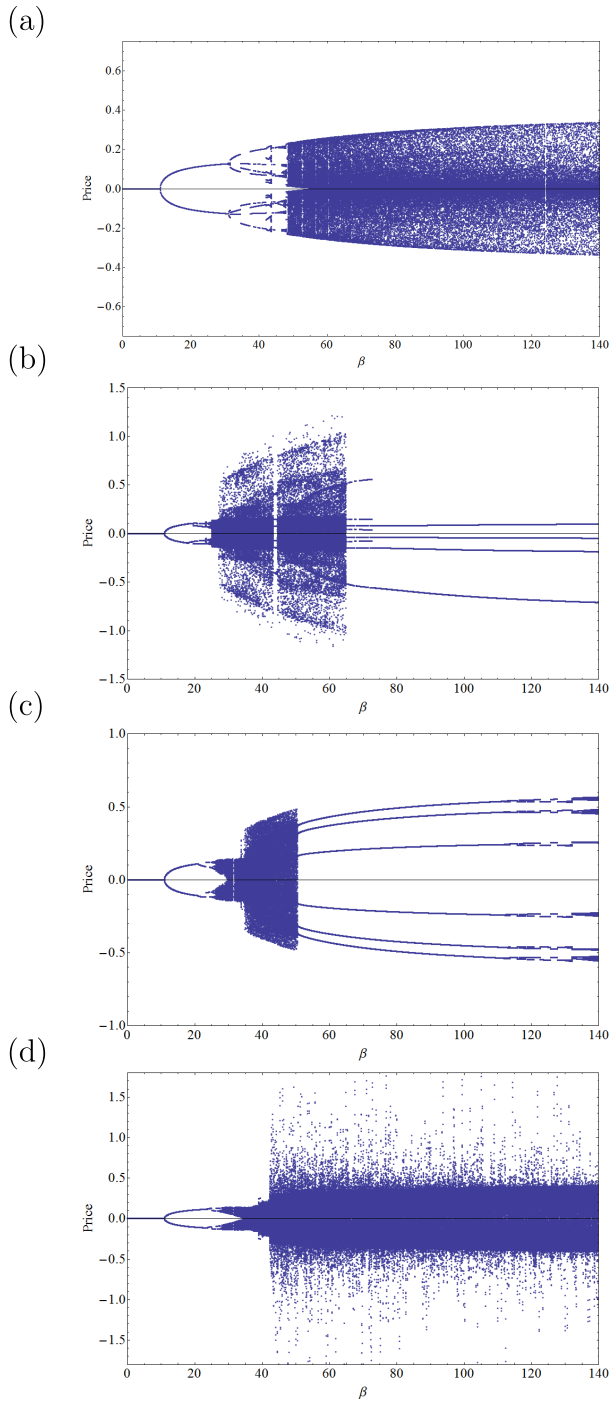

FIGURE 1. Bifurcation diagrams with respect to β. (a) is ARED. (b) is FPED with $\gamma = 0.95$. (c) is FPED with $\gamma = 0.5$. (d) is FPED with $\gamma = 0.1$. The other parameters of the model are $A = 0$, $B = 1$, $b = 2$, $C = 0.1$.

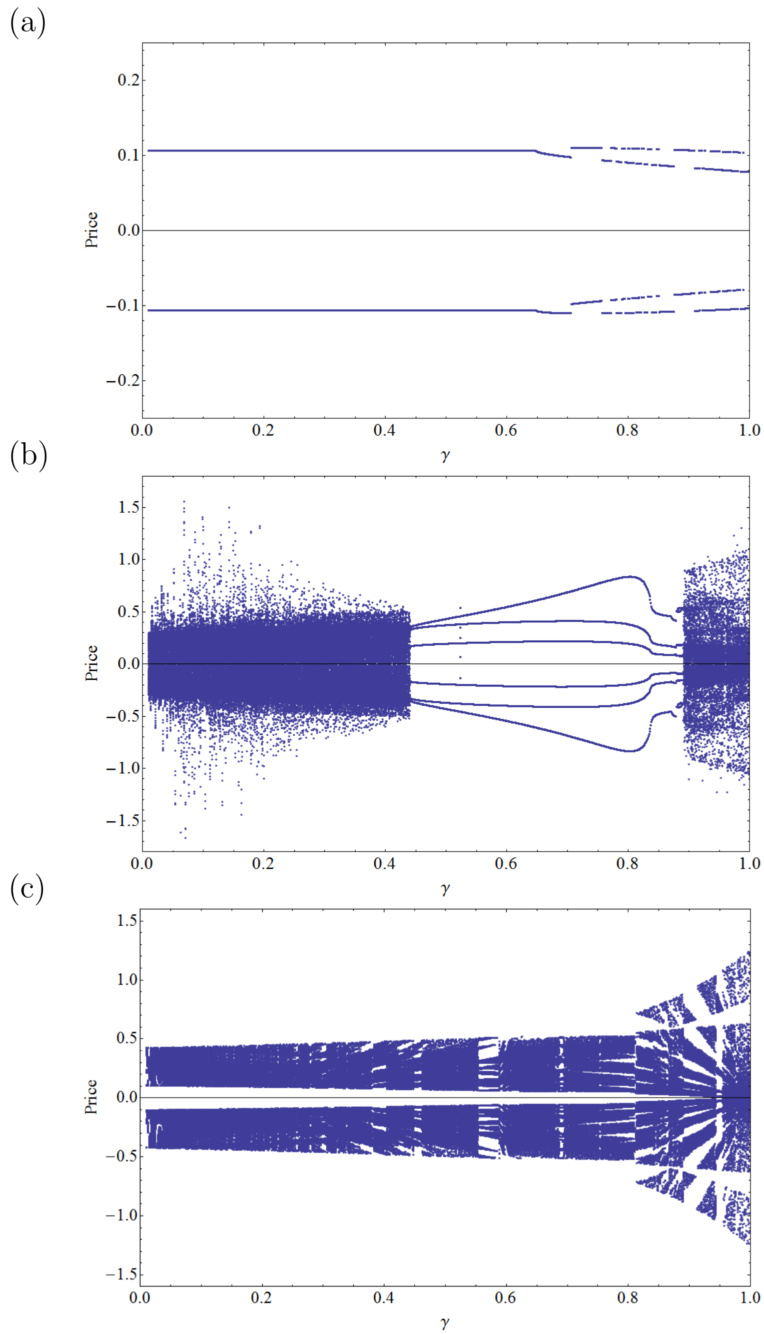

FIGURE 2. Bifurcation diagrams with respect to γ. (a) is FPED with $\beta = 20$. (b) is FPED with $\beta = 60$. (c) is FPED with $\beta = +\infty$. The other parameters of the model are $A = 0$, $B = 1$, $b = 2$, $C = 0.1$.

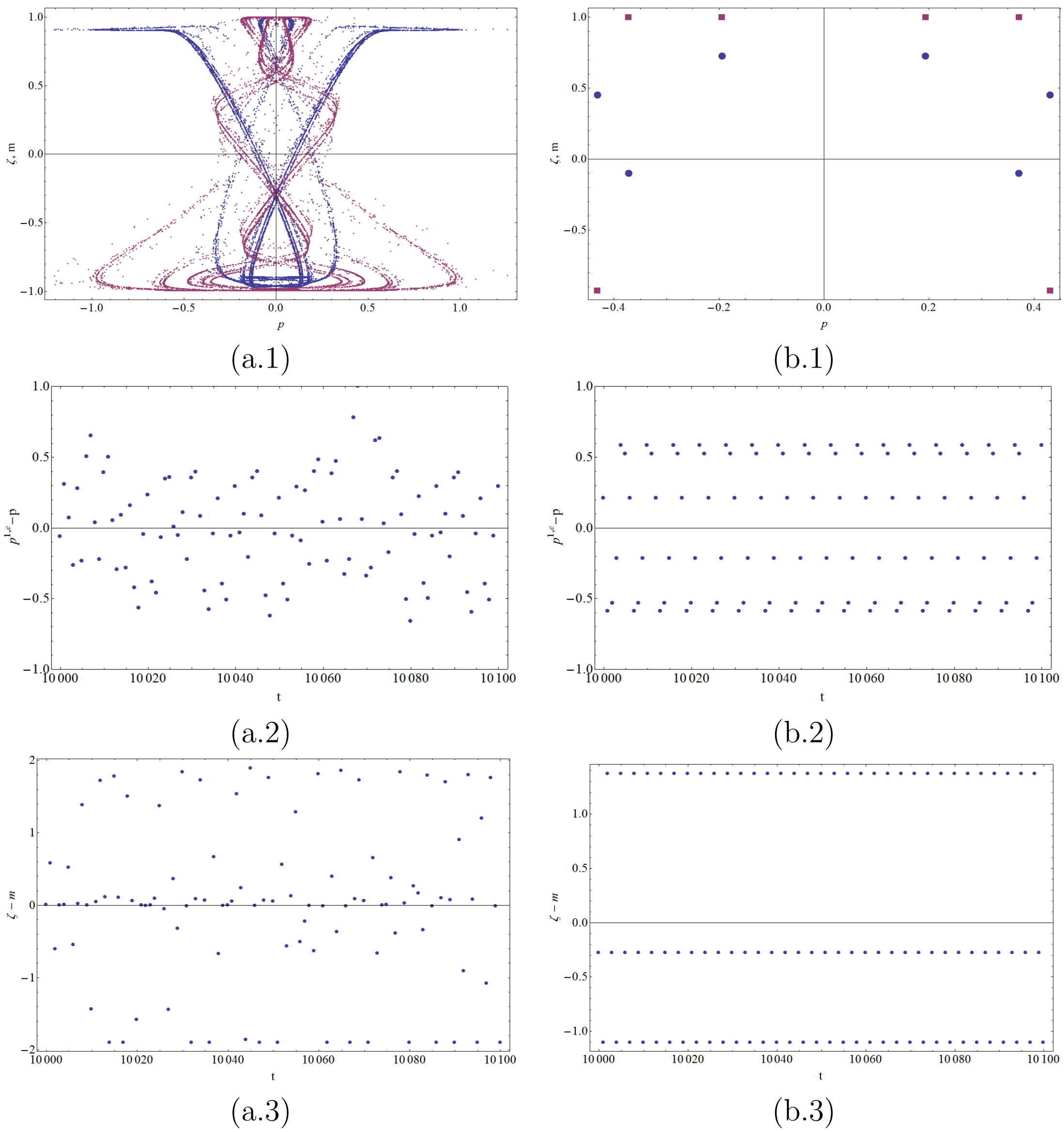

FIGURE 3. The forecast errors of the fictitious play learning rule (2.13) and the expert managers' predictions for $\beta = 60$, $A = 0$, $B = 1$, $b = 2$, $C = 0.1$ and $\gamma = 0.95$ (column (a)) and $\gamma = 0.5$ (column (b)). (a.1) and (b.1.) are the attractors of FPED. Both (p_t, m_t) and (p_t, ζ_t) are plotted. In (b.1) the squares correspond to (p_t, ζ_t).

CHAPTER III

THE INFLATION GAP, FORECAST COMBINATION, AND THE FORECAST COMBINATION PUZZLE

Introduction

Short horizon forecasts of economic measure of the macroeconomy are key inputs into the decision making process of monetary policymakers. The uncertainty surrounding individual forecasts and the plethora of forecasting methods that exist to construct these short horizon forecasts prompt many monetary authorities to consider combined or pooled forecasts.[1] The natural question that arises when considering combined forecasts is how the forecasts should be combined to create an optimal or best forecast.

This paper proposes a new dynamic forecast procedure to create combined forecasts of inflation that is shown to consistently and robustly outperform other common forecast combination strategies in root mean squared forecast error (RMSFE) in real-time out-of-sample forecasting exercises. The proposed method combines recent results from the literature on inflation persistence with research that explains the stylized facts that exist in studies of empirical forecasts to create a simple dynamic combination procedure. The combination procedure proposed is general and potentially could improve the forecast efficiency for combined forecasts for any time series of interest, but I restrict the analysis to inflation so as to draw upon a large literature of empirical observations that can explain the increased efficiency obtained.

[1] Robertson (2000) details the different forecasting strategies employed by a subset of central banks and describes how many explicitly use combined forecasts.

Forecast Combination

Forecast combination or pooling of forecasts is a well-known approach to improve forecast efficiency dating back to the seminal paper of Bates and Granger (1969). This paper launched an entire subfield of econometrics dedicated to explaining the efficiency gains observed from combining forecasts and to devise new schemes to improve efficiency. Surveys of the literature are found in Clemen (1989), Diebold and Lopez (1996), and Timmermann (2006).

A common empirical finding in this literature, including with forecast of inflation (see Stock and Watson (2003)), is that the simple average of the considered forecasts consistently produces the most efficient combined forecasts. The average typically outperforms each individual forecast considered and forecasts generated by more sophisticated combination procedures. The result is found despite the fact that under standard assumptions about the data, equal weights (averaging) is only theoretically optimal under very restrictive assumptions. This empirical finding is called by Stock and Watson (2004), "the forecast combination puzzle."

Hendry and Clements (2004) provide details of a generally accepted hypothesis for the existence of the forecast combination puzzle, which is that equal weights provide a hedge against unanticipated structural breaks and misspecification errors that affect the individual forecasts' efficiencies differently. The majority of the sophisticated weighting strategy proposed in the literature attempt to use past data to choose the combination weight for each forecast. However, the existence of the aforementioned breaks and errors implies that the historical performance of a forecast may not be a guide to its future performance. Employing equal weights is superior in these scenarios because the weights are agnostic about the best forecasts and do not shift weight to forecasts that have performed well in past, leading to suboptimal forecasts in the future when efficiencies change.[2]

The literature studying inflation persistence in the U.S. confirms the existence of both frequent structural breaks (Levin and Piger (2006)) and time-varying misspecification in popular econometric forecasts (Stock and Watson (2010)). However, the literature also finds thats the structural breaks and time-varying misspecification are not necessarily unpredictable.

The Inflation Gap

Cogley, Primiceri, and Sargent (2010) and Stock and Watson (2007) both propose parsimonious univariate models of inflation that decompose inflation into a stochastic trend, a serially uncorrelated transitory component, and which allow for stochastic volatility (UC-SV) to study how inflation persistence has changed over time. These studies focus on a measure termed by Cogley, Primiceri, and Sargent (2010) as the inflation gap. The inflation gap is the deviation of inflation from its estimated stochastic trend. These studies find evidence that the inflation gap is predictable using real activity measures, such as GDP and unemployment. Stock and Watson (2010) argue further that the inflation gap can be used to create improved forecasts of inflation.

The inflation gap is the forecast error from the optimal univariate forecast of inflation generated by the UC-SV model. Therefore, one interpretation of the literature's findings is that the forecast errors of parsimonious forecasting model, which are consistently found to produce the best forecasts of inflation, have time-

[2]A second hypothesis is based on the practical implementation of combination strategies that require additional estimation from past data. The additional parameters add either new sources estimation uncertainty that lead to lower forecast efficiency as found in Smith and Wallis (2009) or contribute to model overfitting.

varying structures that are predictable using real activity measures or other relevant explanatory variables. I explore the possibility that the predicted efficiency could be used to construct intercept corrections of the forecasts similar to those studied in Stock and Watson (2010) or used to pick weights to create an efficient combined forecast. I find that creating combined forecast using the predicted forecast errors of a menu of candidate forecasts can significantly and robustly produce efficient forecasts of inflation that outperform equal weights forecasts and other simple combination strategies in real-time out-of-sample forecasting exercises.

The concept is demonstrated by selecting a list of parsimonious candidate forecast models commonly used in the inflation forecasting literature. I show that the real-time forecast errors generated by the candidate models are serially correlated and can be predicted using real activity measures. I then use the predictions to create combined real-time forecasts for four different measures of inflation.

Predicting Forecast Errors

Stock and Watson (2007 and 2010) consider the following tightly parameterized UC-SV model for inflation:

$$\pi_t = \tau_t + \eta_t, \text{ where } \eta_t = \sigma_{\eta,t}\zeta_{\eta,t} \quad (3.1)$$

$$\tau_t = \tau_{t-1} + \epsilon_t, \text{ where } \epsilon_t = \sigma_{\epsilon,t}\zeta_{\epsilon,t} \quad (3.2)$$

$$\ln\sigma^2_{\eta,t} = \ln\sigma^2_{\eta,t-1} + \psi_{\eta,t} \quad (3.3)$$

$$\ln\sigma^2_{\epsilon,t} = \ln\sigma^2_{\epsilon,t-1} + \psi_{\epsilon,t}, \quad (3.4)$$

where $\zeta_t = (\zeta_{\eta,t}, \zeta_{\epsilon,t})$ is i.i.d. $N(0, I_2)$, $\psi_t = (\psi_{\eta,t}, \psi_{\epsilon,t})$ is i.i.d $N(0, \gamma I_2)$, and ζ_t and ψ_t are independently distributed, and γ is a scalar parameter. The optimal univariate time-t forecast of this model is $\tau_{t|t}$ for all $h \geq 1$ and the inflation gap at time $t+h$ can be defined as $a_{t+h} = \pi_{t+h} - \tau_{t|t}$. Stock and Watson posit that the inflation gap is predictable using real activity measures of the economy such that

$$\pi_{t+h} - \tau_{t|t} = \omega_h x_t + e_{t+h}, \tag{3.5}$$

where x_t is vector of real activity measures and e_{t+h} is the error term.

The justification for their supposition is that if the stochastic volatility in τ_t is ignored, then τ_t can be written as

$$\tau_t = (1-\theta)\sum_{i=1}^{\infty} \theta^i \pi_{t-i} \tag{3.6}$$

and the inflation gap can be expressed as

$$\pi_{t+h} - \pi_t = -\theta \sum_{i=1}^{\infty} \theta^i \Delta\pi_{t-i} + \omega_h x_t + e_{t+h}, \tag{3.7}$$

by substituting (3.6) into (3.5). Stock and Watson note that equation (3.7) is just a tightly parameterized backward-looking Phillips curve, which has been a staple of the inflation forecasting literature for decades. They show that this specification is able to augment and improve the forecasts from the UC-SV model compared to other common forecast methods in pseudo out-of-sample forecasts.

I expand this idea to consider the possibility that the forecast errors of any forecast of inflation are predictable using real activity measures and lags of inflation. If this is true, then the predicted forecast errors can be used to augment the original forecast to improve forecast efficiency or it can be used as a measure of forecast accuracy to rank a set of candidate forecasts. For the purposes of this paper, I consider the forecast error and the inflation gap to be synonymous.

Correcting and Ranking Forecasts

Consider a set of candidate forecasts for inflation based on time-t information $\hat{\pi}_{i,t+h|t}$ for $i = 1, 2, ...n$. Let the inflation gap of the i^{th} forecast be given by

$$\hat{a}_{i,t+h} = \pi_{t+h} - \hat{\pi}_{i,t+h|t}. \tag{3.8}$$

An intercept corrected forecast using the predicted inflation gap is given by

$$\hat{\pi}^{c}_{i,t+h|t} = \hat{\pi}_{i,t+h|t} + \hat{a}_{i,t+h}, \tag{3.9}$$

where $\hat{a}_{i,t+h}$ is added to correct for the presumed bias in the intercept. The correction's efficacy depends on the absolute accuracy of the point forecast, $\hat{a}_{i,t+h}$. The prediction must correctly discriminate the direction and magnitude of the gap to increase the overall forecast efficiency.

A combined forecast is a weighted sum of the candidate forecasts

$$\hat{\pi}^{fc}_{t+h} = \sum_{i=1}^{n} \gamma_i \hat{\pi}_{i,t+h}, \tag{3.10}$$

where γ_i denotes the weight given to the i^{th} forecast. I determine the weights by ranking the forecasts using the square of the inflation gap and assigning the largest weights to those forecasts with the smallest predicted gap. Recall that the inflation gap is a forecast error, so that this ranking implies high weights are assigned to the forecasts with the lowest predicted error. Note that this creates an important distinction between corrected forecasts and the combined forecasts. The combined forecast identifies off the relative magnitudes of the predicted inflation gaps, while the correction depends on both magnitude and direction.

I determine the weights using the multinomial logit functional form which is popular in the discrete-choice econometric literature,

$$\gamma_i = \frac{e^{-\beta a^2_{i,t+h}}}{Z_t}, \; Z_t = \sum_{i=1}^{n} e^{-\beta a^2_{i,t+h}}, \tag{3.11}$$

where β is the shrinkage or intensity of choice parameter that governs the relative weight given to each model for predicted differences in the inflation gap.[3] The β parameter is treated as constant in the analysis of this paper, but it could be estimated as well.

Candidate Forecasts and Data

Data

The forecast exercises I conduct are on the real-time data sets provided by the Philadelphia Federal Reserve and ALFRED. A real-time data set is a panel data set with multiple observations of the a given time series that reflects the actual vintages of data that were available to a forecaster on the middle month of a given quarter.[4] I consider four different measures of inflation. The measures are headline CPI spanning the time period 1947Q1-2011Q2 with real-time data starting in 1994Q2, the Price Index for Personal Consumption Expenditure (PCE) spanning the time period 1947Q1-2011Q1 with real-time data starting in 1965Q4, the Core

[3]The seminal reference in this literature is Manski and Mcfadden (1981).

[4]For further explanation see Croushore and Stark (2001).

Price Index for Personal Consumption Expenditure (Core PCE) spanning the time period 1959Q2-2011Q1 with real-time data starting in 1996Q1, and GDP Deflator (GDPDEF) spanning the time period 1947Q1-2011Q1 with real-time data starting in 1991Q1. Quarterly inflation is defined as

$$\pi_t = ln(\frac{P_t}{P_{t-1}}), \tag{3.12}$$

where P_t is quarterly price index at time t.[5]

I consider two real activity measures, GDP and unemployment, to predict inflation and inflation gaps. I select these two measures because of their common use in constructing Phillips curve inspired forecasts. This data is also from the Philadelphia Federal Reserve's Real-Time data sets. The real GDP (RGDP) spans the time period 1947Q1-2011Q1 with real-time data starting in 1965Q4 and the civilian unemployment rate (CUR) spans the time period 1948Q1-2011Q2 with real-time data starting in 1965Q4.

Real GDP is transformed into two different measures. The first measure is log differenced RGDP,

$$\Delta y_t = ln(\frac{y_t}{y_{t-1}}). \tag{3.13}$$

The second measure uses the Hodrick-Prescott filter to calculate an estimate of the cycle component of RGDP to create a measure of the output gap.[6] This series is called y_t^{gap}.

[5]For price indices that are recorded at a monthly frequency P_t is the three month average.

[6]The smoothing parameter is $\lambda = 1600$. Filters of this type have well known end-of-sample problems for providing accurate measures of the cycle component and thus can be a poor real-time forecast tools. Despite this fact, the gap measure does well in out-sample forecasting compared to the other measures considered in this paper.

The civilian unemployment rate is transformed into a single gap measure. The measure is a one-sided gap developed by Stock and Watson (2010) which only captures increases in unemployment compared to the minimum level of unemployment observed over the current and previous 10 quarters[7]

$$u_t^{gap} = u_t - min\{u_t, u_{t-1}, ..., u_{t-10}\}. \tag{3.14}$$

The measure is meant to capture the non-linearity in Phillips curve specifications where inflation only responds to increases in the unemployment rate.

Candidate Forecasts

I propose a short list of candidate forecast models to test the proposed forecast strategies. The list of models is given in Table 1. The list of candidate forecasts is small because of the recommendations of Granger and Jeon (2004), Aiolfi and Timmermann (2006), and Jose and Winkler (2008) who suggest trimming the list of candidate forecasts to include only a subset of the best performing models to maximize the forecast efficiency of a combined forecast.

The list of candidate forecasts model was constructed by choosing models that are found to be hard to beat in the literature or which would provide a good robustness check for the proposed forecasting strategy. The fourth column of Table 1 gives the reason the model was included. A reference to a paper indicates that the model was recommended by the paper or was the benchmark model considered. For example, the ARMA(1,1) is the benchmark econometric forecast specification used by Ang,

[7]Stock and Watson (2010) use a 12 quarter measure, while this measure is 11 quarters. The time frame I chose comes from the fact that the average peak-to-peak cycle over the postwar period is 66 months. So our measure covers exactly half of the average length of post-war peak-to-peak cycles.

Type	Abr.	Specification	Recommendation
ARMA	ARMA11	ARMA(1,1)	Ang, Bekaert, and Wei (2007)
ARMA	AR4	AR(4)	Marcellino, Stock, and Watson (2006)
Random Walk	RW	$\pi_t = \pi_{t-1} + \epsilon_t$	Atkeson and Ohanian (2001)
VAR	VAR1	VAR(2) with $X_t' = [\pi_t \; y_t^{gap}]'$	Stock and Watson (1999)
VAR	VAR2	VAR(2) with $X_t' = [\pi_t \; u_t^{gap}]'$	Robustness
ARMA	AR1	AR(1)	Robustness
ARMA	AR2	AR(2)	Robustness
ARMA	ARMA44	ARMA(4,4)	Robustness
Direct Forecast	DF1	$\pi_{t,4} = c + \phi_1 \pi_{t-4,4} + \theta_1 y_{t-1}^{gap} + \epsilon_{t,4}$	Robustness
Direct Forecast	DF2	$\pi_{t,4} = c + \phi_1 \pi_{t-4,4} + \theta_1 u_{t-1}^{gap} + \epsilon_{t,4}$	Robustness
Direct Forecast	DF3	$\pi_{t,4} = c + \phi_1 \pi_{t-4,4} + \theta_1 \Delta y_{t-1} + \epsilon_{t,4}$	Robustness
VAR	VAR3	VAR(3) with $X_t' = [\pi_t \; \Delta y_t \; u_t^{gap} \; y_t^{gap}]'$	Robustness

TABLE 1. Candidate forecasts models.

Bekaert, and Wei (2007), who tested dozens of different forecast specifications covering surveys, ARMA models, regressions using real-activity measures, and term structure models and found that a majority of models are unable to improve upon this simple specifications. Similarly, Atkeson and Ohanian (2001) show that a random walk specification systematically outperform better specified Phillips curve specifications during the 1990's.

The "Robustness" notation in the recommendation column denotes that the forecast model is included because either it is found to forecast well or it attempts to allay a concern about the proposed inflation forecasting strategy. The three ARMA models with this moniker fall into the former category, while the remaining models are chosen to include different combinations of the information that is used to predict the inflation gap. If the inflation gap is predictable, then the forecast model is misspecified and the forecast may be improved simply by adding this information into the original forecast model. These forecast accommodate this criticism to show that the gain in forecast efficiency cannot be achieved simply by correcting for omitted variables in the usual way.

Forecast Definitions and Individual Forecasts

The forecast horizon of interest for this paper is four quarters. I employ the definition used by Ang, Bakaert, and Wei (2007):

$$\hat{\pi}_{t+4,4} = E_{t-1}[\sum_{i=1}^{4} \pi_{t+i}], \tag{3.15}$$

where the forecast is the sum of the expected quarterly rates of inflation. The expectation is dated $t-1$ because with real-time data the current quarterly rate is unknown. This definition is nice because it is a proxy for the forecast efficiency for all $h \leq 4$ forecasts.

The benchmark forecast in this paper is an equal weights combined forecast of all considered models. Table 2 presents the out-of-sample RMSFE of the candidate forecasts and the relative RMSFE compared to the benchmark for the four measures of inflation. The real-time RMSFE of a forecast is calculated as

$$RMSFE_i = \sqrt{\frac{1}{M-k}\sum_{t=k}^{M} \pi_{T,t} - \hat{\pi}_{i,t}} \tag{3.16}$$

where k is the first available vintage of real-time data, T is most recent vintage of data, M is equal to $T-4$, $\pi_{T,t}$ is the actual inflation at time t from vintage T data, and $\hat{\pi}_{i,t}$ is the i^{th} forecast of inflation at time t using vintage t data. The results in Table 2 illustrate the forecast combination puzzle with the equal weights forecast outperforming the overwhelming majority of individual forecasts considered.

Predicting Inflation Gaps

The identifying assumption for the forecasts of inflation in this paper is that forecast errors are predictable. This section proposes model specifications for the

	CPI		GDPDEF	
Models	RMSFE	Relative	RMSFE	Relative
AR1	0.0181	1.1220	0.0108	1.3204
ARMA11	0.0175	1.0867	0.0083	1.0214
AR2	0.0180	1.1206	0.0090	1.0938
AR4	0.0152	0.9415	0.0084	1.0305
ARMA44	0.0151	0.9373	0.0083	1.0168
RW	0.0168	1.0437	0.0078	0.9540
DF1	0.0183	1.1337	0.0089	1.0901
DF2	0.0166	1.0303	0.0090	1.0989
DF3	0.0150	0.9345	0.0076	0.9256
VAR1	0.0202	1.2529	0.0093	1.1357
VAR2	0.0185	1.1500	0.0088	1.0712
VAR3	0.0204	1.2684	0.0095	1.1621
Equal Weights	0.0161	1.0000	0.0082	1.0000
Sample Period	1994Q2-2010Q2		1991Q4-2010Q2	

	Core PCE		PCE	
Models	RMSFE	Relative	RMSFE	Relative
AR1	0.0065	1.2022	0.0199	1.1065
ARMA11	0.0057	1.0492	0.0174	0.9619
AR2	0.0059	1.0911	0.0177	0.9839
AR4	0.0056	1.0766	0.0184	1.0223
ARMA44	0.0057	1.0319	0.0192	1.0643
RW	0.0050	0.9348	0.0171	0.9475
DF1	0.0081	1.5098	0.0223	1.2369
DF2	0.0080	1.4827	0.0206	1.1428
DF3	0.0051	0.9348	0.0205	1.1353
VAR1	0.0074	1.3712	0.0198	1.0977
VAR2	0.0067	1.2385	0.0182	1.0096
VAR3	0.0079	1.4651	0.0204	1.1315
Equal Weights	0.0054	1.0000	0.0180	1.0000
Sample Period	1996Q1-2010Q2		1971Q4-2010Q2	

TABLE 2. RMSFE for the candidate forecasts of inflation compared to an equal weights combined forecast of the same candidate forecast models for the four different measures of inflation. Numbers larger than 1 indicate that the equal weights forecast has a lower RMSFE over the relevant period.

inflation gaps to create inflation gap predictions. The forecast errors for annual inflation are given by

$$a_{i,t+4,4} = \pi_{t+4,4} - \hat{\pi}_{i,t+4,4}. \tag{3.17}$$

I consider two different specifications for predicting (3.17). The first is inspired by the autoregressive-distributed lag model given by equation (3.7) and posits that the inflation gap can be predicted by lags of inflation, a real activity measure, and is first order autoregressive,

$$\begin{aligned} a_{i,t+4,4} &= \omega_0 + \omega_1 \pi_t + \omega_2 \pi_{t-1} + \omega_3 \pi_{t-2} + \omega_4 x_t + v_t \\ v_t &= \phi v_{t-1} + \epsilon_{t,4}, \end{aligned} \tag{3.18}$$

where x_t is a real activity measure and $\epsilon_{t,4} \overset{\text{iid}}{\sim} N(0, \sigma^2)$. The second specification is geared towards the annual inflation definition given by equations (3.15). The definition impart an MA(3) structure to $a_{i,t+4,4}$ due to the summation of multiple quarterly forecasts. Therefore, the specification posits that the inflation gaps can be predicted by real activity measures and follows an MA(3) process,

$$\begin{aligned} a_{i,t+4,4} &= \omega_0 + \omega_1 x_{t-1} + v_t \qquad (3.19) \\ v_t &= \theta_3 \epsilon_{t-3,4} + \theta_2 \epsilon_{t-2,4} + \theta_1 \epsilon_{t-1,4} + \epsilon_{t,4} \end{aligned}$$

Equation (3.18) and (3.19) are estimated by maximum likelihood (MLE) using the Kalman Filter. Table 3 and 4 (the remaining tables and figures appear in the supplementary materials section) gives the coefficient estimates for these model on a selection of inflation gaps produced by the candidate forecasts. The table shows that the real activity measures are significant for most inflation gap series considered.

Real-Time Forecasts

The construction of real-time forecasts requires the sample data to be separated into three subsets to produce out-of-sample forecasts. The required divisions are

1. a training subset to estimate the parameters of the candidate forecast models
2. an in-sample forecast subset to recursively forecast the candidate models to construct series of inflation gaps to estimate the inflation gap models
3. and an out-of-sample subset to conduct out-of-sample forecasts.

The sample divisions are dictated by the availability of real-time data for each of the four measures of inflation considered. Table 6 details the sample periods.

Intercept Correction

A natural use of the predicted inflation gap is to add the prediction back to the original point forecast. This is known in the literature as intercept correction. Traditional intercept correction methods suggest adding the most recent forecast error to the model to correct for the deviation in the intercept. Turner (1990), Wallis and Whitney (1991), and Clements and Hendry (1998) report that intercept correction can result in modest improvement in forecast efficiency.

In this paper I go beyond traditional intercept correction by modeling the forecast error dynamics and predicting the path of forecast errors out to the same horizon as the original forecast. This procedure should capture any dynamics in the forecast errors that are detectable in past data. The intercept corrected forecasts are constructed using definition (3.9).

Table 5 shows a representative sample of the relative forecast efficiencies obtained by intercept correction. The forecast errors are predicted using equation (3.18) with Δy_t as the real activity measure. The results are presented relative to an equal weights combined forecast of the uncorrected models. The values marked with asterisks denote intercept corrected forecasts that have improved forecast efficiency. About half of the models considered improve forecast efficiency, but none improve significantly. The row marked "Equal Weights (IC)" combines all forecasts and predicted forecast errors together. The row shows that even in aggregate, the intercept correction approach does not increase relative efficiency compared to equal weights forecasts made with the uncorrected models. The results are consistent across different specifications for modeling forecast error dynamics and are consistent with the previous findings in the literature.

Combined Forecasts

In this section, instead of correcting forecasts, I use predicted inflation gaps to construct dynamics weights to create a combined forecast. The real-time out-of-sample combined forecasts for $\pi_{t+4,4}$ are made recursively using the following procedure at each time period t:

1. Each candidate forecast model is estimated on vintage t data and used to construct the set of n forecasts denoted $\{\hat{\pi}_{i,t+4,4}\}_{i=1}^{n}$

2. The set of past forecasts dated t and earlier is used to construct n time series of inflation gaps using vintage t inflation data.

3. A set of predictions for the inflation gaps $\{\hat{a}_{i,t+4,4}\}_{i=1}^{n}$ is estimated.

4. The set of forecasts $\{\hat{\pi}_{i,t+4,4}\}_{i=1}^{n}$ are combined using $\{\hat{a}_{i,t+4,4}\}_{i=1}^{n}$ according to equation (3.10).

$$\hat{\pi}_{t+4,4} = \sum_{i=1}^{n} \frac{e^{-\beta(\hat{a}_{i,t+4,4})^2}}{Z_t} \hat{\pi}_{i,t+4,4} \tag{3.20}$$

Table 6 shows the results for the RMSFE of the combined forecasts for $\beta = 2,000$ for predicting the inflation gap and for the four measure of inflation. The predicted inflation gap combined forecast results in lower RMSFE than the equal weights forecast in all but five out of twenty-eight specifications.[8]

Shrinkage Parameter

The shrinkage or intensity of choice parameter governs the deviation of the predicted inflation gap combined forecast (PIGC) from an equal weights forecast. A $\beta = 0$ collapses the combined forecast to an equal weights forecast, while larger values of the intensity of choice parameter cause a larger proportion of the weight to be placed on the individual candidate forecasts with the lowest predicted squared forecast error. A known result in the literature, shown by Clemen and Winkler (1986), Deibold and Pauly (1990), and Stock and Watson (2004), is that shrinking the combination weights towards equals weight increases forecast efficiency. This implies that lower β should result in lower RMSFE. The predicted inflation gap combined forecasts do not follow this trend.

Figure 4 plots the time path of PCE inflation, the combined forecast of PCE, and an equal weights forecast of PCE for the entire out-of-sample forecast period using

[8]The chosen β parameter is small for the scaling of the data considered. The β parameter is scalable by the units of $a_{i,t}$. The $a_{i,t}$ is scaled so that 1% is 0.01. If instead 1% is 1, then the corresponding β is 0.2.

four different values of β.[9] The time path of the weights generated by PIGC are shown along the bottom of the figure. Note that as the intensity of choice parameter is increased, the weights become more volatile, indicating that PIGC is placing more weight on the best predicted model. The increased dependence on the predicted best model causes PIGC to perform better, relative to equal weights. The relative RMSFE in plots "A-D" is 0.9668, 0.9447, 0.9303, and 0.9214, respectively. The plots also reveal that time variation in the weights is not uniform throughout the sample period. PIGC appears to be increasing efficiency greatly during the late 1970's and early 1980's, which corroborates the findings in Cogley, Primiceri, and Sargent (2010) for predicting the inflation gap.

To explore the relationship observed in Figure 4 more thoroughly, I conduct comparisons between the combined forecast and equal weights by varying both the sample period and the intensity of choice parameter. For PCE, I start by forecasting the sample period 1971Q4-1978Q1 and calculating the relative RMSFE obtained for values of β between 0 and 50,000. This allows for the relative RMSFE to plotted as a smooth curve against the values of β. Then, I increase the forecast sample period by five quarters to 1971Q4-1979Q2 and repeat the same exercise. I continue to repeat the same routine until all available data is exhausted. This results in 27 out-of-sample test periods with relative efficiency plotted as smooth curve dependent on β. I summarize the data by plotting the RMSFE against β for the mean, minimum, and maximum relative RMSFE observed out of the 27 sample periods. By plotting the mean, minimum, and maximum relative RMSFE a sense for the distribution of possible forecast efficiency, irrespective of the out-of-sample time period, is gained. Figure 5 and 6 show the results for PCE in the third column.

[9] Equation (3.18) is used to predict the forecast errors with y_t^{gap} acting as the real activity measure.

The black lines indicate the mean, minimum, and maximum relative RMSFE observed for the different sample forecast periods for PIGC forecasts. The gray lines indicate the mean, minimum, and maximum relative RMSFE for the limiting case of PIGC forecasts where β is infinite and all weight is placed on single forecast in each period. The same exercise is conducted with CPI, GDPDEF, and Core PCE starting with the initial sample ending in 2001Q4 and each successive sample growing by 5 quarters until the data is exhausted to create 8 out-of-sample periods. Each column of the figure represents a different measure of inflation and each row denotes a different real activity measure used in estimating equation (3.18) or (3.19). The 28 figures overwhelmingly show that on average predicted inflation gap combined forecasts outperforms equal weights.

Discussion

The main criticism to any out-of-sample forecast exercise is the possibility of data mining. Since the pseudo forecaster possess all the data, it is possible to tailor a sample test period specifically to generate a positive result. Figures 5 and 6 are designed to overcome this criticism. The figures shows how data mining is possible with some sample periods considered resulting in little or no increase in relative forecasting performance, while others result in 10% or greater gains. However, ultimately figures 5 and 6 demonstrates that this criticism does not apply to PIGC.

Figures 5 and 6 also demonstrates that PIGC is an outlier in regards to the forecast combination puzzle. The forecast combination puzzle implies that the further a combination method moves away from equal weights in an attempt to account for the time variation in forecast efficiency, the less efficient a forecast should become. PIGC displays the exact opposite behavior, with larger β's resulting in lower MSFE forecasts, indicating that the predicted forecast errors are providing an accurate prediction of future performance. The most striking illustration of this is the gray lines of figures 5 and 6 that represent the minimum, maximum, and mean relative forecasts produced when a weight of 1 is placed on the model that is predicted to yield the lowest RMSFE in each period. The RMSFE produced in these cases is on average as good as equals weights or better, even though placing all weight on a single model eliminates the theoretical advantages of combining forecasts. It can also be shown that if the limiting case is compared to picking a model at random to forecast in each period, then PIGC produces a forecast with on average 30% lower RMSFE when both are compared to equal weights.

Forecast Tournament

This section evaluates the out-of-sample forecast performance of the inflation gap combined forecast compared to four other combination methods employed frequently in the literature as well as equal weights. I tests the strategy's ability to combine different subsets of the 12 candidate forecasts models to produce efficient forecasts to create a robust comparison of the methods. A robust forecast combination strategy should be able to produce the best forecast possible given any set of models. I consider all possible combination of models of size greater than two, which results in 4,083 unique sets of candidate forecasts.

I compare predicted inflation gap combined forecasts to regression weights (GRW), MSFE weights (SW), backward-looking weights (BL) and equal weights (EW). The regression weights method I employ follows Granger and Ramanathan (1984) to obtain theoretically optimal weights for unbiased forecasts. The regression weights are calculated by estimating the following equation

$$\pi_{t,4} = \sum_{i=1}^{n} \gamma_{i,t}\hat{\pi}_{i,t,4} + \epsilon_{t,4}.$$
$$s.t. \sum_{i=1}^{n} \gamma_{i,t} = 1. \tag{3.21}$$

The equation is estimated using constrained least squares. The MSFE weights follow Stock and Watson (2001 and 2004) and weight candidate forecasts according to their relative past mean squared forecast error using

$$\gamma_{i,t} = \frac{(1/MSFE_{i,t})^k}{\sum_{i=1}^{n}(1/MSFE_{i,t})^k}, \tag{3.22}$$

where $MSFE_{i,t} = (1/m)\sum_{\tau=t-m}^{t} e_{i,\tau-4,4}^2$ and k is a shrinkage parameter.[10] The MSFE weights are one of the simple combination procedures that, like equal weights, consistently beats more sophisticated weighting procedures such as regression weights.

The backward-looking weights is a weighting strategy that acts as a robustness measure for PIGC. It uses the last known inflation gap to construct relative weights for the candidate forecast models. The weights are formed using

$$\gamma_{i,t} = \frac{e^{-\beta a_{i,t-4,4}^2}}{\sum_{i=1}^{n} e^{-\beta a_{i,t-4,4}^2}}, \tag{3.23}$$

where β is the same intensity of choice parameter used for the inflation gap weights. It provides a benchmark measure for the amount of forecast efficiency that is gained by forecasting the inflation gap of each model. If there is no useful information in iterating out the path of the forecast errors and all gains are made simply by

[10]For the forecast tournament I choose $k = 2$.

including the ranking information in the last known inflation gap, then this measure should perform as well as the predicted measure.

The predicted inflation gap forecasts for the tournament are estimated using equations (3.18) with Δy_t as the real activity measure and $\beta = 20,000$ for the intensity of choice parameter. Note that Δy_t and $\beta = 20,000$ do not maximize the absolute performance gains found in figures 5 or 6. The parameters are instead chosen because of the general pattern that arises in figures 5 and 6, which show that including a real activity measures increase forecast efficiency and that a large, but finite β also increases forecast efficiency.

The five forecast combination methods are used to construct real-time out-of-sample forecasts of inflation for all 4,083 sets of candidate models for the sample periods given in Table 6. The methods are ranked 1 to 5 for their performance in four categories for each set of candidate forecasts of size n. The four categories are the average RMSFE, the variance of the RMSFE, the minimum RMSFE, and the maximum RMSFE recorded for the out-of-sample forecast period. The average and minimum RMSFEs show the absolute forecasting ability of each method, while the variance of the RMSFEs and maximum RMSFE represent metrics that are important to the risk adverse forecaster by showing the min-max efficiency. The ranking is from low to high in all categories with 1 assigned to the method with lowest observed value.

Tables 7, 8, 9, and 10 shows the individual results for all five methods. The mean, minimum, and maximum results are presented relative to an equal weights forecast of all 12 candidate models. The standard deviation results are for the RMSFE of all the point forecasts produced by a given method, combining n models. The far right columns display the ranking of each method in each category. Table 11 shows the cumulative results for each method across the four categories. The forecast tournament shows that the predicted inflation gap combined forecast is the best forecast procedure considered.

The tournament also demonstrates the appeal of predicted inflation gap combinations to a risk adverse forecaster. If a forecaster is worried about *ex ante* selecting the worst set of forecasts to combine, then this method has a clear advantage because it consistently produces a low standard deviation of point forecasts indicating that it is robust to the inclusion of poor performing forecasts. It also produces on average the lowest maximum MSFE. The lowest maximum is an important benchmark because it shows the worst case scenario under each combination strategy. Since forecast combination is primarily a technique to hedge the risk of using a single forecast model, this is a desirable property of a combination technique.

Conclusion

This paper uses the explanation for the forecast combination puzzle and empirical results from the inflation persistence literature to create dynamic forecast weights that can robustly beat an equal weights forecast. The effectiveness of the proposed procedure demonstrates a promising way to construct efficient combined forecasts. The analysis presented in this paper, however, represent only a proof of concept. It demonstrates the existence of exploitable information in real-time out-of-sample forecast errors, but does so only on inflation. The success of the method demonstrates a promising path for future research into efficient combined forecasts.

Supplementary Materials

This section contains the remaining figures and tables that are commented in the chapter.

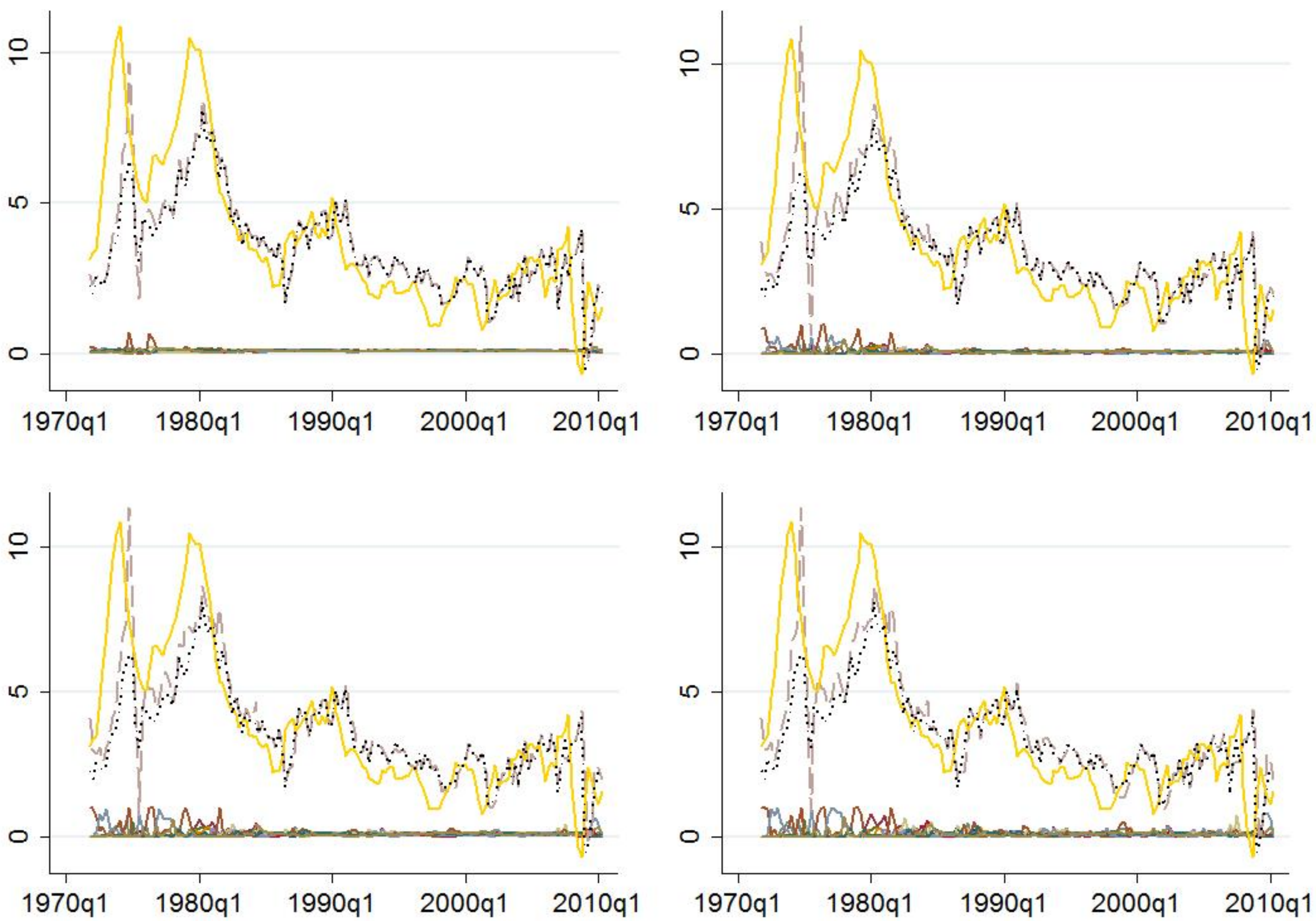

FIGURE 4. PCE inflation (solid line), the PIGC forecast of PCE (dashed line), the equal weights forecast of PCE (dotted line), and the time-varying weights of PIGC (solid lines on the bottom of each figure). Each panel is created using a different intensity of choice parameter for PIGC. A: $\beta = 2,000$, B: $\beta = 10,000$, C: $\beta = 20,000$, and D: $\beta = 50,000$.

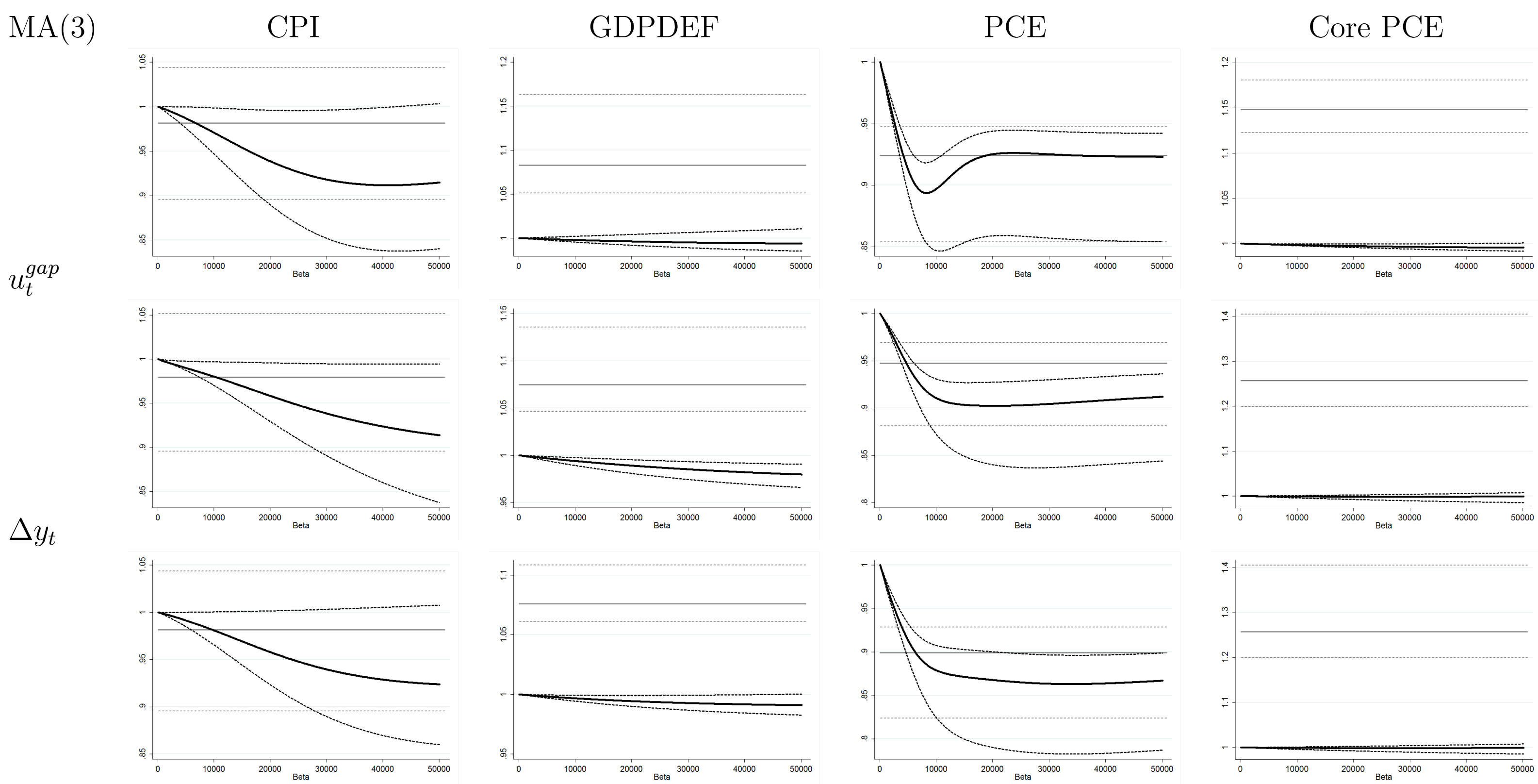

FIGURE 5. Mean, minimum, and maximum relative RMSFE for all out-of-sample forecasts period plotted against the value of intensity of choice parameter used to construct the PIGC forecast. The left-hand column denotes the real activity measure used in estimation equation (3.19).

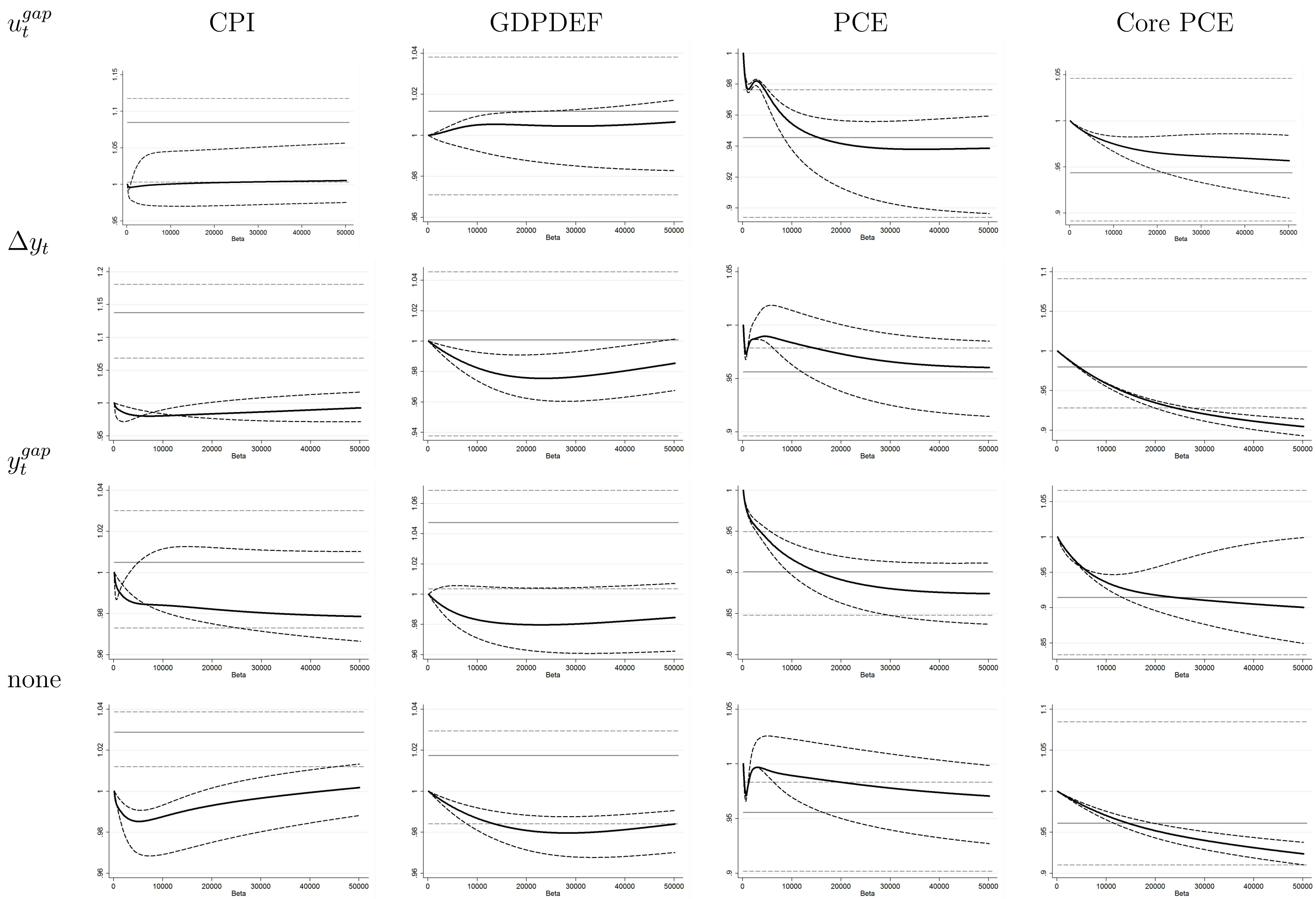

FIGURE 6. Mean, minimum, and maximum relative RMSFE for all out-of-sample forecasts period plotted against the value of intensity of choice parameter used to construct the PIGC forecast. The left-hand column denotes the real activity measure used in estimation equation (3.18).

Parameter Estimates for the Model of Forecast Error Dynamics						
	CPI			GDPDEF		
	AR1-FE	DF1-FE	VAR3-FE	AR2-FE	DF3-FE	VAR1-FE
v_{t-1}	0.9729***	0.9687 ***	0.9707***	0.9669***	0.9400***	0.9397***
	(0.0146)	(0.0181)	(0.0159)	(0.0189)	(0.0326)	(0.0273)
π_{t-1}	-2.8113***	-1.2332***	-2.7822***	-0.4837***	-.7150***	-1.7752***
	(0.1143)	(0.1154)	(0.1523)	(0.1241)	(0.1032)	(0.1524)
π_{t-2}	-0.4274***	-0.7366***	-0.9319***	-1.3372***	-0.5888***	-0.7491***
	(0.1053)	(0.0875)	(0.0899)	(0.1279)	(0.1148)	(0.1948)
π_{t-3}	-0.0910	-0.3968***	-0.6654***	-0.4056***	-0.3078**	-0.1625
	(0.1033)	(0.1199)	(0.1445)	(0.1237)	(0.1292)	(0.1904)
u^{gap}_{t-1}	-0.0063***			-0.0046***		
	(0.0017)			(0.0010)		
y^{gap}_{t-1}		-0.0001***			4.95e-6	
		(9.6e-6)			(6.12e-6)	
Δy_{t-1}			-0.4214***			-0.2465***
			(0.0583)			(0.0602)
c	0.0458**	0.0300	0.0538*	0.0244*	0.0124**	-0.0227*
	(0.0231)	(0.0191)	(0.0299)	(0.0142)	(0.0061)	(0.0124)
Rel. AIC	0.855	0.953	0.839	0.935	1.005	0.950
Obs.	170	170	170	155	155	155
	PCE			Core PCE		
	ARMA11-FE	DF2-FE	VAR2-FE	AR4-FE	ARMA44-FE	RW-FE
v_{t-1}	0.9768***	0.9696***	0.9695***	0.9759***	0.9673***	0.9653***
	(0.0151)	(0.0155)	(0.0177)	(0.0175)	(0.0202)	(0.0232)
π_{t-1}	-0.7769***	-0.9801***	-2.3986***	-0.8238***	-0.7752***	-1.0546***
	(0.1284)	(0.1079)	(0.1014)	(0.1301)	(0.1429)	(0.1279)
π_{t-2}	-1.8654***	-0.7058***	-1.0049***	-1.8895***	-1.7052***	-0.9416***
	(0.1031)	(0.1107)	(0.1589)	(0.1459)	(0.1608)	(0.1309)
π_{t-3}	-0.5236	-0.4539	-0.0591	-0.6415***	-0.3679**	-0.6761
	(0.1263)	(0.1184)	(0.1559)	(0.1631)	(0.1561)	(0.1374)
u^{gap}_{t-1}	-0.0042			-0.0033***		
	(0.0011)			(0.0010)		
y^{gap}_{t-1}		-2.35e-6			1.2e-5	
		(7.96e-6)			(8.30e-6)	
Δy_{t-1}			-0.0741			-0.0206
			(0.0490)			(0.0335)
c	0.0339	0.0239	0.0328***	0.0320	0.02162**	0.01989
	(0.0224)	(0.0171)	(0.0197)	(0.0203)	(0.0169)	(0.0144)
Rel. AIC	0.905	1.006	0.890	0.949	0.941	1.013
Obs.	179	179	179	155	155	155

TABLE 3. This table presents the parameter estimates for the model of forecast error dynamics proposed in Section 4, equations (3.18). The table includes only a selection of the forecast error series. Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1)

Parameter Estimates for the Model of Forecast Error Dynamics

	CPI			GDPDEF		
	AR1-FE	DF1-FE	VAR3-FE	AR2-FE	DF3-FE	VAR1-FE
θ_1	0.748***	1.149***	0.865***	1.135***	0.967***	0.639***
	(0.0634)	(0.0573)	(0.0778)	(0.0661)	(0.0453)	(0.0782)
θ_2	0.341***	1.214***	0.529***	0.738***	0.949***	0.455***
	(0.0719)	(0.0740)	(0.0883)	(0.0921)	(0.0622)	(0.0755)
θ_3	0.381***	1.160***	0.343***	0.294***	0.793***	0.247***
	(0.0684)	(0.0769)	(0.0751)	(0.0777)	(0.0569)	(0.0549)
u_t^{gap}		0.00272*			-0.00145*	
		(0.00151)			(0.00079)	
Δy_t			-0.336***			-0.178
			(0.130)			(0.0651)
c	0.00325	0.0278	0.00926**	-0.00299	-0.00246	-0.00193
	(0.00329)	(0.00328)	(0.00376)	(0.00189)	(0.00167)	(0.00176)
Rel. AIC	1.000	0.999	0.995	1.000	0.999	0.998
Obs.	170	170	170	155	155	155

	PCE			Core PCE		
	ARMA11-FE	DF2-FE	VAR2-FE	AR4-FE	ARMA44-FE	RW-FE
θ_1	1.329***	1.395***	0.952***	1.175***	1.20***	1.163***
	(0.0634)	(0.0598)	(0.0727)	(0.0661)	(0.0726)	(0.0525)
θ_2	0.747***	1.411***	0.555***	0.786***	0.748***	1.104***
	(0.0719)	(0.0787)	(0.0813)	(0.0921)	(0.0957)	(0.0627)
θ_3	0.216***	1.104***	0.369***	0.283***	0.213***	0.818***
	(0.0684)	(0.0566)	(0.0696)	(0.0736)	(0.0712)	(0.0569)
u_t^{gap}		0.00101*			-0.00246**	
		(0.00104)			(0.00116)	
Δy_t			-0.0734			-0.003
			(0.0825)			(0.0249)
c	0.00369	0.00502	0.00526**	-0.00073	0.00138	-0.00153
	(0.00232)	(0.00218)	(0.00257)	(0.00182)	(0.00232)	(0.00147)
Rel. AIC	1.000	1.001	1.001	1.000	0.997	1.002
Obs.	179	179	179	155	155	155

TABLE 4. This table presents the parameter estimates for the model of forecast error dynamics proposed in Section 4, equations (3.19). The table includes only a selection of the forecast error series. Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1)

Intercept Corrected Models Compared to Equal Weights Forecast				
	CPI		GDPDEF	
Models	RMSFE	Relative	RMSFE	Relative
AR1	0.0162*	1.0082	0.0098*	1.1934
ARMA11	0.0169*	1.0524	0.0092	1.1185
AR2	0.0163*	1.0174	0.0096	1.1684
AR4	0.0187	1.1653	0.0092	1.1308
ARMA44	0.0179	1.1170	0.0093	1.1416
RW	0.0199	1.2415	0.0085	1.0469
DF1	0.0200	1.2476	0.0089	1.0102
DF2	0.0200	1.2447	0.0090	1.0989
DF3	0.0197	1.2299	0.0078	0.9242
VAR1	0.0161*	1.0012	0.0089*	1.0925
VAR2	0.0165*	1.0242	0.0102	1.2489
VAR3	0.0176*	1.0966	0.0123	1.5085
Equal Weights (IC)	0.0174	1.0837	0.0088	1.0081
Sample Period	1994Q2-2010Q2		1991Q4-2010Q2	
	Core PCE		PCE	
Models	RMSFE	Relative	RMSFE	Relative
AR1	0.0056*	1.4065	0.0192*	1.0663
ARMA11	0.0057	1.0634	0.0191	1.0615
AR2	0.0057*	1.0592	0.0190	1.0548
AR4	0.0057	1.0588	0.0184*	1.0218
ARMA44	0.0060	1.1238	0.0203	1.1264
RW	0.0057	1.0591	0.0207	1.1453
DF1	0.0079*	1.4746	0.0180*	0.9976
DF2	0.0067*	1.2562	0.0186*	1.0331
DF3	0.0054	1.0049	0.0198*	1.1001
VAR1	0.0065*	1.2228	0.0191*	1.0612
VAR2	0.0057*	1.0601	0.0186	1.0307
VAR3	0.0070*	1.2972	0.0201*	1.1177
Equal Weights (IC)	0.0053	0.9796	0.0183	1.0161
Sample Period	1996Q1-2010Q2		1971Q4-2010Q2	

TABLE 5. Real-time out-of-sample intercept corrected forecasts of inflation. The values are compared to an equal weights combined forecast of the uncorrected models. The asterisks denote models that have improved forecast efficiency with intercept correction compared to Table 2.

Sample Subsets			
	Training Subset	In-sample Subset	Out-of-sample
CPI	1947Q2-1971Q3	1971Q4-1994Q1	1994Q2-2010Q2
GDPDEF	1947Q2-1971Q3	1971Q4-1990Q4	1991Q1-2010Q2
PCE	1947Q2-1965Q3	1965Q4-1971Q3	1971Q4-2010Q2
Core PCE	1947Q2-1971Q3	1971Q4-1996Q1	1996Q1-2010Q2

PIGC compared to Equal Weights

Model/x_{t-1}	CPI	GDPDEF	PCE	Core PCE
MA(3) / none	1.0001	1.0004	0.9691	0.9997
(3.19) / u_t^{gap}	0.9891	0.9987	0.9829	0.9986
(3.19) / Δy_t	0.9963	0.9997	0.9678	1.0002
(3.18) / u_t^{gap}	1.0225	0.9975	0.9823	0.9983
(3.18) / Δy_t	0.9716	0.9981	0.9854	0.9901
(3.18) / y_t^{gap}	0.9959	1.0041	0.9668	0.9734
(3.18) /none	0.9797	0.9978	0.9919	0.9939

TABLE 6. The top portion of the table denotes the data subsets chosen for the implementation of PIGC. The bottom of the table shows the relative RMSFE of PIGC compared to an equal weights combined forecast on the full range of out-of-sample date and for $\beta = 2,000$.

Forecast Tournament Results - Core PCE								
	n	Sets	EW	PIGC	GRW	BL	SW	Ranking
Mean	2	66	1.0952	1.0675	1.1477	1.1061	1.1199	2, 1, 5, 3, 4
	3	220	1.0584	1.0181	1.1507	1.0679	1.0878	2, 1, 5, 3, 4
	4	495	1.0394	0.9924	1.1638	1.0472	1.0695	2, 1, 5, 3, 4
	5	792	1.0278	0.9765	1.1818	1.0342	1.0577	2, 1, 5, 3, 4
	6	924	1.0200	0.9657	1.2067	1.0252	1.0494	2, 1, 5, 3, 4
	7	792	1.0143	0.9578	1.2403	1.0186	1.0433	2, 1, 5, 3, 4
	8	495	1.0101	0.9518	1.2828	1.0136	1.0386	2, 1, 5, 3, 4
	9	220	1.0067	0.9471	1.3323	1.0097	1.0348	2, 1, 5, 3, 4
	10	66	1.0040	0.9432	1.3851	1.0066	1.0318	2, 1, 5, 3, 4
	11	12	1.0018	0.9401	1.4355	1.0041	1.0292	2, 1, 5, 3, 4
	12	1	1.0000	0.9374	1.4788	1.0020	1.0271	2, 1, 5, 3, 4
Std. Dev.	2	66	0.1261	0.1342	0.1571	0.1435	0.1354	1, 2, 5, 4, 3
	3	220	0.0943	0.0952	0.1200	0.1084	0.1021	1, 2, 5, 4, 3
	4	495	0.0753	0.0735	0.0998	0.0871	0.0819	2, 1, 5, 4, 3
	5	792	0.0621	0.0594	0.0956	0.0723	0.0677	2, 1, 5, 4, 3
	6	924	0.0518	0.0491	0.1020	0.0609	0.0567	2, 1, 5, 4, 3
	7	792	0.0434	0.0408	0.1117	0.0514	0.0476	2, 1, 5, 4, 3
	8	495	0.0361	0.0337	0.1182	0.0429	0.0396	2, 1, 5, 4, 3
	9	220	0.0293	0.0273	0.1174	0.0350	0.0322	2, 1, 5, 4, 3
	10	66	0.0227	0.0210	0.1064	0.0272	0.0250	2, 1, 5, 4, 3
	11	12	0.0158	0.0146	0.0835	0.0190	0.0174	2, 1, 5, 4, 3
Minimum	2	66	0.8992	0.8699	0.9211	0.8803	0.9031	3, 1, 5, 2, 4
	3	220	0.8968	0.8607	0.9251	0.8703	0.8893	4, 1, 5, 2, 3
	4	495	0.8994	0.8731	0.9314	0.8859	0.9004	3, 1, 5, 2, 4
	5	792	0.9052	0.8839	0.9330	0.8983	0.9184	3, 1, 5, 2, 4
	6	924	0.9166	0.8924	0.9352	0.9103	0.9355	3, 1, 4, 2, 5
	7	792	0.9241	0.9002	0.9825	0.9248	0.9512	2, 1, 5, 3, 4
	8	495	0.9315	0.9063	1.0301	0.9359	0.9603	2, 1, 5, 3, 4
	9	220	0.9461	0.9127	1.0698	0.9502	0.9673	2, 1, 5, 3, 4
	10	66	0.9623	0.9205	1.1552	0.9653	0.9825	2, 1, 5, 3, 4
	11	12	0.9793	0.9278	1.2755	0.9843	1.0014	2, 1, 5, 3, 4
	12	1	1.0000	0.9374	1.4788	1.0020	1.0271	2, 1, 5, 3, 4
Maximum	2	66	1.4395	1.4431	1.5151	1.4633	1.4623	1, 2, 5, 4, 3
	3	220	1.3883	1.3855	1.4606	1.4139	1.4060	2, 1, 5, 4, 3
	4	495	1.3296	1.3142	1.4274	1.3431	1.3456	2, 1, 5, 3, 4
	5	792	1.2924	1.2462	1.5128	1.3110	1.3051	2, 1, 5, 4, 3
	6	924	1.2100	1.1467	1.5792	1.2379	1.2461	2, 1, 5, 3, 4
	7	792	1.1546	1.0988	1.6240	1.1841	1.1962	2, 1, 5, 3, 4
	8	495	1.1140	1.0577	1.6297	1.1390	1.1536	2, 1, 5, 3, 4
	9	220	1.0858	1.0404	1.5948	1.1113	1.1187	2, 1, 5, 3, 4
	10	66	1.0657	1.0169	1.5600	1.0927	1.0918	2, 1, 5, 4, 3
	11	12	1.0335	0.9753	1.5324	1.0544	1.0560	2, 1, 5, 3, 4
	12	1	1.0000	0.9374	1.4788	1.0020	1.0271	2, 1, 5, 3, 4

TABLE 7. Forecast tournament results for Core PCE. The n indicates the number of candidate forecast models being combined to form the forecast. The column "Sets" denotes the number unique groups of candidate forecasts of size n that are formed out of the 12 models. The "Ranking" column assign a 1 to lowest relative value in each row.

Forecast Tournament Results - CPI								
	n	Sets	EW	PIGC	GRW	BL	SW	Ranking
Mean	2	66	1.0397	1.0296	1.1463	1.0812	1.0556	2, 1, 5, 4, 3
	3	220	1.0240	1.0175	1.1797	1.0880	1.0423	2, 1, 5, 4, 3
	4	495	1.0161	1.0134	1.1921	1.0962	1.0353	2, 1, 5, 4, 3
	5	792	1.0113	1.0119	1.1968	1.1041	1.0310	1, 2, 5, 4, 3
	6	924	1.0081	1.0110	1.2000	1.1110	1.0281	1, 2, 5, 4, 3
	7	792	1.0058	1.0101	1.2045	1.1169	1.0260	1, 2, 5, 4, 3
	8	495	1.0041	1.0089	1.2110	1.1218	1.0244	1, 2, 5, 4, 3
	9	220	1.0027	1.0074	1.2202	1.1258	1.0232	1, 2, 5, 4, 3
	10	66	1.0016	1.0056	1.2334	1.1291	1.0222	1, 2, 5, 4, 3
	11	12	1.0007	1.0035	1.2520	1.1318	1.0214	1, 2, 5, 4, 3
	12	1	1.0000	1.0012	1.2787	1.1339	1.0207	1, 2, 5, 4, 3
Std. Dev.	2	66	0.0722	0.0638	0.1063	0.1075	0.0614	3, 1, 5, 4, 2
	3	220	0.0547	0.0464	0.0971	0.0986	0.0475	3, 1, 5, 4, 2
	4	495	0.0441	0.0354	0.0853	0.0870	0.0393	3, 1, 5, 4, 2
	5	792	0.0366	0.0274	0.0729	0.0745	0.0334	3, 1, 5, 4, 2
	6	924	0.0308	0.0217	0.0629	0.0622	0.0286	3, 1, 5, 4, 2
	7	792	0.0259	0.0176	0.0552	0.0504	0.0245	3, 1, 5, 4, 2
	8	495	0.0216	0.0146	0.0484	0.0394	0.0207	3, 1, 5, 4, 2
	9	220	0.0176	0.0121	0.0421	0.0291	0.0171	3, 1, 5, 4, 2
	10	66	0.0137	0.0100	0.0353	0.0197	0.0134	3, 1, 5, 4, 2
	11	12	0.0095	0.0075	0.0281	0.0112	0.0094	3, 1, 5, 4, 2
Minimum	2	66	0.9120	0.9274	0.9565	0.9176	0.9145	1, 4, 5, 3, 2
	3	220	0.9198	0.9267	0.9945	0.9174	0.9178	3, 4, 5, 1, 2
	4	495	0.9328	0.9318	0.9894	0.9229	0.9513	3, 2, 5, 1, 4
	5	792	0.9393	0.9376	1.0089	0.9321	0.9609	3, 2, 5, 1, 4
	6	924	0.9471	0.9505	1.0339	0.9415	0.9643	2, 3, 5, 1, 4
	7	792	0.9555	0.9590	1.0157	0.9535	0.9698	1, 3, 5, 2, 4
	8	495	0.9623	0.9656	1.0198	0.9631	0.9771	1, 3, 5, 2, 4
	9	220	0.9703	0.9720	1.0514	0.9729	0.9876	1, 2, 5, 3, 4
	10	66	0.9789	0.9893	1.1627	1.0520	0.9986	1, 2, 5, 3, 4
	11	12	0.9887	0.9953	1.2236	1.1029	1.0088	1, 2, 5, 4, 3
	12	1	1.0000	1.0012	1.2787	1.1339	1.0207	1, 2, 5, 4, 3
Maximum	2	66	1.2441	1.2132	1.3798	1.2868	1.2437	3, 1, 5, 4, 2
	3	220	1.2003	1.1884	1.3910	1.2905	1.1962	3, 1, 5, 4, 2
	4	495	1.1643	1.1499	1.3964	1.2872	1.1542	3, 1, 5, 4, 2
	5	792	1.1429	1.1335	1.4054	1.2842	1.1361	3, 1, 5, 4, 2
	6	924	1.1292	1.1156	1.4217	1.2791	1.1219	3, 1, 5, 4, 2
	7	792	1.1029	1.0824	1.4146	1.2605	1.1042	2, 1, 5, 4, 3
	8	495	1.0745	1.0490	1.3545	1.2348	1.0853	2, 1, 5, 4, 3
	9	220	1.0514	1.0352	1.3373	1.1969	1.0700	2, 1, 5, 4, 3
	10	66	1.0336	1.0289	1.3101	1.1709	1.0540	2, 1, 5, 4, 3
	11	12	1.0158	1.0218	1.3220	1.1466	1.0397	2, 1, 5, 4, 3
	12	1	1.0000	1.0012	1.2787	1.1339	1.0207	1, 2, 5, 4, 3

TABLE 8. Forecast tournament results for CPI. The n indicates the number of candidate forecast models being combined to form the forecast. The column "Sets" denotes the number unique groups of candidate forecasts of size n that are formed out of the 12 models. The "Ranking" column assign a 1 to lowest relative value in each row.

Forecast Tournament Results - GDP Deflator								
	n	Set	EW	PIGC	GRW	BL	SW	Ranking
Mean	2	66	1.0349	1.0366	1.0440	1.0175	1.0314	3, 4, 5, 1, 2
	3	220	1.0209	1.0215	1.0575	0.9951	1.0175	3, 4, 5, 1, 2
	4	495	1.0139	1.0130	1.0931	0.9818	1.0107	4, 3, 5, 1, 2
	5	792	1.0097	1.0073	1.1388	0.9722	1.0067	4, 3, 5, 1, 2
	6	924	1.0069	1.0031	1.1864	0.9646	1.0040	4, 2, 5, 1, 3
	7	792	1.0050	1.0000	1.2311	0.9583	1.0021	3, 4, 5, 1, 2
	8	495	1.0035	0.9975	1.2704	0.9529	1.0007	4, 2, 5, 1, 3
	9	220	1.0023	0.9954	1.3034	0.9482	0.9996	4, 2, 5, 1, 3
	10	66	1.0014	0.9937	1.3302	0.9441	0.9988	4, 2, 5, 1, 3
	11	12	1.0006	0.9922	1.3517	0.9404	0.9981	4, 2, 5, 1, 3
	12	1	1.0000	0.9909	1.3694	0.9370	0.9975	4, 2, 5, 1, 3
Std. Dev.	2	66	0.0775	0.0754	0.0669	0.0733	0.0672	5, 4, 1, 3, 2
	3	220	0.0629	0.0619	0.0765	0.0613	0.0520	4, 3, 5, 2, 1
	4	495	0.0527	0.0529	0.0929	0.0539	0.0422	2, 3, 5, 4, 1
	5	792	0.0447	0.0458	0.1072	0.0482	0.0352	2, 3, 5, 4, 1
	6	924	0.0382	0.0398	0.1172	0.0432	0.0296	2, 3, 5, 4, 1
	7	792	0.0325	0.0344	0.1224	0.0384	0.0250	2, 3, 5, 4, 1
	8	495	0.0274	0.0294	0.1224	0.0336	0.0209	2, 3, 5, 4, 1
	9	220	0.0225	0.0244	0.1165	0.0285	0.0170	2, 3, 5, 4, 1
	10	66	0.0176	0.0193	0.1034	0.0230	0.0132	2, 3, 5, 4, 1
	11	12	0.0123	0.0136	0.0810	0.0167	0.0092	2, 3, 5, 4, 1
Minimum	2	66	0.8784	0.8917	0.9066	0.8466	0.8956	2, 3, 5, 1, 4
	3	220	0.8799	0.8919	0.8964	0.8496	0.8949	2, 3, 5, 1, 4
	4	495	0.8982	0.9005	0.8917	0.8513	0.9135	2, 4, 3, 1, 5
	5	792	0.9138	0.9092	0.8879	0.8635	0.9305	4, 3, 2, 1, 5
	6	924	0.9246	0.9204	0.8880	0.8774	0.9430	4, 3, 2, 1, 5
	7	792	0.9348	0.9306	0.9100	0.8897	0.9541	4, 3, 2, 1, 5
	8	495	0.9475	0.9420	0.9316	0.8985	0.9631	4, 3, 2, 1, 5
	9	220	0.9579	0.9513	1.0097	0.9070	0.9703	3, 2, 5, 1, 4
	10	66	0.9683	0.9617	1.1058	0.9162	0.9797	3, 2, 5, 1, 4
	11	12	0.9783	0.9723	1.1642	0.9245	0.9872	4, 3, 5, 1, 2
	12	1	1.0000	0.9909	1.3694	0.9370	0.9975	4, 2, 5, 1, 3
Maximum	2	66	1.1985	1.2561	1.2393	1.1803	1.1861	3, 5, 4, 1, 2
	3	220	1.1495	1.1819	1.3182	1.1364	1.1393	3, 4, 5, 1, 2
	4	495	1.1263	1.1481	1.3815	1.1128	1.1141	3, 4, 5, 1, 2
	5	792	1.1107	1.1279	1.4878	1.0865	1.1006	3, 4, 5, 1, 3
	6	924	1.0920	1.1016	1.5148	1.0693	1.0786	3, 4, 5, 1, 4
	7	792	1.0781	1.0835	1.5450	1.0572	1.0649	3, 4, 5, 1, 5
	8	495	1.0674	1.0708	1.5489	1.0480	1.0540	3, 4, 5, 1, 6
	9	220	1.0582	1.0592	1.5403	1.0269	1.0466	3, 4, 5, 1, 7
	10	66	1.0477	1.0490	1.5369	1.0144	1.0398	3, 4, 5, 1, 8
	11	12	1.0303	1.0280	1.4639	0.9863	1.0213	4, 3, 5, 1, 2
	12	1	1.0000	0.9909	1.3694	0.9370	0.9975	4, 2, 5, 1, 3

TABLE 9. Forecast tournament results for GDP Def. The n indicates the number of candidate forecast models being combined to form the forecast. The column "Sets" denotes the number unique groups of candidate forecasts of size n that are formed out of the 12 models. The "Ranking" column assign a 1 to lowest relative value in each row.

Forecast Tournament Results - PCE								
	n	Set	EW	PIGC	GRW	BL	SW	Ranking
Mean	2	66	1.0318	1.0037	1.1216	1.0300	1.0180	4, 1, 5, 3, 2
	3	220	1.0190	0.9855	1.2185	1.0214	0.9960	3, 1, 5, 4, 2
	4	495	1.0126	0.9787	1.3622	1.0189	0.9837	3, 1, 5, 4, 2
	5	792	1.0088	0.9763	1.5174	1.0185	0.9757	3, 1, 5, 4, 2
	6	924	1.0063	0.9756	1.6546	1.0187	0.9701	3, 2, 5, 4, 1
	7	792	1.0045	0.9758	1.7658	1.0192	0.9659	3, 2, 5, 4, 1
	8	495	1.0031	0.9764	1.8592	1.0199	0.9626	3, 2, 5, 4, 1
	9	220	1.0021	0.9772	1.9390	1.0208	0.9600	3, 2, 5, 4, 1
	10	66	1.0013	0.9782	1.9877	1.0218	0.9578	3, 2, 5, 4, 1
	11	12	1.0006	0.9792	1.9423	1.0228	0.9560	3, 2, 5, 4, 1
	12	1	1.0000	0.9802	1.6601	1.0235	0.9545	3, 2, 5, 4, 1
Std. Dev.	2	66	0.0674	0.0425	0.1248	0.0448	0.0700	3, 1, 5, 2, 4
	3	220	0.0556	0.0274	0.2153	0.0305	0.0642	3, 1, 5, 2, 4
	4	495	0.0468	0.0199	0.2954	0.0245	0.0596	3, 1, 5, 2, 4
	5	792	0.0399	0.0159	0.3281	0.0219	0.0546	3, 1, 5, 2, 4
	6	924	0.0341	0.0128	0.3316	0.0205	0.0492	3, 1, 5, 2, 4
	7	792	0.0291	0.0103	0.3359	0.0194	0.0436	3, 1, 5, 2, 4
	8	495	0.0245	0.0081	0.3507	0.0181	0.0379	3, 1, 5, 2, 4
	9	220	0.0201	0.0060	0.3705	0.0165	0.0319	3, 1, 5, 2, 4
	10	66	0.0157	0.0040	0.3913	0.0142	0.0254	3, 1, 5, 2, 4
	11	12	0.0110	0.0019	0.4123	0.0110	0.0181	3, 1, 5, 2, 4
Minimum	2	66	0.8775	0.9296	0.9248	0.9527	0.8643	2, 4, 3, 5, 1
	3	220	0.9002	0.9284	0.9239	0.9589	0.8663	2, 4, 3, 5, 1
	4	495	0.9177	0.9306	0.9457	0.9645	0.8804	2, 3, 5, 4, 1
	5	792	0.9327	0.9410	0.9774	0.9678	0.8926	2, 3, 5, 4, 1
	6	924	0.9436	0.9478	0.9862	0.9706	0.9046	2, 3, 5, 4, 1
	7	792	0.9531	0.9499	1.0145	0.9747	0.9135	3, 2, 5, 4, 1
	8	495	0.9637	0.9544	1.2302	0.9788	0.9231	3, 2, 5, 4, 1
	9	220	0.9728	0.9581	1.3168	0.9821	0.9307	3, 2, 5, 4, 1
	10	66	0.9823	0.9609	1.4877	0.9981	0.9372	3, 2, 5, 4, 1
	11	12	0.9909	0.9764	1.5393	1.0057	0.9434	3, 2, 5, 4, 1
	12	1	1.0000	0.9802	1.6601	1.0235	0.9545	3, 2, 5, 4, 1
Maximum	2	66	1.1663	1.1042	1.4596	1.1366	1.1712	3 ,1, 5, 2, 4
	3	220	1.1444	1.0936	1.9071	1.1242	1.1494	3, 1, 5, 2, 4
	4	495	1.1183	1.0476	2.4570	1.1057	1.1204	3, 1, 5, 2, 4
	5	792	1.1015	1.0251	2.9710	1.0876	1.0897	4, 1, 5, 2, 3
	6	924	1.0908	1.0153	3.3534	1.0759	1.0792	4, 1, 5, 2, 3
	7	792	1.0760	1.0088	3.4842	1.0690	1.0632	4, 1, 5, 3, 2
	8	495	1.0644	1.0033	3.4546	1.0622	1.0515	4, 1, 5, 3, 2
	9	220	1.0548	0.9962	3.0778	1.0577	1.0416	3, 1, 5, 4, 2
	10	66	1.0434	0.9907	2.8351	1.0549	1.0264	3, 1, 5, 4, 2
	11	12	1.0323	0.9824	2.7105	1.0469	1.0129	3, 1, 5, 4, 2
	12	1	1.0000	0.9802	1.6601	1.0235	0.9545	3, 2, 5, 4, 1

TABLE 10. Forecast tournament results for PCE. The n indicates the number of candidate forecast models being combined to form the forecast. The column "Sets" denotes the number unique groups of candidate forecasts of size n that are formed out of the 12 models. The "Ranking" column assign a 1 to lowest relative value in each row.

Forecast Tournament Summary Results					
	EW	PIGC	GRW	BL	SW
PCE	3.00	**1.60**	4.91	3.28	2.21
Core PCE	2.07	**1.07**	4.98	3.21	3.67
CPI	2.05	**1.63**	5.00	3.56	2.77
GDPDEF	3.19	**3.07**	4.56	1.63	2.51
Mean	2.58	**1.84**	4.86	2.92	2.79

TABLE 11. The cumulative tournament ranking of each combination method on the four measures of inflation. Each entry is the average ranking for all subsets and categories.

CHAPTER IV

FORECAST COMBINATION IN THE MACROECONOMY

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 this justification the cognitive consistency principle.

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.

The standard implementation of econometric learning outlined in Evans and Honkapohja (2001) is to assume that all agents possess a single forecast model or perceived law of motion for the economy to form expectations. In the empirical practice of forecasting, however, econometricians often possess a menu of different subjective forecast models form which to choose. The menu of different models may 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 has two recourses when presented with multiple suitable models, she can select among the models by assigning a fitness measure to each or she can devise a way to combine them.

The forecasting literature has studied both solutions extensively and has consistently found that combination is the more robust and efficient solution to the model selection problem. The seminal paper demonstrating the result is Bates and Granger (1969). This paper spawned an entire subfield of econometrics dedicated to developing new forecast 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), and Branch and McGough (2008 and 2010). 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 interesting and relevant economic phenomena such as multiple equilibria, time-varying volatility, and exotic dynamics, which can match dynamics observed in actual economic data. However, It remains 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 this paper I propose a general framework to study homogeneous agents who consider a menu of different forecast models to form forecasts of endogenous state variables to determine the equilibrium and dynamic effects of employing forecast combination strategies to the model selection problem. The assumption of agents who use forecast combination strategies is more closely aligned with the spirit of the cognitive consistency principle because combined forecasts are favored by both practitioners and policymakers.[1] In addition, combined forecasts are some of the most widely distributed forecasts such as Survey of Professional Forecasters, the Michigan Inflation Expectations Survey, or the Blue Chip Consensus Forecasts, which are all either mean or median combined forecasts.

Contribution

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. 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 using a predetermined strategy. The resulting equilibrium beliefs and dynamics are explored. The forecast combination strategies are judged on their ability to obtain

[1]Monetary and fiscal policymakers typically rely on consensus or combined forecasts. As an example, Robertson (2000) compares the forecasting methods of different central banks including the Federal Reserve and the Bank of England and reports that in general they rely on combined forecasts for policy decisions.

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

or approximate the Rational Expectations Equilibrium (REE) in keeping with the traditional use of learning techniques as selection and robustness measures.

I use the equilibrium concept to study a popular reduced form macroeconomic model that is consistent with either a Lucas-type aggregate supply model of Lucas (1973) or a Muth cobweb model of Muth (1961) to demonstrate the possible congruence or divergence in outcomes that occurs compared to the standard learning model and rational expectations. I consider agents who form combined forecast using a weighted sum of the menu of forecasts they consider. The goal of the agents is to choose weights for the sum to produce an optimal forecast. I explore two basic strategies that agents can use to pick weights: an optimal weighting scheme that attempts to minimize a quadratic loss function and simple averaging (equal weights).

I show that different forecast combination strategies can result in fundamentally different equilibrium outcomes from one another and from rational expectations. I also show that under an the optimal weight combination strategy, where the weights become an endogenous variable, that there may exist multiple equilibra. In addition, if there exists multiple equilibria and the rational expectations equilibria is one of those equilibria, then the rational expectations equilibria is never learnable in the sense that if agents are estimating the parameters of their forecast models and the optimal weights, they will never converge to the REE from nearby initial points. However, many of the other equilibria of the model are shown to stable under learning.

The multiplicity of equilibria in the model can generate dynamics under real time econometric learning that are similar to the results found in the dynamic predictor selection literature. In particular, an economy that only is subject to small white noise shocks can exhibit stochastic volatility as in Brock and Hommes (1998) and Branch and Evans (2007) and endogenous structural change. The result is novel because the agents take into account the full history of past data when choosing weights, the weights are not confined to unit interval as in dynamic predictor selection where they represent population shares, and the agents hold homogeneous beliefs.

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 for a particular specification of the model under study. 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.

A General Framework

To fix ideas I present a macroeconomy that has a unique Rational Expectations Equilibrium (REE) 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.

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,

$$y_t = \mu + \alpha E_{t-1} y_t + \zeta' x_{t-1} + v_t, \tag{4.1}$$

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.[3] 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 rational expectations solution of this model can be represented as linear combination of the exogenous observable shocks:

$$E_{t-1} y_t = \phi' z_{t-1}, \tag{4.2}$$

where ϕ is a $(n+1) \times 1$ vector of coefficients that reflect agents beliefs about the effect of the shocks 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,

$$E z_{t-1}(y_t - \phi' z_{t-1}) = 0, \tag{4.3}$$

where E is the unconditional expectations operator and 0 is an $(n+1)$ vector of zeros. The unique beliefs that satisfy (4.3) and constitute a rational expectations equilibrium are $\phi = (1-\alpha)^{-1}(\mu \;\; \zeta')'$.

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

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 (4.2) that each omit one or more of the exogenous shocks in x_{t-1}. 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 assumption 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 and 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 forecasts using a weighted sum approach that is standard in the forecasting literature. The weighted sum of the k underparameterized model is given by

$$E_{t-1} y_t = \sum_{i=1}^{k} \gamma_i \phi_i' z_{i,t-1}, \tag{4.4}$$

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

Forecast Combination Equilibrium

The equilibrium concept I propose 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 for each model on their menu. Each model on the menu is required to be optimal given the information set used to create it. These individual forecasts are, however, not necessarily optimal given the total information set of the agents. The definition reflects the behavior of actual forecasters who optimally fit different misspecified models conditional on their included information and then combine the forecasts.

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

$$Ez_{i,t-1}(y_t - \phi_i' z_{i,t-1}) = 0 \tag{4.5}$$

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

The basic definition takes the weights as an exogenous variable. This is done so that different weights or different optimality requirements for weights can be studied under a single equilibrium definition. The selection of weights is a non-trivial problem in the empirical practice of forecasting and one of the questions of interest is how different exogenous or endogenous weighting strategies may alters equilibrium outcomes.

This equilibrium concept is related to another equilibrium definition proposed in the literature, the Misspecification Equilibrium (ME) concept developed in Branch and Evans (2006). An ME is employed to 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 where the weights are 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. However, the Forecast Combination Equilibrium concept is distinct from ME because in the case under study there is no heterogeneity in forecasts among agents choice and no restrictions on the value of the weights.

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 $Y_t = (y_{1,t}\ y_{2,t}\ \dots\ y_{k,t})' \in \mathbb{R}^k$ and they choose to combine them to form a single forecast using the weights $\Gamma = (\gamma_1\ \gamma_2\ \dots\ \gamma_k)' \in \mathbb{R}^k$. The expectation of the agents, $E_{t-1}y_t$, is given by equation (4.4) and the economy under the combined forecasts can be written as

$$y_t = \mu + \alpha \sum_{i=1}^{k} \gamma_i \phi_i' z_{i,t-1} + \zeta' x_{t-1} + v_t. \tag{4.6}$$

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 (4.15) given in Definition 1. This implies the beliefs of the agents represented by the ϕ_i's must satisfy the following system of equations

$$
\begin{aligned}
Ez_{1,t-1}(\mu + \alpha \sum_{i=1}^{k} \gamma_i \phi_i' z_{i,t-1} + \zeta' x_{t-1} + v_t - \phi_1' z_{1,t-1}) &= 0 \\
&... \\
Ez_{k,t-1}(\mu + \alpha \sum_{i=1}^{k} \gamma_i \phi_i' z_{i,t-1} + \zeta' x_{t-1} + v_t - \phi_k' z_{k,t-1}) &= 0.
\end{aligned} \tag{4.7}
$$

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.[4]

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

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

to choose. The next section studies different possible ways agents may choose weights based on recommendations from the forecasting literature.

Exogenous and Endogenous Selection of Weights

I consider the possibility that agents either exogenously impose recommended weights or endogenously choose weights according to an optimality criterion to form a 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.[5] Given a set of underparmeterized models Y_t, the forecast combination problem is

$$\min_{\{\Gamma\}} E[(y_t - \Gamma' Y_t)^2]. \tag{4.8}$$

The agents solve the minimization problem the same way they form a forecast by positing a form of the solution and finding optimal coefficients Γ. I study three different solutions to (4.8) proposed by Granger and Ramanathan (1984) who consider regression models for estimating the solution from past data. The three specifications

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

are

$$y_t = \gamma_1 y_{1,t} + \gamma_2 y_{2,t} + ... + \gamma_k y_{k,t} + e_t \tag{4.9}$$

$$y_t = \gamma_1 y_{1,t} + \gamma_2 y_{2,t} + ... + \gamma_k y_{k,t} + e_t \ : \ s.t. \sum_{i=1}^{k} \gamma_i = 1 \tag{4.10}$$

$$y_t = \gamma_0 + \gamma_1 y_{1,t} + \gamma_2 y_{2,t} + ... + \gamma_k y_{k,t} + e_t, \tag{4.11}$$

where e_t represents the error term. The first specification, equation (4.9), is argued by Granger and Ramanathan to be the optimal solution when the menu of forecasts is believed to be unbiased. This specification is the main one of interest and I will denote it as the *optimal weights* (OW) specification.

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 significantly reduce forecast efficiency.[6] The modification is also not

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

an obvious addition based on the objective function given by (4.8). Although, it is a more natural modification when considered in a regression framework.

In addition to the optimal weights, I consider the choice of static 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 (4.8) under very specific conditions, but it is found to work well in empirical 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.[7] The four different strategies are used to modify the FCE definition to incorporate the selection of weights as an equilibrium condition.

Assessing FCEs

Rational expectations is the natural benchmark for the characterization of FCEs formed under different combination strategies. The resulting FCEs 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.

[7] 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).

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 $E^{REE}_{t-1}y_t = E^{FCE}_{t-1}y_t$.

The notation $E^{REE}_{t-1}y_t = E^{FCE}_{t-1}y_t$ 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 (4.1) 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.

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

$$Ex'_{t-1}x_{t-1} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}. \tag{4.12}$$

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

$$y_{1,t} = a_1 + b_1 x_{1,t-1} \tag{4.13}$$

$$y_{2,t} = a_2 + b_2 x_{2,t-1}., \tag{4.14}$$

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.

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

$$Ez_{i,t-1}(y_t - \phi_i' z_{i,t-1}) = 0 \tag{4.15}$$

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

$$\begin{aligned} Ez_{1,t-1}(y_t - \phi_1' z_{1,t-1}) &= 0 \\ Ez_{2,t-1}(y_t - \phi_2' z_{2,t-1}) &= 0. \end{aligned} \tag{4.16}$$

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,

$$E_{t-1}y_t = \frac{1}{2}\phi_1' z_{1,t-1} + \frac{1}{2}\phi_2' z_{2,t-1}. \tag{4.17}$$

The PLM represents how agents form $E_{t-1}y_t$ in equation (4.1). The PLM can be substituted in for $E_{t-1}y_t$ to produce the actual law of motion (ALM) of the economy,

$$y_t = \mu + \alpha(\frac{1}{2}\phi_1' z_{1,t-1} + \frac{1}{2}\phi_2' z_{2,t-1}) + \zeta' x_{t-1} + v_t. \tag{4.18}$$

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 T-map under equal weights is

$$T\begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \mu + \frac{\alpha}{2}(a_1 + a_2) \\ \mu + \frac{\alpha}{2}(a_1 + a_2) \\ \frac{\alpha}{2}(b_1 + b_2\frac{\sigma_{12}}{\sigma_1^2}) + \zeta_1 \\ \frac{\alpha}{2}(b_2 + b_1\frac{\sigma_{12}}{\sigma_2^2}) + \zeta_2 \end{pmatrix}, \tag{4.19}$$

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

$$\begin{aligned} a_1 = a_2 &= \frac{\mu}{1-\alpha} \\ b_1 &= \frac{(\frac{\alpha}{2}-1)\zeta_1\sigma_1^2\sigma_2^2 - \frac{\alpha}{2}\zeta_1\sigma_{12}^2 - \zeta_2\sigma_{12}\sigma_2^2}{\frac{1}{4}\sigma_{12}^2\alpha^2 - (1-\frac{\alpha}{2})^2\sigma_1^2\sigma_2^2} \\ b_2 &= \frac{(\frac{\alpha}{2}-1)\zeta_2\sigma_1^2\sigma_2^2 - \frac{\alpha}{2}\zeta_2\sigma_{12}^2 - \zeta_1\sigma_{12}\sigma_1^2}{\frac{1}{4}\sigma_{12}^2\alpha^2 - (1-\frac{\alpha}{2})^2\sigma_1^2\sigma_2^2}, \end{aligned} \tag{4.20}$$

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 the reflected in agents beliefs in the EWFCE.

The combination weights drive a wedge between the EWFCE forecasts and the RE forecasts. Under RE, the expected squared forecast error of the agents is given by $Ev_t^2 = \sigma_v^2$, while the expected squared forecast error in the EWFCE is

$$E(y_t - \frac{1}{2}\sum_{i=1}^{2} y_{i,t})^2 = \sigma_v^2 + \xi_1\sigma_1^2 + \xi_2\sigma_2^2 + \xi_1\xi_2\sigma_{12}, \tag{4.21}$$

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 as expected will have higher expected squared forecast errors than under rational expectations.

Optimal Weights

The optimal weights case uses the regression specification (4.9) 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

$$\begin{aligned} EY_t(y_t - \Gamma' Y_t) &= 0 \\ Ez_{1,t-1}(y_t - \phi_1' z_{1,t-1}) &= 0 \\ Ez_{2,t-1}(y_t - \phi_2' z_{2,t-1}) &= 0, \end{aligned} \tag{4.22}$$

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

$$E_{t-1}y_t = \gamma_1 \phi_1' z_{1,t-1} + \gamma_2 \phi_1' z_{2,t-1}, \tag{4.23}$$

and the corresponding ALM is

$$y_t = \mu + \alpha(\gamma_1 \phi_1' z_{1,t-1} + \gamma_2 \phi_1' z_{2,t-1}) + \zeta' x_{t-1} + v_t. \tag{4.24}$$

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

$$T \begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \mu + \alpha(a_1\gamma_1 + a_2\gamma_2) \\ \mu + \alpha(a_1\gamma_1 + a_2\gamma_2) \\ \alpha\gamma_1 b_1 + (\alpha\gamma_2 + \zeta_2)\frac{\sigma_{12}}{\sigma_1^2} + \zeta_1 \\ \alpha\gamma_2 b_2 + (\alpha\gamma_1 + \zeta_1)\frac{\sigma_{12}}{\sigma_2^2} + \zeta_2 \\ \frac{\alpha\gamma_1 a_1^2 + b_1(\alpha\gamma_1\sigma_1^2 b_1 + (\alpha-1)\gamma_2\sigma_{12}b_2 + \sigma_1\zeta_1 + \sigma_{12}\zeta_2) + a_1((\alpha-1)\gamma_2 a_2 + \mu)}{(a_1^2 + \sigma_1^2 b_1^2)} \\ \frac{(\alpha-1)(\gamma_1 a_1 a_2 + \gamma_1\sigma_{12}b_1 b_2) + \alpha\gamma_2 a_2^2 + \alpha\gamma_2\sigma_2 b_2^2 + \zeta_1\sigma_{12}b_2 + \zeta_2\sigma_2^2 b_2 + \mu a_2}{(a_2^2 + \sigma_2^2 b_2^2)} \end{pmatrix}. \tag{4.25}$$

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

The T-map is a system of polynomial equations which suggest the potential for multiple equilibria to exist. However, obtaining analytic solutions to systems of polynomial equations is difficult (see Sturnfels (2002)). It is possible to solve this system analytically with the aid of computers, but the solutions are large and impractical to study. So 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 equilibrium. The possible equilibria that exist outside of this neighborhood are then 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

$$\dot{\Theta} = T(\Theta) - \Theta. \tag{4.26}$$

The differential equation governs the dynamics of $\Theta = (\phi_1' \; \phi_2' \; \Gamma')'$ in notional time.[8] 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 affect on the fundamental FCE is analyzed by using the center manifold reduction 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 (7).

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

[8] 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.

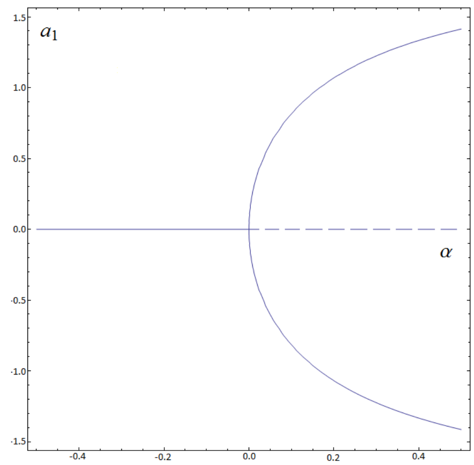

FIGURE 7. A pitchfork bifurcation on the approximate center manifold of the T-map (4.26). The bifurcation has been normalized to occur at $(0, 0)$. The solid line indicate stable fixed points and the dashed line indicates unstable fixed points.

outcomes that are equivalent to rational expectations. But, as α becomes large, optimal weights can also provide equilibrium outcomes that diverge from rational expectations.[9] The economic intuition for the existence of the non-fundamental OWFCEs when the feedback parameter on expectations is positive is that positive feedback 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 optimal weights to create new fixed points of the system.

[9]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.

To illustrate the multiple OWFCEs that exist graphically and show their dependence on α, I create a pseudo-bifurcation diagram. The diagram fixes a realization of $\hat{x}_{t-1}$ and plots the different equilibrium forecasts that the agents would hold for the set of OWFCES that exist for a given α. Figure 8 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.[10] 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.

Further Exploration

No fundamental OWFCE exists when either $\sigma_{12} \neq 0$ or $\mu \neq 0$. Figure 9 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 8. 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.

[10] 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 $\alpha = -2$, and Branch and Evans (2007) who use $\alpha = 0.6$.

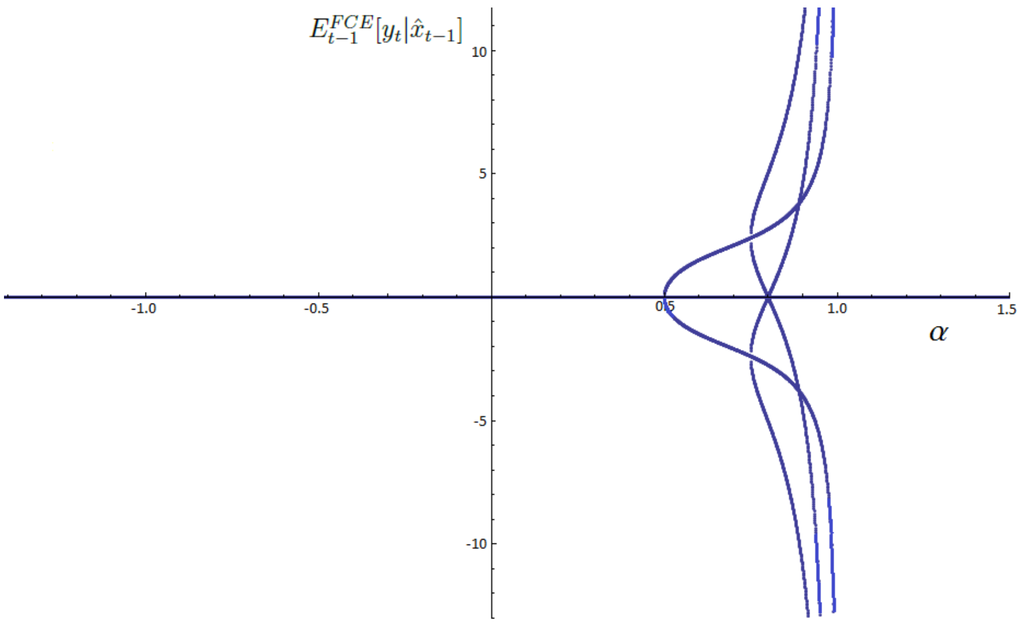

FIGURE 8. A plot of the OWFCE forecasts for a specific realization of x_{t-1} and for different values of α.

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. 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.

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 8. There now exists multiple OWFCEs for $\alpha > 1$. The maximum number of OWFCEs remains at seven.

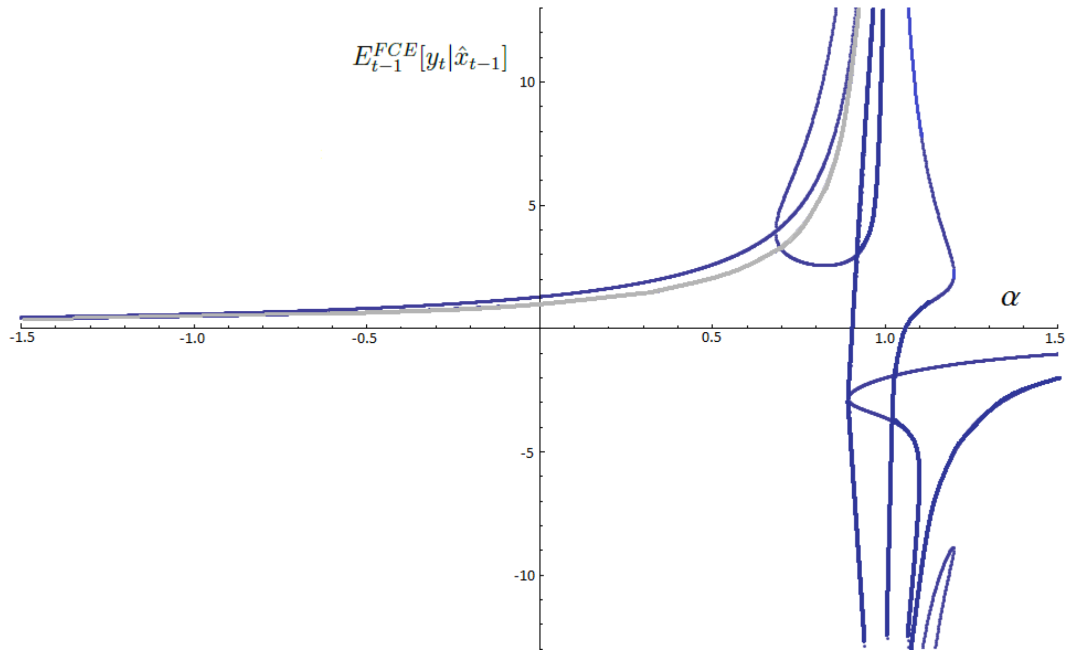

FIGURE 9. A plot of the OWFCE forecasts (black) and the REE forecast (gray) for a specific realization of x_{t-1} and for different values of α.

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

$$E(y_{1,t} - y_{2,t})[(y_t - y_{2,t}) - \gamma_1(y_{1,t} - y_{2,t})]. \tag{4.27}$$

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

$$T\begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \\ \gamma_1 \end{pmatrix} = \begin{pmatrix} \mu + \alpha\gamma_1 a_1 + \alpha(1-\gamma_1)a_2 \\ \mu + \alpha\gamma_1 a_1 + \alpha(1-\gamma_1)a_2 \\ \alpha\gamma_1 b_1 + (\zeta_2 + \alpha(1-\gamma_1)b_2)\frac{\sigma_{12}}{\sigma_1^2} + \zeta_1 \\ \alpha(1-\gamma_1)b_2 + (\zeta_1 + \alpha\gamma_1 b_1)\frac{\sigma_{12}}{\sigma_2^2} + \zeta_2 \\ \frac{b_1\Omega_1 - b_2\Omega_2 + (a_1-a_2)\Omega_3}{a_1^2 - 2a_1a_2 + a_2^2 + \sigma_1^2 b_1^2 - 2b_1b_2 + \sigma_2^2 b_2^2} \end{pmatrix} \tag{4.28}$$

where Ω_1, Ω_2, and Ω_3 are expanded in the footnote.[11]

The possible FCEs under restricted optimal weights are similar to the OW case with a unique FCE that experience a bifurcation resulting in multiple FCEs for positive α above a threshold. Figure 10 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 as Figure 8, 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 10 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

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

restriction that the weights sum to one restricts the possible beliefs and forecasts from ever being equivalent to rational expectations.

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 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

$$T\begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \\ \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \mu + \alpha(\gamma_0 + \gamma_1 a_1 + \gamma_2 a_2) \\ \mu + \alpha(\gamma_0 + \gamma_1 a_1 + \gamma_2 a_2) \\ \alpha\gamma_1 b_1 + (\alpha\gamma_2 b_2 + \zeta_2)\frac{\sigma_{12}}{\sigma_1^2} + \zeta_1 \\ \alpha\gamma_2 b_2 + (\alpha\gamma_1 b_1 + \zeta_1)\frac{\sigma_{12}}{\sigma_2^2} + \zeta_2 \\ \mu - \gamma_1 a_1 - \gamma_2 a_2 + \alpha(\gamma_0 + \gamma_1 a_1 + \gamma_2 a_2) \\ \frac{\alpha\gamma_1 a_1^2 + b_1(\alpha\gamma_1\sigma_1^2 b_1 + (\alpha-1)\gamma_2\sigma_{12}b_2 + \sigma_1^2\zeta_1 + \sigma_{12}\zeta_2) + a_1((\alpha-1)(\gamma_0+\gamma_2 a_2)+\mu)}{a_1^2 + b_1\sigma_1^2} \\ \frac{\alpha\gamma_2 a_2^2 + b_2((\alpha-1)\sigma_{12}b_1 + \alpha\gamma_2\sigma_2 b_2 + \sigma_{12}\zeta_1 + \sigma_2\zeta_2) + a_2((\alpha-1)(\gamma_0+\gamma_1 a_1)+\mu)}{a_2^2 + \sigma_2^2 b_2^2} \end{pmatrix}. \quad (4.29)$$

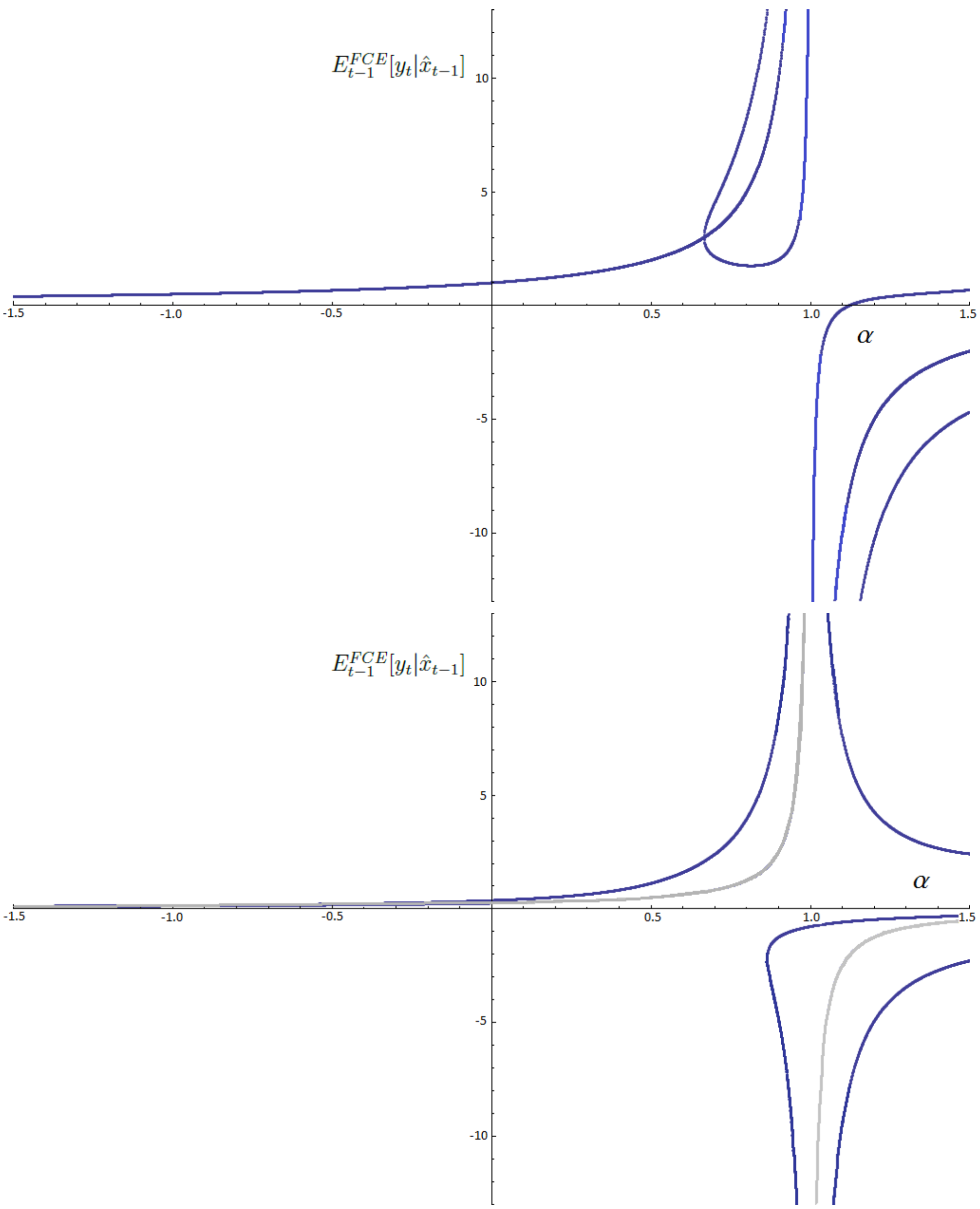

FIGURE 10. Plots of the ROWFCE forecasts. The top plot demonstrates multiple ROWFCEs for the same parameter values used in Figure (8), but with $\mu = 1$ and $\sigma_{12} = 0.1$. The bottom plot compares ROWFCE forecasts (black) to the REE forecast (gray) for different values of α and $\zeta = (.9 \; -.65)'$, $\sigma_1^2 = \sigma_2^2 = 1$, $\sigma_{12} = 0$, and $\mu = 0$.

Lemma 4: There exists a unique OWCFCE given by

$$a_1 = a_2 = \frac{\mu}{1-\alpha} \qquad \gamma_0 = \frac{(\sigma_1^2\sigma_2^2 - (\sigma_{12})^2)\zeta_1\zeta_2\mu}{(\alpha-1)(\sigma_1^2\zeta_1 + \sigma_{12}\zeta_2)(\sigma_{12}\zeta_1 + \sigma_2^2\zeta_2)}$$

$$b_1 = \frac{\sigma_1^2\zeta_1 + \sigma_{12}\zeta_2}{\sigma_1^2(1-\alpha)} \qquad \gamma_1 = \frac{\sigma_1^2\zeta_1}{\sigma_1^2\zeta_1 + \sigma_{12}\zeta_2}$$

$$b_2 = \frac{\sigma_{12}\zeta_1 + \sigma_2^2\zeta_2}{\sigma_2^2(1-\alpha)} \qquad \gamma_2 = \frac{\sigma_2^2\zeta_2}{\sigma_2^2\zeta_2 + \sigma_{12}\zeta_1}$$

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 8 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.

Results Summary

The findings of the section are summarized in Table 12. 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.

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

FCE Representative Results Summary

		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	$\sigma_{12} = 0$

TABLE 12. 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.

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.

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

$$\begin{aligned} \Theta_t &= \Theta_{t-1} + \kappa_t R_t^{-1} \mathbf{z_{t-1}} (\mathbf{y_t} - \mathbf{z'_{t-1}} \Theta_{t-1}) \\ R_t &= R_{t-1} + \kappa_t (\mathbf{z_{t-1}} \mathbf{z'_{t-1}} - R_{t-1}), \end{aligned} \tag{4.30}$$

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

$$\mathbf{z_t} = \begin{pmatrix} z_{1,t} & 0 & 0 \\ 0 & z_{2,t} & 0 \\ 0 & 0 & Y_{t+1} \end{pmatrix}, \tag{4.31}$$

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 (4.30) are equivalent to the FCEs determined in Section 3. The stability of these FCEs are determined by appealing to the E-stability principle.[12] The E-stability principle states that the stability of a rest point of (4.30) 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 (4.30) as t goes to infinity.[13] The resulting system is

$$\begin{aligned} \frac{d\Theta}{d\tau} &= R^{-1}E\mathbf{z}\mathbf{z}'(T(\Theta) - \Theta) \\ \frac{dR}{d\tau} &= E\mathbf{z}\mathbf{z}' - R. \end{aligned} \tag{4.32}$$

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

$$\frac{d\Theta}{d\tau} = T(\Theta) - \Theta, \tag{4.33}$$

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

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

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

fixed point governs the stability of (4.30). 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 benchmark result in the literature is that if agent consider a single correctly specified model that is of the form of the REE, then there exists a single equilibrium that is stable under learning if $\alpha < 1$.

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.

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 11 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 12 plots a grid of the parameter space 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

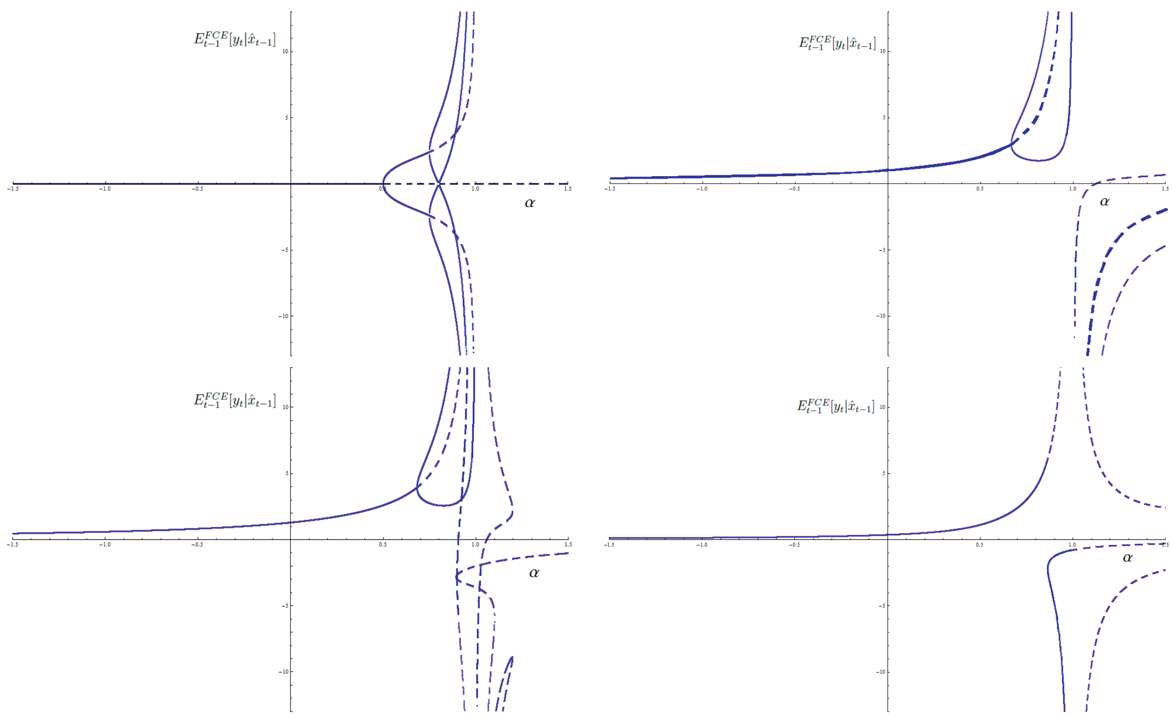

FIGURE 11. The E-stability of the OWFCEs and ROWFCEs. The $E_{t-1}^{FCE}[y_t|\hat{x}_{t-1}]$ under OW (left) and ROW (right) for different values of α. E-stability is indicated by solid lines and E-instability is indicated by dashed lines.

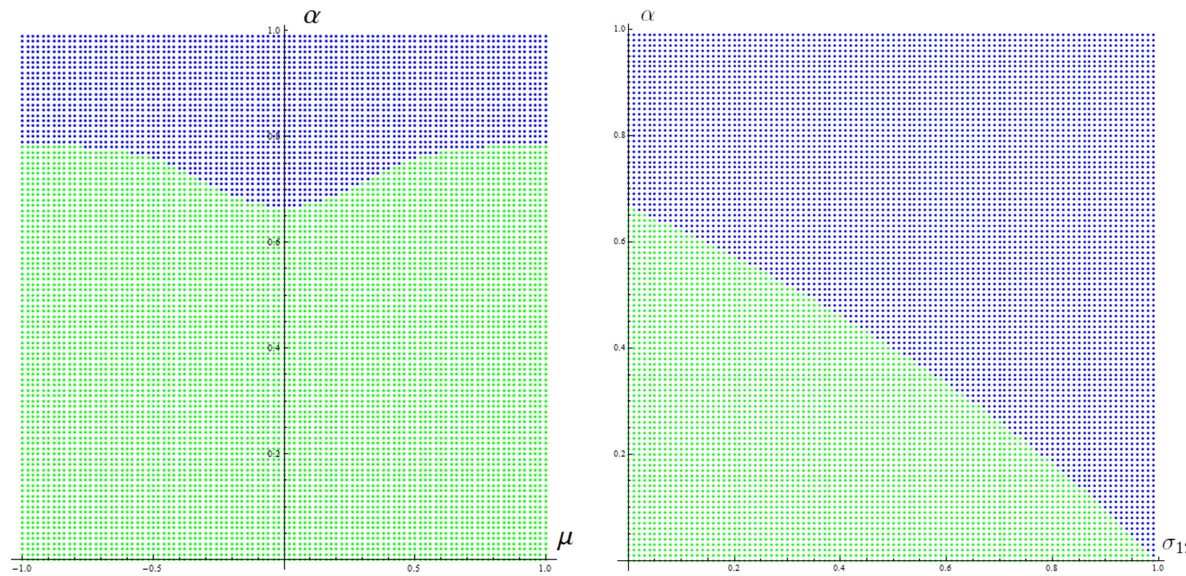

FIGURE 12. E-stability of the fundamental FCE under OWC for pairs of parameters (α, σ_{12}), and (α, μ). The light portion corresponds to the parameter space that is E-stable.

indicates E-instability. The figures demonstrates a non-linear relationship between E-stability and the parameters of the model.

Discussion

The E-stability results are summarized in Table 13. 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 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 condition 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.

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 13. 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 (12)). **The E-stability condition is a non-linear function of σ_{12}, ζ, and μ that ranges between $(0, 1)$.

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.

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.

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 economy may transition from one stable FCE to another.

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 (4.30).

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 $\kappa = 0.05$.[14] There are four E-stable OWFCEs under this parameterization of the model. The OWFCEs are

[14]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.

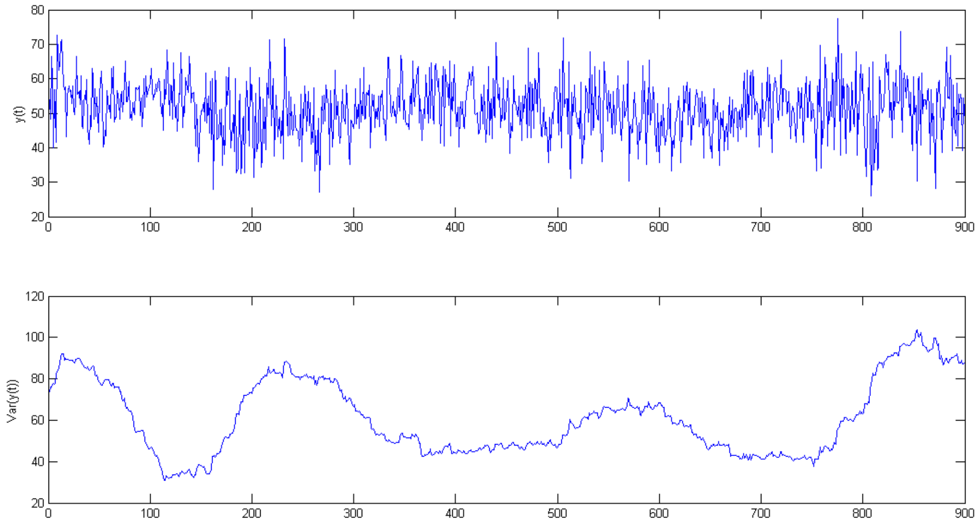

FIGURE 13. Time-varying volatility generated by OW forecast combination under constant gain learning with $v_t \sim N(0,1)$.

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

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 (13) 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

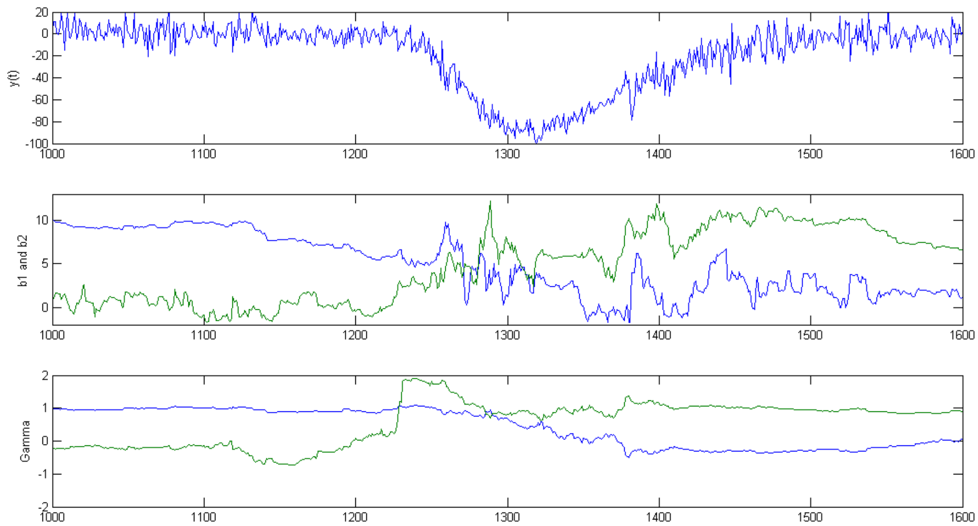

FIGURE 14. Time-varying volatility generated by OW forecast combination under constant gain learning with $v_t \sim N(0, 2)$.

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 (14) 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.

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 complicated 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.

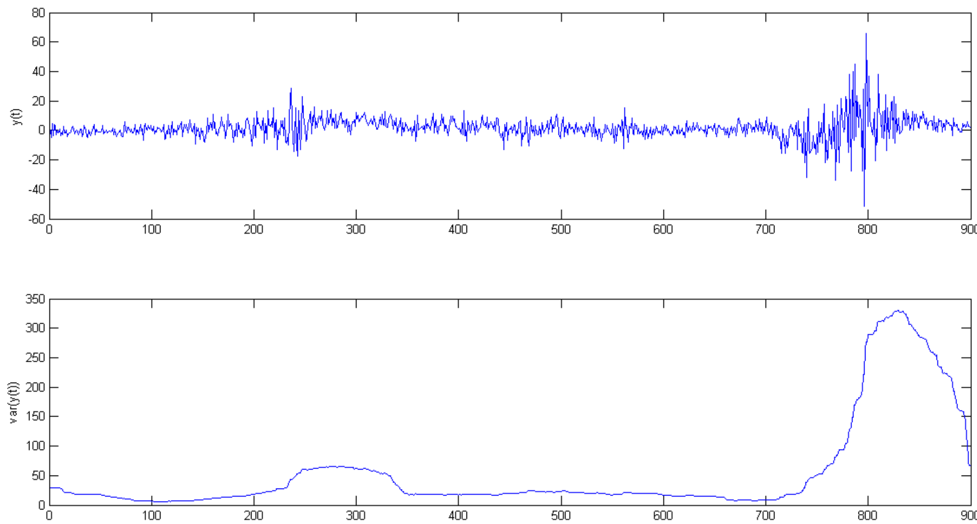

FIGURE 15. Time-varying volatility generated by OW forecast combination under constant gain learning with x_t following a VAR(1) process.

The exogenous shocks of the economy are

$$x_t = Ax_{t-1} + \epsilon_t, \tag{4.34}$$

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

$$A = \begin{pmatrix} .5 & .001 \\ .001 & .3 \end{pmatrix}, \Sigma_x = \begin{pmatrix} .2668 & .1190 \\ .1190 & 3.5166 \end{pmatrix}, \Sigma_\epsilon = \begin{pmatrix} .2 & .1 \\ .1 & 3.2 \end{pmatrix}, \tag{4.35}$$

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 (15) 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.

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 observed in the empirical literature has actually generated a puzzle.[15] 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. The out-of-sample subset is then recursively forecasted. 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

[15] 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.

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 found in past data can often fail to have the intended effect as described by Lucas (1976).

The Forecast Combination Equilibrium concept 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.

Forecast Combination and the Lucas Critique

This paper demonstrates that forecast combination strategies suffer from a 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 forecasts. This form of the Lucas Critique is perverse because instead of past correlations in the data being disconfirmed through poor forecasting performance, the agents receive confirming information for the erroneous correlations. The confirming information moves agents' beliefs away from the fundamentals of the economy and prevents agents from recognizing their mistakes.

Given that one of the goals of research in forecasting is to publish and 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 also 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 when widely employed. 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.

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 and 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 in a simple specification of the model under study, 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, some of these non-fundamental equilibria are found to be stable under learning.

In addition, 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.

Supplementary Materials and Proofs

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

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

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

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

Then simplify

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

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

$$-\alpha \sum_{i=1}^{k} \gamma_i u_1 \Sigma_z u_i' \phi_i + u_1 \Sigma_z u_1' \phi_1 = u_1 \Sigma_z B$$
$$...$$
$$-\alpha \sum_{i=1}^{k} \gamma_i u_k \Sigma_z u_i' \phi_i) + u_k \Sigma_z u_k' \phi_k = u_k \Sigma_z B.$$

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

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

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

$$\begin{aligned} \dot{x} &= Ax + f(x, y, \epsilon), \\ \dot{y} &= By + f(x, y, \epsilon), \qquad (x, y, \epsilon) \in \mathbb{R}^c \times \mathbb{R}^s \times \mathbb{R}, \\ \dot{\epsilon} &= 0, \end{aligned} \tag{4.36}$$

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

$$\dot{y} = D_x h(x, \epsilon)\dot{x} + D_\epsilon h(x, \epsilon)\dot{\epsilon} = Bh(x, \epsilon) + g(x, h(x, \epsilon), \epsilon). \tag{4.37}$$

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

$$f(0,0,0)=0 \qquad \frac{\partial f}{\partial x}(0,0,0)=0 \qquad \frac{\partial f}{\partial \epsilon}(0,0,0)=0$$
$$\frac{\partial^2 f}{\partial x^2}(0,0,0)=0 \quad \frac{\partial^2 f}{\partial x \partial \epsilon}(0,0,0)\neq 0 \qquad \frac{\partial^3 f}{\partial x^3}\neq 0.$$

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 (4.36) 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 (4.37) the center manifold must satisfy

$$D_u h(u,\epsilon)[Au + f(u, ha(u,\epsilon), hb(u,\epsilon),\epsilon)] - Bh(u,\epsilon) - g(u, ha(u,\epsilon), hb(u,\epsilon),\epsilon) = 0. \tag{4.38}$$

Equation (4.38) 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 7 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 (4.19) 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 (4.25) 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 $\alpha < \frac{1}{2}$.

CHAPTER V

CONCLUSION

The forecast selection problem affects to some degree every actor in the macroeconomy. Individuals as well as policymakers find themselves in an increasingly data rich environment, where the number of way to form a forecast of a variable of interest grows by the day. This dissertation provides two major theoretical and one major empirical contribution to our understanding of this problem and how its proposed solutions may affect economic dynamics and forecast efficiency.

The first theoretical contribution is that the common assumptions that underpin rational expectations are not sufficient to yield a rational forecast when agents explicitly face the model selection problem. The knowledge that other agents may deviate from rational expectations prohibits agents from individually coordinating without the assumption of shared information about individual choices. The lack of coordination may result in chaotic market dynamics and persistent forecast errors that follow from the miscoordination of the agents at the aggregate level.

The second theoretical contribution is that the use of forecast combination strategies to overcome the model selection problem by agents economy-wide may result in multiple equilibria and endogenous volatility. The use of optimal weighting strategies in this context results in self-fulfilling equilibria where erroneous beliefs about the forecasting efficiency of a model are reinforced by agents behaving as if the forecast is correct. The findings contribute insights into the possible expectational causes of economic volitility and raises questions about the theoretical underpinning of many forecast combination strategies if they were employed at the macroeconomic level.

The empirical contribution of the paper is to show the common explanation for the forecast combination puzzle can be used to create a combine forecast that does not exhibit the puzzle for inflation. The strategy also demonstrates the value of predicting the future efficiency of a forecast to rank and weight models to create combined forecasts. The strategy provides a promising way to create combined forecasts for many economic variables of interest.

The contributions of this dissertation also provide a number of interesting avenues for future research. Chapter 2 provides justification for heterogeneous agent models that use bounded rationality by showing the strong implicit assumptions that exist under traditional assumptions. The chapter justifies current research in the field and provides a new type of predictor that uses the rational assumptions, but does not necessarily yield a correct forecast. In addition, Chapters 3 and 4 present new forecasting and modeling concepts respectively, which have numerous unexplored applications and extensions that may improve our ability to forecast, while providing justifications based in macroeconomic theory.

REFERENCES CITED

[1] Aiolfi, M., Capistran, C., and A., T. Forecast combinations. *mimeo* (2010).

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

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

[4] Aumann, R., and Dreze, J. Rational expectations in games. *The American Economic Review 98*, 1 (2008), 72–86.

[5] Bates, J., and Granger, C. The combination of forecasts. *Operations Research* (1969), 451–468.

[6] Branch, W. Local convergence properties of a cobweb model with rationally heterogeneous expectations. *Journal of Economic Dynamics and Control 27*, 1 (2002), 63–85.

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

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

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

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

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

[12] Branch, W., and McGough, B. Multiple equilibria in heterogeneous expectations models. *The BE Journal of Macroeconomics 4*, 1 (2004), 12.

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

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

[15] Brock, W., Dindo, P., and Hommes, C. Adaptive rational equilibrium with forward looking agents. *International Journal of Economic Theory 2*, 3-4 (2006), 241–278.

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

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

[18] Brock, W., Hommes, C., and Wagener, F. More hedging instruments may destabilize markets. *Journal of Economic Dynamics and Control 33*, 11 (2009), 1912–1928.

[19] Brown, G. Iterative solution of games by fictitious play. *Activity analysis of production and allocation 13* (1951), 374–376.

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

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

[22] Clemen, R., and Winkler, R. Combining economic forecasts. *Journal of Business & Economic Statistics* (1986), 39–46.

[23] Clements, M., and Hendry, D. Intercept corrections and structural change. *Journal of Applied Econometrics 11*, 5 (1996), 475–494.

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

[25] Cogley, T., Primiceri, G., and Sargent, T. Inflation-gap persistence in the us. *American Economic Journal: Macroeconomics 2* (2010), 43–69.

[26] Croushore, D., and Stark, T. A real-time data set for macroeconomists. *Journal of Econometrics 105*, 1 (2001), 111–130.

[27] Diebold, F. Serial correlation and the combination of forecasts. *Journal of Business & Economic Statistics* (1988), 105–111.

[28] Diebold, F., and Lopez, J. Forecast evaluation and combination. *Handbook of Statistics 14* (1996), 241–68.

[29] Diebold, F. X., and Pauly, P. The use of prior information in forecast combination. *International Journal of Forecasting 6*, 4 (1990), 503–508.

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

[31] Evans, G., and Honkapohja, S. Learning as a rational foundation for macroeconomics and finance. In *Rethinking Expectations: The Way Forward for Macroeconomics*, R. Frydman and E. S. Phelps, Eds. Princeton Univ Pr, 2013.

[32] Evans, G., Honkapohja, S., Sargent, T., and Williams, N. Bayesian Model Averaging, Learning, and Model Selection. In *Macroeconomics at the Service of Public Policy*, T. Sargent and J. Vilmunen, Eds. Oxford University Press, 2013.

[33] Evans, G., and Ramey, G. Expectation calculation and macroeconomic dynamics. *The American Economic Review 82*, 1 (1992), 207–224.

[34] Evans, G., and Ramey, G. Calculation, adaptation and rational expectations. *Macroeconomic Dynamics 2*, 02 (1998), 156–182.

[35] Fudenberg, D., and Kreps, D. Learning mixed equilibria. *Games and Economic Behavior 5*, 3 (1993), 320–367.

[36] Gaunersdorfer, A. Endogenous fluctuations in a simple asset pricing model with heterogeneous agents. *Journal of Economic Dynamics and Control 24*, 5-7 (2000), 799–831.

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

[38] Granger, C., and Jeon, Y. Thick modeling. *Economic Modelling 21*, 2 (2004), 323–343.

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

[40] Guesnerie, R. An exploration of the eductive justifications of the rational-expectations hypothesis. *The American Economic Review 82*, 5 (1992), 1254–1278.

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

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

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

[44] Hommes, C., Sonnemans, J., Tuinstra, J., and Van de Velden, H. Coordination of expectations in asset pricing experiments. *Review of Financial Studies 18*, 3 (2005), 955.

[45] Hommes, C., Sonnemans, J., Tuinstra, J., and Van De Velden, H. Learning in cobweb experiments. *Macroeconomic Dynamics 11*, S1 (2007), 8–33.

[46] Hommes, C., and Wagener, F. Does eductive stability imply evolutionary stability? *Journal of Economic Behavior & Organization 75*, 1 (2010), 25–39.

[47] Jose, V., and Winkler, R. Simple robust averages of forecasts: Some empirical results. *International Journal of Forecasting 24*, 1 (2008), 163–169.

[48] Levin, A., and Piger, J. Is inflation persistence intrinsic in industrial economies? *mimeo* (2006).

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

[50] Lucas, R. Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy* (1976), vol. 1, pp. 19–46.

[51] Manski, C. F., McFadden, D., et al. *Structural analysis of discrete data with econometric applications*. MIT Press Cambridge, MA, 1981.

[52] Marcellino, M., Stock, J., and Watson, M. A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series. *Journal of Econometrics 135*, 1-2 (2006), 499–526.

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

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

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

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

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

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

[59] Stock, J., and Watson, M. A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. *in R.F. Engle and H. White, eds,. Festschrift in Honour of Clive Granger.* (2001), 1–44.

[60] Stock, J., and Watson, M. Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature 41*, 3 (2003), 788–829.

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

[62] Stock, J., and Watson, M. Forecasting with many predictors. *Handbook of economic forecasting 1* (2006), 515–554.

[63] Stock, J., and Watson, M. Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking 39* (2007), 3–33.

[64] Stock, J., and Watson, M. Modeling inflation after the crisis. *mimeo* (2010).

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

[66] Thorlund-Petersen, L. Iterative computation of cournot equilibrium. *Games and Economic Behavior 2*, 1 (1990), 61–75.

[67] Timmermann, A. Forecast combinations. *Handbook of economic forecasting 1* (2006), 135–196.

[68] Turner, D. The role of judgment in macroeconomic forecasting. *Journal of Forecasting 9*, 4 (1990), 315–345.

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

[70] Wallis, K., and Whitley, J. Sources of error in forecasts and expectations: Uk economic models, 1984-88. *Journal of Forecasting 10*, 3 (1991), 231–253.

[71] Waters, G. Chaos in the cobweb model with a new learning dynamic. *Journal of Economic Dynamics and Control 33*, 6 (2009), 1201–1216.

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

[73] Young, P. *Individual strategy and social structure: An evolutionary theory of institutions*. Princeton University Press, 2001.